25 Feb 2019

Molecular Dynamics Simulations of Substrate Release from Trypanosoma Cruzi UDP-Galactopyranose Mutase

Rodrigo Cossio-Pérez

Gustavo Pierdominici-Sottile

Pablo Sobrado
Missouri University of Science and Technology, psobrado@mst.edu

Juliana Palma

Recommended Citation

R. Cossio-Pérez et al., "Molecular Dynamics Simulations of Substrate Release from Trypanosoma Cruzi UDP-Galactopyranose Mutase," *Journal of Chemical Information and Modeling*, vol. 59, no. 2, pp. 809 - 817, American Chemical Society, Feb 2019.
The definitive version is available at https://doi.org/10.1021/acs.jcim.8b00675

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

Cite This: J. Chem. Inf. Model. 2019, 59, 809−817

pubs.acs.org/jcim

Molecular Dynamics Simulations of Substrate Release from *Trypanosoma cruzi* UDP-Galactopyranose Mutase

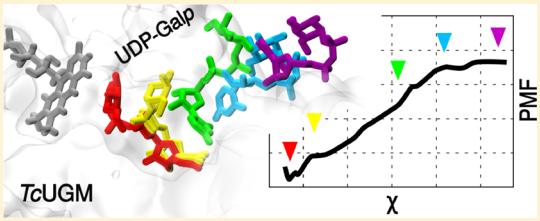
Rodrigo Cossio-Pérez,[†] Gustavo Pierdominici-Sottile,[†] Pablo Sobrado,[‡] and Juliana Palma*[,†]

[†]Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, CONICET, Bernal, Buenos Aires B1876BXD, Argentina

[‡]Department of Biochemistry, Virginia Tech, Blacksburg, Virginia 24060, United States

Supporting Information

ABSTRACT: The enzyme UDP-galactopyranose mutase (UGM) represents a promising drug target for the treatment of infections with *Trypanosoma cruzi*. We have computed the Potential of Mean Force for the release of UDP-galactopyranose from UGM, using Umbrella Sampling simulations. The simulations revealed the conformational changes that both substrate and enzyme undergo during the process. It was determined that the galactopyranose portion of the substrate is highly mobile and that the opening/closing of the active site occurs in stages. Previously uncharacterized interactions with highly conserved residues were also identified. These findings provide new pieces of information that contribute to the rational design of drugs against *T. cruzi*.



INTRODUCTION

Trypanosoma cruzi is the protozoan parasite that causes Chagas disease. This disease affects 8 million people worldwide and leads to about 10,000 annual deaths.[1] Nitroheterocyclic drugs are currently used to treat this illness but have considerable drawbacks: they are expensive and present severe side effects, and drug resistance has been reported.[2,3] Galactofuranose (Gal*f*) is a cyclic sugar absent in humans but found in *T. cruzi* glycocalyx, as part of glycoproteins and glycoinositolphospholipids. Glycoproteins are involved in the attachment and invasion of the host cells,[4−8] while glycoinositolphospholipids help to suppress its immune system.[9,10] These actions are recognized as key factors for the pathogenesis and permanence of the parasite in the human body.[11−13] Because of this, it is believed that inhibiting Gal*f* formation would be a target for chemical intervention.[11,14]

The biosynthesis of Gal*f* begins with the isomerization of UDP-galactopyranose (UDP-Gal*p*) into UDP-Gal*f*, which is catalyzed by the enzyme UDP-galactopyranose mutase (UGM).[11] Since UGM is not present in humans, it has been considered a valuable drug target, and, accordingly, it has been thoroughly studied.[14−17] Experiments revealed that the UGM of *T. cruzi* (*Tc*UGM) functions as a monomer and contains the FAD cofactor. FAD can either be reduced or oxidized, but the reduced state is required for enzyme activity.[18] Structures of the enzyme with the cofactor in both redox states have been solved by X-ray crystallography.[19] Molecular dynamics (MD) simulations were used to characterize the conformations of the enzyme bound to different ligands as well as those corresponding to the unbound state.[20] In addition, experimental and computational studies were able to identify intermediates of the catalytic mechanism.[18,21,22] A comparative analysis of the crystal structures of different UGMs revealed residues involved in substrate recognition.[15,19] Finally, a mutagenesis and structural study of eukaryotic UGMs unveiled crucial information about the conformational changes that take place when the active site closes.[23]

In this Article, we present the results of a computational study focused on the release of UDP-Gal*p* from *Tc*UGM. The study is based on the calculation of the Potential of Mean Force (PMF) using Umbrella Sampling (US) simulations. Because the simulations were performed under equilibrium conditions, the profile equally describes the binding of UDP-Gal*p* to *Tc*UGM. Analysis of the results allowed us to identify relevant interactions between the substrate and the enzyme, which have not been hitherto observed. Additionally, we were able to characterize the conformational changes that take place in the enzyme−substrate complex during the binding/release process. The outcomes of this Article contribute to the general knowledge of eukaryotic UGMs and should assist further studies aimed at providing inhibitors for *Tc*UGM.

METHODS

System Setup. The crystal structure of reduced *Tc*UGM was taken from the Protein Data Bank[24] (PDB) code 4DSH.[19] The FAD cofactor was set in its reduced monoprotonated form, as suggested by experiments performed on other UGMs.[17] The coordinates of the UDP-Gal*p* atoms were generated from the cocrystallized UDP molecule, using the structure of UDP-Gal*p* bound to *Aspergillus fumigatus* UGM (*Af*UGM) as a template.[25] All crystallographic water molecules were kept in the model. Standard protonation states were assigned to titratable residues. Histidines residues were

Received: October 1, 2018
Published: January 4, 2019

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

protonated in the N-epsilon atom. The only exception was done for His62 which was set in the doubly protonated state, as suggested by the analysis of the crystal structure.[19] The complex was neutralized with Na^+ and solvated with water molecules in a truncated octahedron box of about 560 nm^3. We used the AMBER ff99SB force field for the protein,[26] the TIP3P model for the water molecules, and the GAFF force field for UDP-Gal*p* and FAD.[27] The partial charges of the last two molecules were computed using the RESP methodology at the B3LYP/6-31G* level of theory.[28]

The system was first minimized with harmonic restraints to the backbone atoms of the protein. The force constant was set to 20.0 kcal/mol/Å^2. Then, the system was heated at constant pressure from 0.0 to 310.0 K in 500.0 ps using the Langevin thermostat with a collision frequency of 1.0 ps^{-1} and the Berendsen barostat with a relaxation time of 1.0 ps. For the heating stage, the restraint constant was reduced to 6.0 kcal/mol/Å^2, and the time-step was set to 1.0 fs. The Particle Mesh Edwald method with a cutoff of 10.0 Å was used to calculate long-range electrostatic interactions. Afterward, we introduced the SHAKE algorithm to fix the lengths of bonds involving hydrogen atoms. This allowed us to increase the time-step to 2.0 fs. Next, we gradually reduced the restraint constants using four simulations of 250.0 ps with constants of 6.0, 4.0, 2.0, and 0.4 kcal/mol/Å^2. Finally, for the equilibration stage, we removed all the restraints and simulated the system for 10.0 ns. The stability of the model obtained with this protocol was assessed by evaluating different properties. We computed the RMSD of the C_α atoms of the enzyme with respect to its initial structure. This parameter reached a plateau value of 1.18 ± 0.08 Å after ∼3 ns of the equilibration stage. The radius of gyration of the complex hardly varied in this period, with an average of 23.25 ± 0.74 Å. In addition, we calculated the Connolly molecular surface of the complex employing a spherical probe with a radius of 1.4 Å. This calculation afforded a surface of 19107 ± 641 Å^2. We note that the substrate remained firmly attached to the active site along the whole equilibration period, with its galactose moiety in strong interaction with the isoalloxazine ring of the cofactor. The parameters employed in the production runs were the same as those implemented in the equilibration period. The preparation of the system was performed with TLEAP. The simulations were run with PMEMD. Both modules belong to the AMBER16 package.[29]

Umbrella Sampling Simulations. The US technique was utilized to force the unbinding of UDP-Gal*p* from *Tc*UGM.[30] The reaction coordinate of the process (χ) was defined as the distance between two centers of mass (COMs). One of them contained the heavy atoms of UDP-Gal*p*, while the other COM involved the C_α atoms of residues 293−297, 350−352, and 390−392. These residues belong to a stable β-sheet that forms the rear wall of the binding pocket. Figure S1 of the Supporting Information illustrates the definition of the reaction coordinate. The value of χ was sequentially increased from 15.3 to 36.3 Å, with a spacing of 0.1 Å. The force constant of the harmonic bias potential was set to 500.0 kcal/mol/Å^2. We checked that, with this spacing and force constant, histograms of the reaction coordinate corresponding to adjacent windows have a good overlap. The last structure of a simulation was employed as the initial structure of the following. An equilibration period of 1.0 ns was allowed within each US window before starting to collect production data.

The amount of sampling required to achieve convergence variates with the value of χ. Therefore, we divided the whole χ-range into three segments named s_1, s_2, and s_3 and employed different simulation times for each of them. The ranges of the segments were 15.3−19.5 Å for s_1, 19.6−27.1 Å for s_2, and 27.2−36.3 Å for s_3. The simulations lasted 35.0 ns in s_1, 20.0 ns in s_2, and 15.0 ns in s_3. The total simulation time adds up to 4.4 μs. In all cases, snapshots were taken every 10.0 ps, and the χ value was recorded every 4.0 fs. To compute the PMF we used the last 20.0 ns of simulations from s_1 and the last 10.0 ns of simulations from s_2 and s_3. It should be noted that the computation of the PMF employs equilibrium conditions for each value of χ. Therefore, the profile describes both binding and release processes.

Potential of Mean Force. The Weighted Histogram Analysis Method (WHAM)[31] and the Dynamic Histogram Analysis Method (DHAM)[32] were implemented to compute the PMF. A good agreement between the profiles obtained with both approaches is considered as a test of the correctness of the parameters used in each calculation.[32] WHAM computations were performed with the program developed by Alan Grossfield.[33] We found that 100000 iterations were enough to achieve convergence. DHAM computations were carried out with our own FORTRAN code. In the two cases, data were binned from 15.3 to 36.3 Å using a spacing of 0.02 Å between bin centers. To test the consistency of the PMF, two additional assessments were carried out. Both were applied along with the simulations, to decide whether further calculations were needed or not. The first approach measures consistency between adjacent simulations and involves the observation of the function[34]

$$f_i(\chi) = \frac{1}{\beta} \ln\left(\frac{P_{i+1}(\chi)}{P_i(\chi)}\right) + \Delta U_i(\chi) \tag{1}$$

where $\beta = 1/k_B T$, with k_B being the Boltzmann constant and T = 310.0 K. $P_i(\chi)$ and $P_{i+1}(\chi)$ are the probability distributions for χ, calculated from the samples of simulations i and $i + 1$, respectively. $\Delta U_i(\chi)$ represents the difference between the bias potentials employed in the two simulations for a given value of χ. The function $f_i(\chi)$ should be constant between the centers of simulations i and $i + 1$.[34] Accordingly, we calculated $f_i(\chi)$ for all pairs of adjacent windows and visually verified that it was almost constant in the required range. The second approach assesses the consistency between the distributions actually observed in the US simulations and the biased probabilities calculated from results of WHAM or DHAM. It employs the Kullback−Leibler divergence[35] which measures the distance between two probability densities P_0 and P_1 as

$$D(P_0, P_1) = \sum_{k=1}^{N} P_0(\chi_k) \ln \frac{P_0(\chi_k)}{P_1(\chi_k)} \tag{2}$$

where N represents the number of bins employed in a discretized representation of the probability densities, while χ_k is the value of the random variable at the center of bin k. In our case, we compared the probability distribution of χ observed in the i-th US simulation, $P_i(\chi)$, with the biased distribution computed from the result of WHAM, $\mu_i(\chi)$. To consider both functions on the same footing we evaluated the symmetric Kullback−Leibler (sKL) divergence

$$S_i = \frac{1}{2} D(P_i, \mu_i) + \frac{1}{2} D(\mu_i, P_i) \tag{3}$$

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

The lower the value of the sKL-divergence, the better the agreement.

Finally, to estimate the statistical uncertainty of the results, we divided the whole batch of data into 10 sets of equal size and computed a PMF with the data of each set. The profiles so obtained were shifted so that the free energy at χ = 15.6 Å was zero in all cases. Afterward, we calculated the standard deviation of the 10 PMFs at several other points. These standard deviations were employed as a measure of the statistical uncertainty of the global PMF.[36]

Analysis. The analysis of the data collected along the simulations was performed either with CPPTRAJ, with other auxiliary programs of the AMBERTOOLS16 package, or with our own PYTHON and FORTRAN codes. For the visualization of the model, we employed VMD 1.9.3.[37] The graphical representation of the data was done with GNUPLOT 5.0, with the assistance of program INKSCAPE 0.92 for adding explanatory details and for assembling more than one image in the same picture.

We counted the number of water molecules within the active site pocket, at different stages of the unbinding process. For this purpose we applied the following algorithm. First, we determined the middle point between the N_5 atom of FAD and the C_α atom of Phe102. We considered this point as an estimation of the center of the cavity. Then, we selected the water molecules located within a sphere of r = 8.0 Å centered at this point. Next, we added to the selected set those molecules that fulfilled the following requirements: 1) they were separated by less than 4.4 Å from at least two molecules already included; 2) their distances to the C_α atom of Thr295 were shorter than 27.0 Å. Since Thr295 is situated in the rear wall of the cavity, the latter action avoids including molecules that lie outside it. The last step is applied iteratively until convergence is reached. Besides, we determined the locations where the probability of finding water molecules is larger than 60%. To this end, we employed the VolMap 1.1 extension of VMD 1.9.3.[37]

As detailed below, the analysis of the US simulations unveiled important interactions between the substrate and numerous active site residues. We accordingly studied the degree of conservation of those residues among eukaryotic UGMs. We collected 124 protein sequences from the UniProt database[38] and performed a multiple alignment using the program MUSCLE with default parameters.[39] Then, we built a sequence logo employing the WebLogo server.[40]

RESULTS AND DISCUSSION

Potential of Mean Force. The computed PMF is depicted in Figure 1. The results of the calculations carried out to analyze its accuracy are shown in the Supporting Information. Figure S2 compares the PMFs obtained with WHAM and DHAM; Figure S3 shows typical plots of $f_i(\chi)$ (eq 1), and Figure S4 illustrates the variation of the symmetric sKL-divergence (eq 3) along the whole range of the reaction coordinate. All the assessments afford satisfactory results.

It can be readily appreciated that the PMF presents regions of very different slopes. It increases by ∼5 kcal/mol between the equilibrium position, located at χ = 15.6 Å, and χ = ∼18 Å. This is followed by a short plateau between ∼18 Å and ∼20 Å. Beyond that, there is a sustained increase in the free energy, while χ variates between ∼20 Å and ∼30 Å. Finally, when the substrate is leaving the enzyme, the profile reaches a new plateau. The total free energy change from the equilibrium position to the exit is 21.4 ± 3.0 kcal/mol.

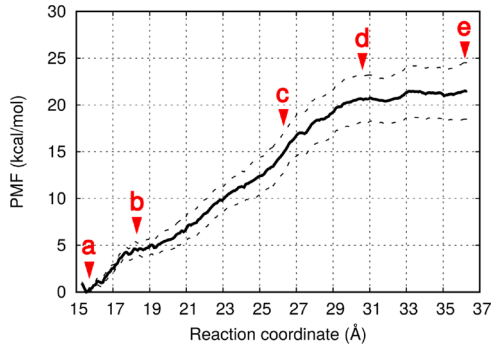


Figure 1. Solid line indicates the PMF for the binding/unbinding process of UDP-Gal*p* in *Tc*UGM. The dashed lines show the limits of the statistical uncertainty. The red triangles with labels indicate the χ values of the typical structures presented in Figure 2.

We note that the barrier for the release of the substrate is slightly smaller than the one reported for the rate-determining step of the isomerization reaction (23.4 ± 0.4 kcal/mol).[22] Assuming that the release of UDP-Gal*p* has a PMF similar to that of UDP-Gal*f*, this finding implies that the chemical reaction constitutes the rate-determining step of the whole process, instead of product release. This is in line with the experiments of Oppenheimer et al.[18] However, the statistical uncertainty of the present calculations is similar to the difference between the barriers of the two stages. Therefore, our results provide support to the previous experiments but are not conclusive. On the other hand, the fact that the barriers are close to each other suggests that small variations in the chemical groups that compose the substrate and enzyme active site can switch the rate-determining step between chemical reaction and product release. This phenomenon has been recently observed in *Af*UGM.[41]

Substrate Displacement. UDP-Gal*p* is a large molecule with several chemical groups: Gal*p*, diphosphate, ribose, and uracil. The analysis of snapshots collected along the complete range of the reaction coordinate shows that these groups move very differently during the binding/unbinding process. Figure 2 depicts representative conformations of the substrate at alternative values of χ. Different colors have been used for each chemical moiety to highlight their individual behavior. At the minimum of the PMF, the substrate presents an extended linear conformation with Gal*p* close to the cofactor and the uracil group in the opposite extreme. During the first stages of the unbinding process, Gal*p* and diphosphate perform large displacements while ribose and uracil hardly move. This indicates that the last two groups have stronger interactions with the enzyme binding pocket. We note that uracil and ribose lie on a cavity of the binding pocket, interacting with the so-called uridine wall pattern formed by Phe152, Tyr156, Asn157, and Trp161.[15] Additionally, the uracil moiety interacts with Phe102 and Gln103.

The displacements observed between 15.6 Å and ∼18 Å can be roughly described as a bending motion of the substrate, caused by small rotations around the bonds of the diphosphate group (Figure 2, panels (a) → (b)). After that, for χ between ∼18 Å and ∼30 Å, there is a global rotation of the substrate around an axis that passes near the anchoring point of the uracil group (Figure 2, panels (b) → (c)). This movement leaves Gal*p* and the diphosphate close to the exit of the binding

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

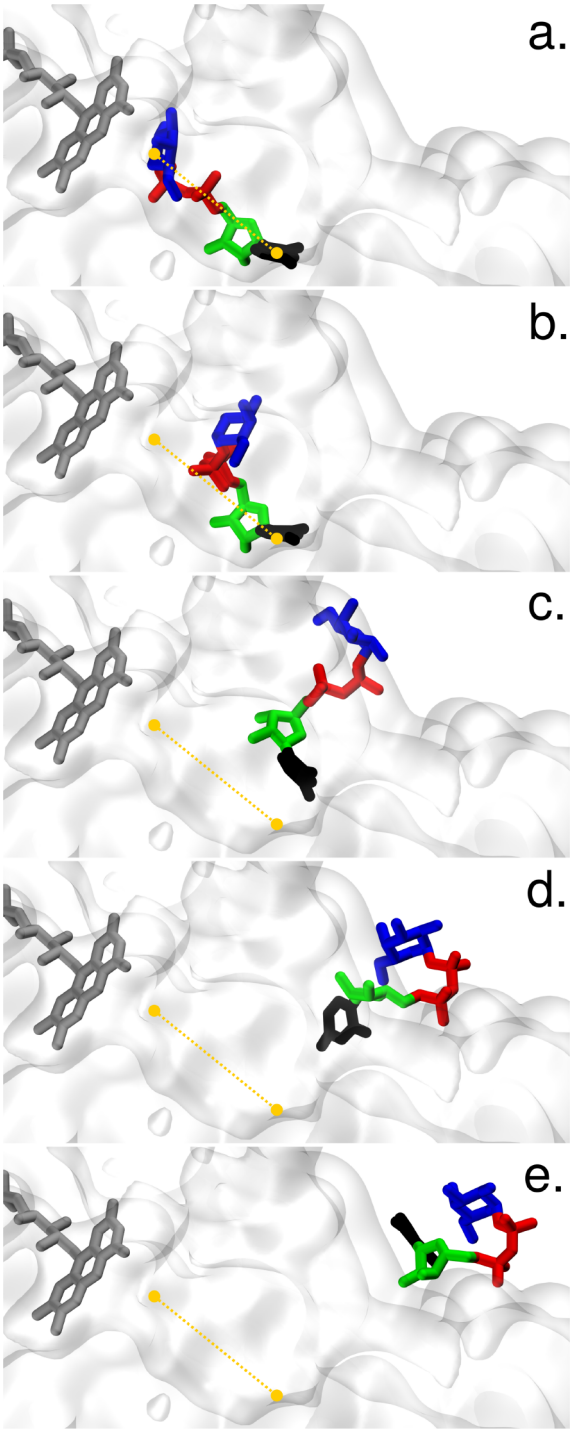


Figure 2. Typical conformations adopted by the substrate along the binding/unbinding process. The scheme of UDP-Gal*p* identifies each chemical moiety with a characteristic color: Gal*p* (blue), diphosphate (red), ribose (green), and uracil (black). Reference marks (yellow) were added to help visual recognition of the conformational changes.

pocket, while uracil and ribose make smaller displacements. However, they have already lost their interactions with the residues of the uridine wall. Beyond this point, uracil and ribose perform larger translations and the substrate bends even further, reducing the distance between its ends (Figure 2, panels (d) and (e)). Finally, only the ribose and phosphate group interact with the enzyme. The phosphate is the last group to leave the enzyme. This observation agrees with the proposal of Da Fonseca et al., who suggested that the interactions between substrate and enzyme start with the binding of the phosphate group.[23] At χ values greater than ∼30 Å the substrate gets exposed to the solvent adopting a large variety of conformations. Movie S1 in the Supporting Information shows the conformations of the substrate along the substrate-release process.

Gal*p*-FAD Interactions. We have previously shown that Gal*p* is the group that makes the largest displacements throughout the substrate release process. It should also be noted that this group adopts various conformations even when the system is close to the minimum of the PMF. To quantify this observation, we collected substrate structures from US simulations with χ between 15.6 and 16.4 Å and aligned them to the structure corresponding to the minimum of the PMF. From this set, we calculated the root-mean-squared-fluctuation (RMSF) of each moiety. The RMSFs were 1.5 Å for Gal*p*, 0.9 Å for diphosphate, 0.7 Å for ribose, and 0.6 Å for uracil, confirming that Gal*p* presents a much larger mobility than the other groups of the substrate.

To further characterize the conformations of Gal*p* in this region we computed its Cremer-Pople angles.[42,43] In what follows, atoms will be named according to the guidelines provided in ref 44. The analysis of the Cremer-Pople angles revealed that Gal*p* always adopts a $^4C_1$ sugar ring conformation. However, variations in the torsional angles around the $C_1'-O_3B$, O_3B-PB, and $PA-O_5D$ bonds modify the way in which Gal*p* interacts with the FAD cofactor. Each of these alternative conformations presents a characteristic H-bond between O_4 of FAD and one of the −OH groups of Gal*p*. Panel (a) of Figure 3 presents the probabilities of the three relevant H-bonds as a function of χ, while panels (b), (c), and (d) show the alternative conformations of Gal*p*. For $\chi \sim 15.6$ Å, the atom involved in the H-bond with FAD is O_3' (Figure 3b). When $\chi \sim 16.4$ Å, the atom involved is O_6' (Figure 3d). Between these two values, the sugar adopts an intermediate conformation in which O_4' of Gal*p* interacts with O_4 of FAD (Figure 3c). It should also be noted that, in going from $\chi \sim 15.6$ Å to 16.4 Å, the plane formed by the sugar ring rotates with respect to the isoalloxazine ring. At the shortest values of χ both planes are almost parallel, forming an angle of ∼10°. At the largest values of χ the angle between the planes is ∼50°.

The existence of three alternative H-bonds between Gal*p* and the isoalloxazine ring of the cofactor was corroborated by QM/MM computations that included both moieties in the QM subsystem. The level of theory employed was scc-DFTB.[45] Technical details regarding the implementation of these calculations can be found in the Supporting Information. Table S1 of the Supporting Information presents a comparison between the classical and QM/MM results. In agreement with the classic results, QM/MM simulations show that the prevailing H-bonds at χ = 15.6, 15.9, and 16.3 Å are those involving O_3', O_4', and O_6', respectively. There are also minor differences between the two calculations. QM/MM simulations afford somewhat larger probabilities for the H-bonds between O_4' and FAD at χ = 15.6 Å and between O_6' and FAD at χ = 15.9 Å.

The different conformations achieved by Gal*p* in its interaction with FAD are consistent with the excess of void volume around the sugar, already described for UGMs.[46] This excess of volume allows the enzyme to accommodate alternative saccharides. Moreover, the conformations of Gal*p* described here are similar to those reported for *Af*UGM with

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

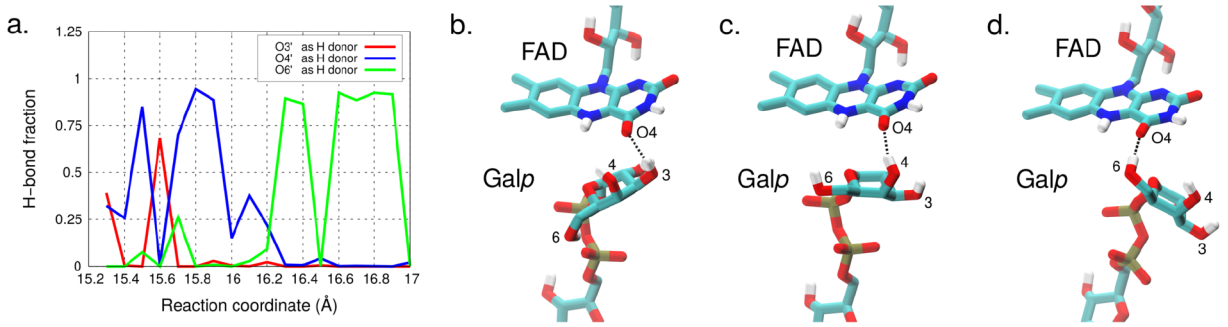


Figure 3. (a) Probabilities for the H-bonds between atom O_4 of FAD and different −OH groups of Gal*p*, as a function of χ. (b), (c), and (d) alternative conformations of the substrate for χ between 15.6 and 16.4 Å. H-bonds are displayed with dashed lines.

UDP-arabinopyranose and UDP-Galp[41] but different than the one adopted by the nonreactive sugar UDP-glucopyranose in the UGM of *Klebsiella pneumoniae*. The main difference is that, in the present case, the distance between N_5 of FAD and C_1' of the sugar is significantly shorter. In addition, the UDP group always adopts an orientation appropriate to act as leaving group in the first step of the isomerization reaction.

Enzyme Conformations. Molecular dynamics simulations of substrate free *Tc*UGM demonstrated that the volume of its active site changes dramatically upon substrate binding.[20] A similar observation was performed after comparing the crystal structures of closed and open *Af*UGM.[47] This volume-change is accompanied by the movement of two flaps that block the entrance of the active site when the substrate is bound.[20,47] The 180s flap contains residues 173 to 181, and the 200s flap contains residues 195 to 203. The information collected from the US simulations allowed us to analyze the conformational changes occurring in the active site of *Tc*UGM along the complete binding process. In this section, we present the results of this analysis and compare our findings with previous studies that provided valuable and closely related insights. We note that another mobile flap was identified in previous MD simulations of the substrate-free enzyme.[20] It contains residues 461 to 471. We determined that the displacements of this flap bear no relation to the volume of the active site. Accordingly, we did not analyze its behavior any further.

In order to follow the closure of the mobile flaps upon substrate binding, we measured the distance between the C_αs of their central residues. We used Ala178 as the central residue of the 180s flap and Pro200 as that of the 200s flap. The distance between their C_αs is denoted as $d_{178-200}$. In addition, we measured the pocket volume using the POVME software.[48] Thus, we noted that the two variables are highly correlated and therefore provide redundant information. The high correlation between them can be appreciated in the inset of Figure 4a. Their Pearson correlation coefficient is 0.86. Since the measurement of $d_{178-200}$ is simpler than that of the pocket volume, we used it to further characterize the state of the binding pocket.

Figure 4a displays the variations of $d_{178-200}$ along the reaction coordinate. For the analysis we divide the curve into five regions, from (**I**) to (**V**), spanning the following successive ranges of χ: [15.3 Å−18 Å), [18 Å−20 Å), [20 Å−30 Å), [30 Å−33 Å), and χ > 33 Å. In regions (**I**), (**III**), and (**V**), $d_{178-200}$ performs small fluctuations around a nearly constant value. In the other two regions, it presents abrupt jumps that reflect significant conformational changes in the enzyme. In region (**I**), $d_{178-200}$ is ∼10 Å, while flaps 180s and 200s are close to

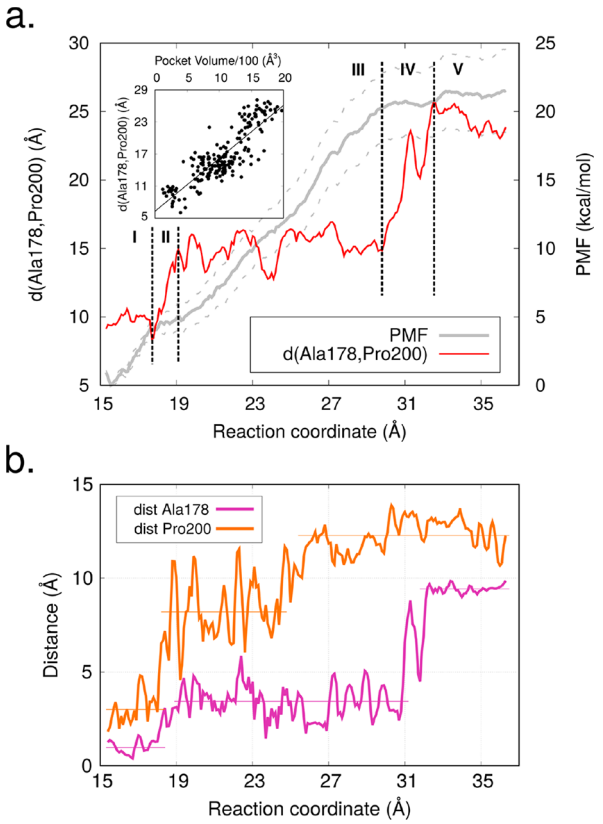


Figure 4. (a) Distance between the C_αs of residues Ala178 and Pro200, $d_{178-200}$, as a function of χ. For comparison, the PMF is shown in gray. The inset shows the correlation between $d_{178-200}$ and the pocket volume. (b) Distance between the average position of the C_αs of residues Ala178 and Pro200 at different values of χ and that corresponding to that of the minimum of the PMF. Distances were smoothed with the Savitzky−Golay filter.

each other, covering the entrance of the pocket. In region (**V**), $d_{178-200}$ is ∼24 Å, the flaps are separated from each other, and the enzyme is widely open. We also note that, in this region, the substrate gets exposed to the solution. Finally, in region (**III**), $d_{178-200}$ adopts an intermediate value of ∼15 Å, certainly different from the distances observed in regions (**I**) and (**V**). We, therefore, conclude that the enzyme achieves a semiopen state in this range. This state can be clearly distinguished from the open and closed states that have been characterized previously.[20]

The existence of a semiopen state was put forward by Da Fonseca et al.[23] The authors obtained the crystal structure of

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

the N207A mutant of *Af*UGM (N201A in *Tc*UGM) bound to UDP. This mutation avoids the formation of an H-bond between Asn207 and Arg91 (Asn201-Arg87 in *Tc*UGM) that stabilizes the closed state of the enzyme. They observed that the 180s flap of the mutant achieved the closed conformation that is typical of the substrate−enzyme complex. However, the 200s flap adopted an intermediate position in which residues 206−207 (200−201 in *Tc*UGM) were poised 5 Å from the closed conformation and 10 Å from the open conformation. In addition, based on the analysis of various mutants, these authors proposed that the closing of the enzyme takes place according to the following sequence of events. First, the uridine and phosphate groups dock into the active site. Next, the 180s flap closes, while the 200s flap adopts an intermediate state. Finally, the 200s flap completes the closing.

To analyze the individual behavior of flaps 180s and 200s, we computed the average position of the C_αs of Ala178 and Pro200 for each US window. Then, we measured the distance between these positions and those achieved by the same atoms at the minimum of the PMF. Figure 4b depicts the variations of these distances along the whole range of the reaction coordinate. Several features in the figure are worth mentioning. The most striking one is that the displacement of Ala178 closely resembles that of $d_{178-200}$, while that of the 200s flap differs significantly (see Figure 4a). In parallel, we found that the correlation coefficient between the displacement of the 180s flap and the pocket volume is 0.82, while that of the 200s flap is merely 0.68. Therefore, the changes of $d_{178-200}$ and pocket volume are mostly caused by the displacement of the 180s flap, while the 200s flap provides a minor contribution. Another aspect worth mentioning is that the two flaps go from the open to the closed conformation through an intermediate stage. However, the transitions between stages are not always simultaneous. Following the curves in the sense of the binding process, one notes that the 180s flap moves toward the intermediate position much earlier than the 200s flap ($\chi \sim 30-33$ vs $\chi \sim 24-25$). This agrees with the proposal of Da Fonseca et al.[23] in regard to the order of both movements. However, there is a qualitative disagreement because these transitions lead to an intermediate position instead of the final position of each flap in the closed state. Finally, the displacement from the intermediate position to that of the closed state is performed simultaneously by the two flaps at $\chi \sim 18-20$ Å.

It is interesting to compare the changes in the PMF and $d_{178-200}$ along the reaction coordinate. In regions (**I**) and (**III**), where the interflaps distance presents small fluctuations, the PMF increases steadily (Figure 4a). On the contrary in regions (**II**) and (**IV**), where the interflaps distance presents large jumps, the PMF is flat. Region (**V**) is set aside in the comparison because at that point the substrate is nearly outside the enzyme. The behavior observed in regions (**I**) and (**III**) can be explained intuitively. Trying to move the substrate out of the enzyme, while the pocket volume remains the same, produces clashes between the substrate and the walls of the cavity. Simultaneously, these displacements break the attractive interactions that keep the substrate bound to the enzyme. All these actions tend to increase the energy of the system, while they should not significantly affect its entropy. Accordingly, the PMF increases. The behavior observed in regions (**II**) and (**IV**) is more curious and difficult to explain. It could be speculated that the transition from the closed to the semiopen state, or from the semiopen to the open state, should be accompanied by an increase in both the potential energy and the entropy of the system. The fact that the PMF remains virtually constant in these regions seems to suggest that the two effects cancel each other out. However, this is just a preliminary and tentative explanation. More-demanding additional calculations would be necessary to assess its suitability.

Enzyme−Substrate Interactions in the Open and Semiopen States. Interactions between UDP-Gal*p* and *Tc*UGM, when the enzyme is closed, have been thoroughly analyzed in previous experimental[19] and computational studies.[22] They agree with the general recognition pattern reported for eukaryotic UGMs.[15] On the other hand, several important interactions occurring in the semiopen and open states have been described by Da Fonseca et al.[23] We have employed the data collected from the US simulations to scrutinize this feature. In particular, we focused on enzyme−substrate interactions taking place in the semiopen and open states because their previous characterization is less detailed.

When the enzyme is open, the substrate interacts via H-bonds with Arg87, Asn104, Asp181, Arg184, Arg187, Asp195, and Asn201 (see Figure 5a). In particular, Arg184 and Arg187

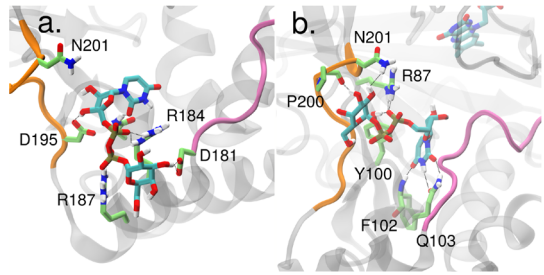


Figure 5. Representative structures highlighting residues that interact with the substrate. (a) Open state of the enzyme and (b) semiopen state. The 180s flap is shown in pink, and the 200s flap is shown in orange. H-bonds are indicated with a dashed line.

strongly interact with the diphosphate group, while Asp195 interacts with ribose. There are also weaker H-bonds between uridine and residues Arg87 and Asn104 and between Gal*p* and residues Asn201 and Asp181. Then, when the 180s flap moves to the intermediate position and the semiopen state is reached, the substrate interacts with Arg87, Tyr100, Phe102, Gln103, Trp198, Pro200, and Asn201 (see Figure 5b). At this stage, the uracil moiety has already achieved its final conformation, tightly bound to Phe102 and Gln103. Tyr100 and Arg87 interact with the diphosphate group, while Gal*p* makes less frequent H-bonds with Trp198, Pro200, and Asn201. Tables S2 and S3 of the Supporting Information show the strength of the relevant H-bonds taking place in the semiopen and open states, at different ranges of the reaction coordinate.

It has to be noted that all the residues mentioned in this paragraph belong to or lie close to the 200s flap. No strong interactions were detected between the substrate and residues of the 180s flap. A similar pattern of interactions, involving the 200s flap but not the 180s flap, has been described for the entrance of NAD(P)H into the active site of *Af*UGM.[49] Finally, it is interesting to note that Arg87 and Asn201 interact with the substrate in all the stages of the process. Moreover, in the closed state, they strongly interact with each other via an H-bond that stabilizes the closed conformation.[15,47]

The degree of conservation among eukaryotic UGMs of residues that significantly interact with UDP-Gal*p* was studied

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

as explained in the Methods section. Figure 6 presents the relevant portions of the sequence logo built as described there.

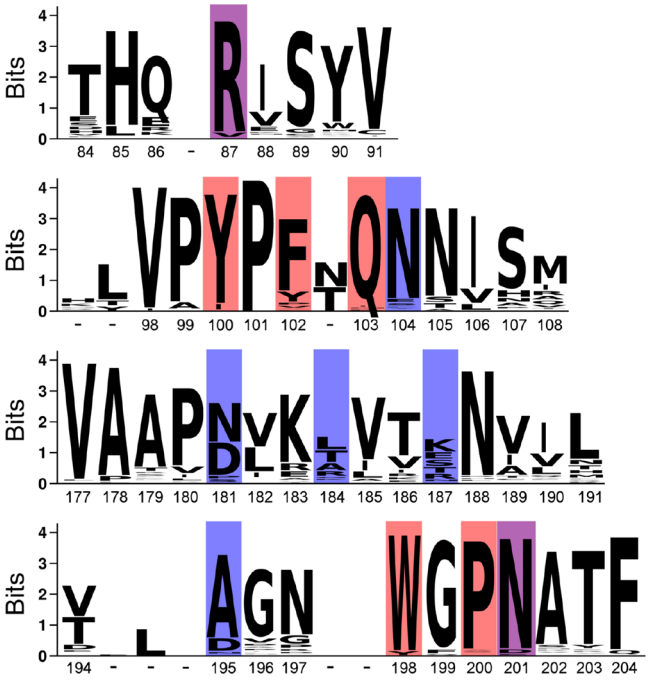


Figure 6. Sequence logo built from 124 sequences of eukaryotic UGMs. Only sections containing the relevant residues are shown. Residues interacting with the substrate when the enzyme is open are highlighted in blue; those interacting when the enzyme is semiopen are highlighted in red. Residues that interact in all the states are highlighted in purple. The numbering of the residues corresponds to *Tc*UGM.

The results show that residues interacting with the substrate when the enzyme is open have poor conservation. Only Asn104 is highly conserved, while position 181 is either occupied by aspartate or asparagine, whose side chains can both establish H-bonds. On the contrary, residues that hold significant interactions when the enzyme is in the semiopen state show a high degree of conservation. Apart from Phe102 and Gln103, which were discussed above, it was shown that Tyr100 (Tyr104 in *Af*UGM) is important for catalysis (The Y100A mutation decreases the enzyme catalytic efficiency by ∼100-fold).[23] Also, residues Arg87 and Asn201 that interact in all the stages are highly conserved. This result is not surprising since previous experiments have demonstrated that these residues are crucial for enzyme functioning. The nonpolar part of the side chain of Arg87 (Arg91 in *Af*UGM) defines the NADPH binding pocket,[15] while the mutation of Asn201 (Asn207 in *Af*UGM) by alanine notoriously increases the Michaelis constant and reduces the catalytic efficiency.[23] In addition, it has been suggested that the interaction between Arg87-Asn201 is essential for the full closure of the 200s flap.[23] In this work, we observe that these residues present H-bond interactions mainly in the closed state of the enzyme, for $\chi <$ 18 Å.

Active-Site Waters. In this section we describe the behavior of the water molecules located within the active site during the binding/unbinding process. As described previously, we used frames from all the US simulations to identify the sites with higher water occupancy. The analysis revealed nine sites with an occupancy probability higher than 0.6. They are depicted in Figure 7. We note that the identity of the molecules located at these places is not always the same.

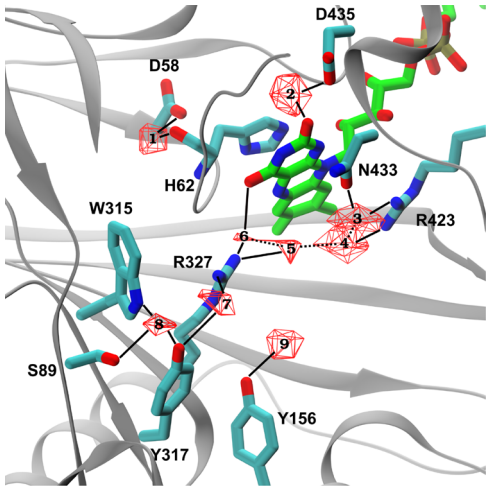


Figure 7. Sites with high occupancy probability for water molecules. The isosurfaces indicate an occupancy probability of 0.6. *Tc*UGM residues are shown in cyan, while the FAD cofactor is shown in green. H-bonds between water at the selected sites and the enzyme or the cofactor are marked with a solid line; H-bonds between water molecules are marked with a dotted line.

Rather, they occasionally exchange with other molecules of the surroundings. Sites **1**, **2**, **3**, **4**, **8**, and **9** are occupied in the crystal structure of *Tc*UGM and persist with high occupancy throughout the whole range of the reaction coordinate. Sites **5**, **6**, and **7** are not occupied in the crystal structure. Waters at sites **5** and **6** interact with Arg327 and with atom O_4 of the cofactor. Besides, they establish H-bonds with each other and with the water molecule located at site **4**. These two molecules are removed from the active site when the substrate binds and the enzyme closes, since the Gal*p* and diphosphate moieties occupy their places. Water molecules located at site **7** are coordinated by Arg327 and Tyr317. When the substrate is bound and the enzyme is closed, they also interact with the diphosphate group of UDP-Gal*p*.

We calculated the average number of water molecules in the active site, as a function of the reaction coordinate. The results are presented at Figure S5 of the Supporting Information. The curve reassembles the behavior of $d_{178-200}$ (see Figure 4a). The number of water molecules fluctuates about a constant value in regions (**I**), (**III**), and (**V**), with ∼30, ∼65, and ∼95 water molecules in the pocket, respectively. Conversely, there are clear increments in the number of water molecules in regions (**II**) and (**IV**), where the enzyme goes from closed to semiopen and from semiopen to open states, respectively. Thus, as the substrate leaves the enzyme, water molecules enter from the solution to the active site cavity. These molecules fill the space previously occupied by the substrate and the extra space created by the opening of the enzyme. Most water molecules enter the cavity passing through flaps 180 and 200. However, a few molecules get to the active site sliding through thin channels that communicate the pocket with the external solution.

CONCLUSIONS

We have presented a detailed characterization for the release of UDP-Gal*p* from *Tc*UGM. The study was based on the analysis

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817

of data collected from extensive molecular dynamics simulations that employed the US technique to take the substrate out of the enzyme. The PMF obtained from these simulations shows that the barrier for the unbinding is 21.4 ± 3.0 kcal/mol. This value is somewhat smaller than the barrier of the rate-determining step for the conversion of UDP-Gal*p* into UDP-Gal*f*, occurring in *Tc*UGM. Under the assumption that the unbinding of UDP-Gal*f* has a PMF similar to that of UDP-Gal*p*, this finding indicates that the chemical reaction is the rate-determining step of the whole process. In addition, the fact that the two barriers are similar suggests that small variations in the enzyme active site or the substrate can change the rate-determining step from the isomerization reaction to product release. This is consistent with biochemical studies that have shown that a chemical step is partially rate-determining in *Tc*UGM and *Af*UGM with UDP-Gal*p*, while in *Af*UGM with UDP-arabinopyranose the product release becomes the rate-determining step.[18,41]

We found that the substrate distortion is significant as it moves from its equilibrium position in the binding pocket to the surface of the enzyme. The Gal*p* moiety undergoes the largest displacements. This group is able to change its conformation even when the substrate is tightly bound to the enzyme. However, the distance between N_5 of FAD and C_1' of Gal*p* is always short so as to facilitate the nucleophilic attack on the sugar. On the other hand, the uracil group is the one that remains more tightly bound and only displaces significantly when the enzyme is open.

Analysis of the conformational changes of the enzyme along the binding process showed that the transition from the open to the closed conformation is not direct. Instead, the enzyme passes through a semiopen state that remains stable in a broad range of the reaction coordinate. This state is achieved because of the movement of the 180s flap, which moves about 6 Å from its position in the open state to reach an intermediate location. The intermediate location is 2.5 Å away from its position in the closed state.

Significant interactions are established between the substrate and the enzyme when the latter is in the semiopen state. Among these, the strongest ones are those between the uracil group with residues Phe102 and Gln103. Other important interactions involve Arg87, Tyr100, Trp198, Pro200, and Asn201. All of these residues are highly conserved among eukaryotic UGMs.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.8b00675.

Figure S1, schematic representation of reaction coordinate; Figures S2−S4, results of different calculations employed to assess consistency and statistical significance of PMF; Figure S5, number of water molecules in active site pocket as a function of reaction coordinate; Table S1, results of classical and QM/MM simulations regarding H-bonds between substrate and cofactor in bound state; Table S2, strength of H-bonds between enzyme and substrate in open state; Table S3, strength of H-bonds between enzyme and substrate in semiopen state; text describing implementation of QM/MM computations (PDF)

Movie S1, typical structures observed during release of UDP-Gal*p* from *Tc*UGM, as revealed by US simulations (MPG)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: juliana@unq.edu.ar.

ORCID

Pablo Sobrado: 0000-0003-1494-5382

Juliana Palma: 0000-0002-0761-0253

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the computational resources provided by the Centro de Cómputos de Alto Rendimiento (CeCAR). This work was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) and the Universidad Nacional de Quilmes (to R.C.P., G.P.S., and J.P.) and the National Institute of General Medical Sciences (NIGMS) award R01GM094469 and the National Science Foundation award CHE-1506206 (to P.S.).

■ REFERENCES

(1) World Heath Organization - Chagas disease (American Trypanosomiasis). http://www.who.int/chagas/en/ (accessed Dec 2018).

(2) Wilkinson, S. R.; Taylor, M. C.; Horn, D.; Kelly, J. M.; Cheeseman, I. A Mechanism for Cross-Resistance to Nifurtimox and Benznidazole in Trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5022−5027.

(3) Campos, M. C.; Leon, L. L.; Taylor, M. C.; Kelly, J. M. Benznidazole-Resistance in Trypanosoma cruzi: Evidence that Distinct Mechanisms Can Act in Concert. *Mol. Biochem. Parasitol.* **2014**, *193*, 17−19.

(4) Turner, C. W.; Lima, M. F.; Villalta, F. Trypanosoma cruzi Uses a 45-kDa Mucin for Adhesion to Mammalian Cells. *Biochem. Biophys. Res. Commun.* **2002**, *290*, 29−34.

(5) Villalta, F.; Madison, M. N.; Kleshchenko, Y. Y.; Nde, P. N.; Lima, M. F. Molecular Analysis of Early Host Cell Infection by Trypanosoma cruzi. *Front. Biosci., Landmark Ed.* **2008**, *Volume* (13), 3714.

(6) Arruda, M. V.; Colli, W.; Zingales, B. Terminal β-d-Galactofuranosyl Epitopes Recognized by Antibodies that Inhibit Trypanosoma cruzi Internalization into Mammalian Cells. *Eur. J. Biochem.* **1989**, *182*, 413−421.

(7) Bertini, S.; Takahashi, H.; Straus, A. Inhibition of Trypomastigotes Infectivity by Monoclonal Antibody BST-1, Directed to Epimastigotes GIPCs. Cross-Reactivity with High Molecular Weight Antigen Present in Trypomastigotes. *Rev. Inst. Med. Trop. Sao Paulo* **2003**, *45* (Suppl.13), 83.

(8) de Castro Oliveira, M. P.; Ramos, T. C. P.; Pinheiro, A. M. V.; Bertini, S.; Takahashi, H. K.; Straus, A. H.; Haapalainen, E. F. Tridimensional Ultrastructure and Glycolipid Pattern Studies of Trypanosoma dionisii. *Acta Trop.* **2013**, *128*, 548−556.

(9) Barreto-Bergter, E.; Vermelho, A. B. Structures of Glycolipids Found in Trypanosomatids: Contribution to Parasite Functions. *Open Parasitol. J.* **2010**, *4*, 84−97.

(10) Suzuki, E.; Mortara, R. A.; Takahashi, H. K.; Straus, A. H. Reactivity of MEST-1 (Antigalactofuranose) with Trypanosoma cruzi Glycosylinositol Phosphorylceramides (GIPCs): Immunolocalization of GIPCs in Acidic Vesicles of Epimastigotes. *Clin. Di- agn. Lab. Immunol.* **2001**, *8*, 1031−1035.

(11) Oppenheimer, M.; Valenciano, A. L.; Sobrado, P. Biosynthesis of Galactofuranose in Kinetoplastids: Novel Therapeutic Targets for

Treating Leishmaniasis and Chagas' Disease. *Enzyme Res.* **2011**, *2011*, 415976.
(12) Fernandes, M. C.; Andrews, N. W. Host Cell Invasion by Trypanosoma cruzi: a Unique Strategy that Promotes Persistence. *FEMS Microbiol. Rev.* **2012**, *36*, 734−747.
(13) Maya, J. D.; Cassels, B. K.; Iturriaga-Vásquez, P.; Ferreira, J.; Faúndez, M.; Galanti, N.; Ferreira, A.; Morello, A. Mode of Action of Natural and Synthetic Drugs Against Trypanosoma cruzi and Their Interaction With the Mammalian Host. *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **2007**, *146*, 601−620.
(14) Kizjakina, K.; Tanner, J.; Sobrado, P. Targeting UDP-Galactopyranose Mutases from Eukaryotic Human Pathogens. *Curr. Pharm. Des.* **2013**, *19*, 2561−2573.
(15) Tanner, J. J.; Boechi, L.; McCammon, J. A.; Sobrado, P. Structure, Mechanism, and Dynamics of UDP-Galactopyranose Mutase. *Arch. Biochem. Biophys.* **2014**, *544*, 128−141.
(16) Beverley, S. M.; Owens, K. L.; Showalter, M.; Griffith, C. L.; Doering, T. L.; Jones, V. C.; McNeil, M. R. Eukaryotic UDP-Galactopyranose Mutase (GLF Gene) in Microbial and Metazoal Pathogens. *Eukaryotic Cell* **2005**, *4*, 1147−1154.
(17) Soltero-Higgin, M.; Carlson, E. E.; Gruber, T. D.; Kiessling, L. L. A Unique Catalytic Mechanism for UDP-Galactopyranose Mutase. *Nat. Struct. Mol. Biol.* **2004**, *11*, 539.
(18) Oppenheimer, M.; Valenciano, A. L.; Kizjakina, K.; Qi, J.; Sobrado, P. Chemical Mechanism of UDP-Galactopyranose Mutase from Trypanosoma Cruzi: a Potential Drug Target Against Chagas' Disease. *PLoS One* **2012**, *7*, No. e32918.
(19) Dhatwalia, R.; Singh, H.; Oppenheimer, M.; Sobrado, P.; Tanner, J. J. Crystal Structures of Trypanosoma cruzi UDP-Galactopyranose Mutase Implicate Flexibility of the Histidine Loop in Enzyme Activation. *Biochemistry* **2012**, *51*, 4968−4979.
(20) Boechi, L.; Oliveira, C. A. F.; Fonseca, I.; Kizjakina, K.; Sobrado, P.; Tanner, J. J.; McCammon, J. A. Substrate-Dependent Dynamics of UDP-Galactopyranose Mutase: Implications for Drug Design. *Protein Sci.* **2013**, *22*, 1490−1501.
(21) Huang, W.; Gauld, J. W. Tautomerization in the UDP-Galactopyranose Mutase Mechanism: A DFT-Cluster and QM/MM Investigation. *J. Phys. Chem. B* **2012**, *116*, 14040−14050.
(22) Pierdominici-Sottile, G.; Cossio-Pérez, R.; Galindo, J. F.; Palma, J. QM/MM Molecular Dynamics Study of the Galactopyranose → Galactofuranose Reaction Catalysed by Trypanosoma cruzi UDP-Galactopyranose Mutase. *PLoS One* **2014**, *9*, No. e109559.
(23) Da Fonseca, I.; Qureshi, I. A.; Mehra-Chaudhary, R.; Kizjakina, K.; Tanner, J. J.; Sobrado, P. Contributions of Unique Active Site Residues of Eukaryotic UDP- Galactopyranose Mutases to Substrate Recognition and Active Site Dynamics. *Biochemistry* **2014**, *53*, 7794−7804.
(24) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank, 1999−. In *International Tables for Crystallography Vol. F: Crystallography of biological macromolecules*, 1st online ed.; Rossmann, M. G., Arnold, E., Eds.; Springer: 2006; pp 675−684, DOI: 10.1107/97809553602060000722.
(25) van Straaten, K. E.; Routier, F. H.; Sanders, D. A. Structural Insight into the Unique Substrate Binding Mechanism and Flavin Redox State of UDP-Galactopyranose Mutase from Aspergillus fumigatus. *J. Biol. Chem.* **2012**, *287*, 10780−10790.
(26) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple AMBER Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.
(27) 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.
(28) 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.
(29) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. *AMBER 16*; University of California: San Francisco, 2016.
(30) Kästner, J. Umbrella Sampling. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 932−942.
(31) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(32) Rosta, E.; Hummer, G. Free Energies from Dynamic Weighted Histogram Analysis Using Unbiased Markov State Model. *J. Chem. Theory Comput.* **2015**, *11*, 276−285.
(33) Grossfield, A. *An implementation of WHAM: the Weighted Histogram Analysis Method*, version 2.0.9. http://membrane.urmc.rochester.edu/content/wham (accessed Jan 2017).
(34) Pohorille, A.; Jarzynski, C.; Chipot, C. Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(35) Kullback, S.; Leibler, R. A. On Information and Sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79−86.
(36) Zhu, F.; Hummer, G. Convergence and Error Estimation in Free Energy Calculations Using the Weighted Histogram Analysis Method. *J. Comput. Chem.* **2012**, *33*, 453−465.
(37) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(38) Uniprot Consortium.. UniProt: the Universal Protein Knowledgebase. *Nucleic Acids Res.* **2017**, *45*, D158−D169.
(39) Edgar, R. C. MUSCLE: a Multiple Sequence Alignment Method with Reduced Time and Space Complexity. *BMC Bioinf.* **2004**, *5*, 113.
(40) Crooks, G. E.; Hon, G.; Chandonia, J.-M.; Brenner, S. E. WebLogo: a Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188−1190.
(41) Pierdominici-Sottile, G.; Cossio-Pérez, R.; Da Fonseca, I.; Kizjakina, K.; Tanner, J. J.; Sobrado, P. Steric Control of the Rate-Limiting Step of UDP-Galactopyranose Mutase. *Biochemistry* **2018**, *57*, 3713−3721.
(42) Cremer, D. T.; Pople, J. General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.
(43) Fushinobu, S. T. *Cremer-Pople Parameter Calculator*; 2006. http://enzyme13.bt.a.u-tokyo.ac.jp/CP/ (accessed Jan 2018).
(44) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. Recommendations for the Presentation of NMR Structures of Proteins and Nucleic Acids. *J. Mol. Biol.* **1998**, *280*, 933−952.
(45) Elstner, M. The SCC-DFTB Method and its Application to Biological Systems. *Theor. Chem. Acc.* **2006**, *116*, 316−325.
(46) Gruber, T. D.; Borrok, M. J.; Westler, W. M.; Forest, K. T.; Kiessling, L. L. Ligand Binding and Substrate Discrimination by UDP-Galactopyranose Mutase. *J. Mol. Biol.* **2009**, *391*, 327−340.
(47) Dhatwalia, R.; Singh, H.; Oppenheimer, M.; Karr, D. B.; Nix, J. C.; Sobrado, P.; Tanner, J. J. Crystal Structures and Small-Angle X-Ray Scattering Analysis of UDP- Galactopyranose Mutase from the Pathogenic Fungus Aspergillus fumigatus. *J. Biol. Chem.* **2012**, *287*, 9041−9051.
(48) Durrant, J. D.; Votapka, L.; Sørensen, J.; Amaro, R. E. POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics. *J. Chem. Theory Comput.* **2014**, *10*, 5047−5056.
(49) Dhatwalia, R.; Singh, H.; Solano, L. M.; Oppenheimer, M.; Robinson, R. M.; Eller-brock, J. F.; Sobrado, P.; Tanner, J. J. Identification of the NAD(P)H Binding Site of Eukaryotic UDP-Galactopyranose Mutase. *J. Am. Chem. Soc.* **2012**, *134*, 18132−18138.

DOI: 10.1021/acs.jcim.8b00675
J. Chem. Inf. Model. 2019, 59, 809−817