bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Energy Landscape of the Domain Movement in *Staphylococcus aureus* UDP-N-acetylglucosamine 2-epimerase

Erika Chang de Azevedo[1] and Alessandro S. Nascimento[1*].

1. Instituto de Física de São Carlos, Universidade de São Paulo. Av. Trabalhador SaoCarlense, 400. Parque Arnold Schmidit. São Carlos, SP, Brazil. 13566-590.

* Corresponding author. E-mail: asnascimento@ifsc.usp.br. Phone: +55-16-3373-8709.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Staphylococcus aureus is an important cause of resistant healthcare-associated infections. It has been shown that the Wall Teichoic Acid (WTA) may be an important drug target acting on antibiotic-resistant cells. The UDP-N-Acetylglucosamine 2-epimerase, MnaA, is one of the first enzymes on the pathway for the biosynthesis of the WTA. Here, detailed molecular dynamics simulations of *S. aureus* MnaA were used to characterize the conformational changes that occur in the presence of UDP and UDP-GlcNac and also the energetic landscape associated with these changes. Using different simulation techniques, such as ABMD and GAMD, it was possible to assess the energetic profile for the protein with and without ligands in its active site. We found that there is a dynamic energy landscape that has its minimum changed by the presence of the ligands, with a closed structure of the enzyme being more frequently observed for the bound state while the unbound enzyme favors an opened conformation. Further structural analysis indicated that positively charged amino acids associated with UDP and UDP-GlcNac interactions play a major role in the enzyme opening movement. Finally, the energy landscape profiled in this work provides important conclusions for the design of inhibitor candidates targeting *S. aureus* MnaA.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Staphylococcus aureus is a naturally occurring Gram-positive bacteria that can be often found in the skin and mucosa of humans and other animals [1]. Although usually harmless, this organism can cause a wide range of infections in humans, such as folliculitis, furuncles, and carbuncles, impetigo, mastitis, wound infections, and staphylococcal scalded skin syndrome. More dangerous pathological conditions caused by *S. aureus* include bacteremia, pneumonia, endocarditis, bone, and joint infections, and toxic shock syndrome [2]. *S. aureus* is also an important cause of hospital infections. According to data collected by the CDC between 2011 and 2014, this microorganism was responsible for approximately 10% of the overall healthcare-associated infections analyzed in the United States [3].

The β–lactams, such as methicillin and penicillin, are among the antibiotics considered as the first choice for the treatment of this Gram-positive bacterial infection. These antibiotics act on the peptidoglycan layer of the cell, inhibiting transglycosylases and transpeptidases, collectively known as penicillin-binding proteins or PBPs, that are responsible for the synthesis and crosslink of specific polymers of the cell wall [4]. However, the treatment of Staphylococcal infections is becoming more challenging due to the emergence of drug-resistant strains [5]. These resistant strains were first reported only a few years after the introduction of penicillin and have increased in number of occurrences faster after the introduction of new antibiotics [6]. The CDC data, taken from 4,515 reporting hospitals and over 365,000 healthcare-associated infections, show that about 50% of the infections caused by *S. aureus* in healthcare units had isolated strains that were reported to be resistant to the β-lactam antibiotics like oxacillin, methicillin, and cofoxitin [3].

S. aureus antibiotic resistance has a direct impact in the human lives but also accounts for an important burden in the health systems. Data taken in 2007 from 31 countries that participate in the European Antimicrobial Resistance Surveillance System (EARSS) showed that 27,711 bloodstream infections caused by MRSA resulted in 5,503 excess deaths and 255,683 excess hospital days, thus generating in a burden of 44 million Euros (63.1 million dollars) [7].

The emergence of antibiotic resistance highlights the need for new drugs or even new ways to re-sensitize resistant bacterium. In this context, the cell wall remains as an interesting target for the discovery of new drugs and agents that fight β–lactam resistance [4]. The cell wall of Gram-positive bacteria, such as *S. aureus*, is typically composed by peptidoglycan and wall teichoic acid (WTA). The WTA is located in the outer layer of the peptidoglycan and its

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

role is associated the maintenance of the cell shape, virulence, coordination of the peptidoglycan synthesis, biofilm formation and, more recently, it has been linked to antibiotic resistance [6,8–10]. In *S. aureus*, the WTA is typically made of a poly(ribitol-phosphate) polymer that is synthesized by the sequential action of several enzymes, called *Tar* (Teichoic acid ribitol) enzymes, starting with the glycosyltransferases TarA and TarO. Another important enzyme involved in the biosynthesis of WTA and in the *S. aureus* resistance to β-lactams is the N-acetylglucosamine 2-epimerase or MnaA [11]. This enzyme interconverts UDP-N-acetylglucosamine and UDP-N-acetylmannosamine, which are the substrates for TarO and TarA, respectively, thus regulating the balance of the starting points of the Tar biosynthesis pathway [11].

The *S. aureus* MnaA (SaMnaA) binds simultaneously to a UDP molecule and to the substrate UDP-N-acetylglucosamine (UDP-GlcNac) or UDP-N-acetylmannosamine (UDP-ManNac) [11,12]. Analysis of *S. aureus* MnaA (SaMnaA) crystal structure [11] and comparison with previously determined crystal structures revealed that the protein exhibits a domain movement upon binding the substrate, defining at least three observable conformations for the enzyme as opened, semi-closed and closed [12]. The opened state was observed in the absence of the substrate or product in the active site. The semi-closed state is observed in the presence of UDP. The closed state is typically observed when the protein has both UDP and UDP-GlcNac bound in its active site [11,12]. Interestingly, this movement between domains is also observed for several enzymes, including some enzymes of the WTA biosynthesis, such as TarM [10].

Here, we evaluated the domain movement of SaMnaA at the atomic level using different molecular dynamics simulations and assessing the free energy landscape associated to such movement. The simulations showed that the protein has a shift in the minimum of the free energy landscape profile when associated with the ligands UDP and UDP-GlcNac. Additionally, a major driving force that causes this shift seems to come from the electrostatic interactions between the enzyme and the substrate. These results can lead to a binding energy profile that indicates a dynamic energy landscape [13] for the open/close movement of the protein, relevant for the enzymatic action.

Methods

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Protein and Ligand Parameterization

The chain A of the *S. aureus* MnaA structure (PDB Id 5ENZ) was parameterized using AMBER FF14SB force-field [14]. Hydrogen atoms were added and the histidine protonation states determined at the optimum pH of 7.0 using the H^{++} server [15,16]. The protein was immersed in a water box using TIP3P water model [17] and sodium ions were also included to keep the systems electrically neutral. UDP and UDP-GlcNac atomic charges were calculated using the Hartree-Fock method, with 6-31G* basis set as available in Gaussian 09 [18]. The obtained atomic charges were then refitted with RESP (Restrained Electrostatics Potential)[19] procedure using ANTECHAMBER [20]. The Lennard-Jones parameters for these ligands defined using GAFF force field [21]. The crystal structure of SaMnaA (PDB Id 5ENZ [11]) was used as the starting structure for the APO MD simulations. The UDP and UDP-GlcNac bound structure (HOLO) was modeled using the chain A of the crystal structure of the SaMnaA (PDB Id 5ENZ) and the crystal structure of *Bacillus anthracis* MnaA (BaMnaA) bound with UDP and UDP-GlcNac (PDB Id 3BEO[12]), using the program Modeller [22,23]. The purpose of this initial modeling was to generate a 'bound-state' (closed) conformation of SaMnaA with UDP and UDP-GlcNac coordinates from BaMnaA. The final systems comprised 45,519 and 55,479 atoms for the APO and HOLO system, respectively.

Equilibrium Simulations

The initial energy of the APO system was minimized in 1,000 cycles of conjugated gradient minimization. The system was then heated to 300 K in 50 ps, keeping the volume constant. The density of the system was afterward equilibrated to 1 g/cm^3 by keeping constant pressure at 1 atm in 400 ps. During the minimization, heating and density equilibration, harmonic restraints were applied to all residues, with weight of 2 kcal.mol^{-1}. Å^{-2}. Finally, the system was equilibrated for 2 ns, without restraints, and the productive simulation was performed for 1 µs in an NPT ensemble and, afterward, extended up to 5 µs. Periodic boundary conditions were applied using a 12 Å cutoff radius for truncating short-range potentials. The simulation protocol for the HOLO system followed a similar procedure, with 4000 cycles of steepest descent energy minimization followed by 4000 steps of conjugated gradient minimization. The system was then heated to 300 k in 50 ps, with constant volume. The density was equilibrated until 1 g/cm^3 in 400 ps keeping constant pressure at 1 atm. For

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

minimization, heating and density equilibration, a harmonic positional restraint was applied to the entire protein and for the two ligands with a weight of 2 kcal.mol^{-1}.Å^{-2}. Finally, the system was equilibrated for 10 ns and the productive simulation was performed for 1 µs without restraints. All the simulations were performed using AMBER16 [24,25].

Gaussian Accelerated Molecular Dynamics

In Gaussian Accelerated Molecular Dynamics (GAMD) [26], the trajectory is propagated in a non-negative Gaussian boost potential. This potential is applied to reduce the potential barriers and accelerate the state transitions, thus enhancing the conformational sampling. When the system potential V is lower than a threshold E, a harmonic boost potential ΔV is introduced to the total energy. The energy boost ΔV is chosen to result in the highest acceleration as possible but still allowing an accurate reweighting and free energy profile estimation. As shown in Equation 1, ΔV is defined as a harmonic potential with force constant k, which depends on the constant σ_0 (Equations 2 and 3) [26]. In Equations 2 and 3, V_{max} and V_{min} are the maximum and minimum system potential collected before the GAMD begins. σ_v and V_{av} are the standard deviation and average of the system potential energy, respectively. As previously observed, σ_0 has to be chosen carefully in a compromise to get the best possible acceleration but still allowing an accurate reweighting of the accelerated potentials. One advantage associated with GAMD and also with Accelerated MD (aMD) is the lack of a pre-chosen *reaction coordinate* or a *collective variable* (CV). Also, as pointed out by Miao and coworkers, GAMD preserves the original potential surface, allowing the computation of a PMF with less noise as compared to aMD [27]. Here, we applied the boost potential to the dihedral and the total potential of the system, in a dual-boost approach. For both potentials, the σ_0 was set to 3 kcal/mol, based on preliminary simulations with varying values for σ_0, where 3 kcal/mol provided the higher acceleration.

$$\Delta V(\vec{r}) = \frac{1}{2}k(E - V(\vec{r}))^2, V(\vec{r}) < E \qquad (1)$$

$$k = k_0 \left(\frac{1}{Vmax - Vmin}\right) \qquad (2)$$

$$k_0 = \min(1.0, k_0') = min\left(1.0, \frac{\sigma_0}{\sigma_V}\frac{V_{max} - V_{min}}{V_{max} - V_{av}}\right) \qquad (3)$$

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The resulting boost potential followed a Gaussian distribution that allowed a reweight and the calculation of a potential of mean force. The reweighting was made by obtaining the second-order of the cumulant expansion for the potential boost, as previously described [28]. This calculation was made based on an opening angle computed between 10-80 degrees, with angle increments of 4 degrees.

Adaptively Biased Molecular Dynamics

The Adaptively Biased Molecular Dynamics (ABMD) is an umbrella sampling method similar, in spirit, to metadynamics. The biasing force compensates for the free energy gradient, reproducing the previously unknown free energy surface for the system analyzed in the direction of the collective variable (CV), chosen by the user [29]. Here, we used the angle between the α-carbon of the residues S128, G349, and P213, describing the opening angle of the SaMnaA structure. The simulation was conducted with a well-tempered umbrella sampling method [30], with pseudo-temperature T defined at 5,000 K.

The total set of simulations used in this work are provided in Table 1.

Table 1. Simulations used in this work. The simulation methods and times used are shown.

Method	Protein System	Time (ns)
Equilibrium	APO	1,000 + 4,000
Equilibrium	HOLO	1,000
GAMD	APO	3 x 500
GAMD	HOLO	3 x 250
ABMD	APO	3 x 250
ABMD	HOLO	3 x 250

Structural Analysis

The APO system simulation of 5 µs was analyzed using the algorithms available in the pyEMMA library [31]. Only the carbon alpha coordinates from the protein were used in the analysis, and the lag time used in the time-lagged independent component analysis (TICA) was set 200

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

frames or 10 ns. The electrostatic forces between the C-terminal and N-terminal domains of the protein were conducted using the Delphi Force Web server [32], where the first domain was defined as the protein region encompassing residues 1 to 181. Three representative states of the protein were used in the calculations, open, intermediate and closed, with the ligands taken into account with the C-terminal domain. Geometric analysis of the MD trajectories were carried out with CPPTRAJ [33] and interaction energies were computed with AmberEnergy++, a locally written program to compute interaction potential energies according to the AMBER force field [34].

Results

The states defined as open, semi-closed and closed can be better characterized by the opening angle observed for protein structures available on the PDB. The UDP-GlcNac 2-epimerase of *Methanocaldococcus janaschii* structure was solved without ligands (PDB id 4NEQ) [35] and with both UDP and UDP-GlcNac (PDB id 4NES) [35]. The first structure is in the open state, with an opening angle (CV) of 54 degrees. Here, the opening angle is defined as the angle between the Cα atoms of the residues S128, G349, and P213 in SaMnaA, or residues in similar positions for the other epimerases. The second structure is in a closed state, with a CV angle of 28 degrees. The SaMnaA was in the semi-closed state, with CV angle of 31 degrees, for chain A. A description of the opening angles, defined by our chosen CV, for all the UDP-GlcNac 2-epimerases with available structures to date is given in Table 2.

Table 2 – Opening state of known crystal structures of UDP-GlcNac 2-epimerases. *chain without ligands

PDB ID	Organism	Ligands	Oligomeric state	CV Angle Chain A (deg)	CV Angle Chain B (deg)
			Enzymes in the unbound state		
4NEQ	*M. jannaschii*	-	Monomer	53	-
1V4V	*T. thermophilus*	-	Dimer	53	40
3OT5	*L. monocytogenes*	-	Dimer	47	33
1O6C	*B. subtilis*	-	Dimer	53	58
4HWG	*R. bellii*	-	Monomer	25	-
3DZC	*V. cholerae*	-	Dimer	48	46

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Enzymes bound to UDP and UDP-GlcNac					
1VGV	*E. coli*	UDP-GlcNac	Dimer	40*	25
3BEO	*E. coli*	UDP, UDP-GlcNac	Dimer	30	30
4FKZ	*B. subtilis*	UDP, UDP-GlcNac	Dimer	25	25
4NES	*M. jannaschii*	UDP, UDP-GlcNac	Monomer	29	-
5DLD	*B. vietnamiensis*	UDP, UDP-GlcNac	Monomer	24	-
Enzyme bound to UDP only					
5ENZ	*S. aureus*	UDP	Dimer	31	26
1F6D	*E. coli*	UDP	Dimer	40	27

From the analysis of the results shown in Table 2, it can be observed that (*i*) for crystal structures of UDP-N-acetylglucosamine 2-epimerases in the absence of ligands (1O6C, 1V4V, 3OT5, 4NEQ and 3DZC) the opening angle varies between 47 and 53°, indicating the tendency to a opened state without substrate and UDP. The exception here is the crystal structure of *Rickettsia bellii* (4HWG). We hypothesize that the cubic symmetry observed in this crystal structure may stabilize the closed conformation in the absence of a ligand, due to an increased number of interchain contacts. We also observe that (ii) in all cases where both substrate (UDP-GlcNac) and cofactor (UDP) are found bound in the crystal structure (1VGV, 3BEO, 4FKZ, 4NES, and 5DLD) the observed opening angle is reduced to 24-30°, significantly reduced as compared to the unbound state for MnaA. Finally, (iii) SaMnaA (5ENZ) and 1F6D are bound with UDP only. Here, both proteins have different opening angles for each monomer, leaving opening questions for this scenario. A possible explanation is that UDP could promote a variety of intermediate states going from a closed state with an opening angle of 26-27° up to more opened states with angles of 30-40°.

Altogether, the analysis of the available structural data indicates the existence of a major structural change in the MnaA that seems to be associated with the binding of UDP/UDP-GlcNac. However, the crystal structures are single conformations of an ensemble of thermally accessible structures that could be sampled in solution. In order to get a more detailed perspective of this structural change, we used MD simulations to generate an ensemble of conformations in unbound (APO) and ligand-bound (HOLO) scenarios.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Equilibrium Simulations

The APO SaMnaA enzyme was simulated for 1 µs. During this simulation, a clear domain movement was observed in the MD trajectory and its magnitude could be measured by the change in the opening angle chosen as the collective variable (CV). As shown in Figures 1A and 1B, the enzyme accessed the opened conformation with an opening angle of about 53 degrees and also a more intermediate state with CV angle of about 40 degrees. The time evolution of the opening was very fast, with opening happening still during the equilibration. On the other hand, once the structure was opened, it remained opened (or in the intermediary state) during the entire simulation, never visiting the closed state over the 1 µs simulation time.

In order to increase the sampling, the initial 1 µs simulation was further extended up to 5 µs. Even in this longer timescale, the closed conformation was never visited (Figure 1E and 1F). Here, the intermediary state was mostly populated, resulting in a bimodal distribution of the CV angle with peaks at about 40 and 53 degrees, as shown in Figure 1F.

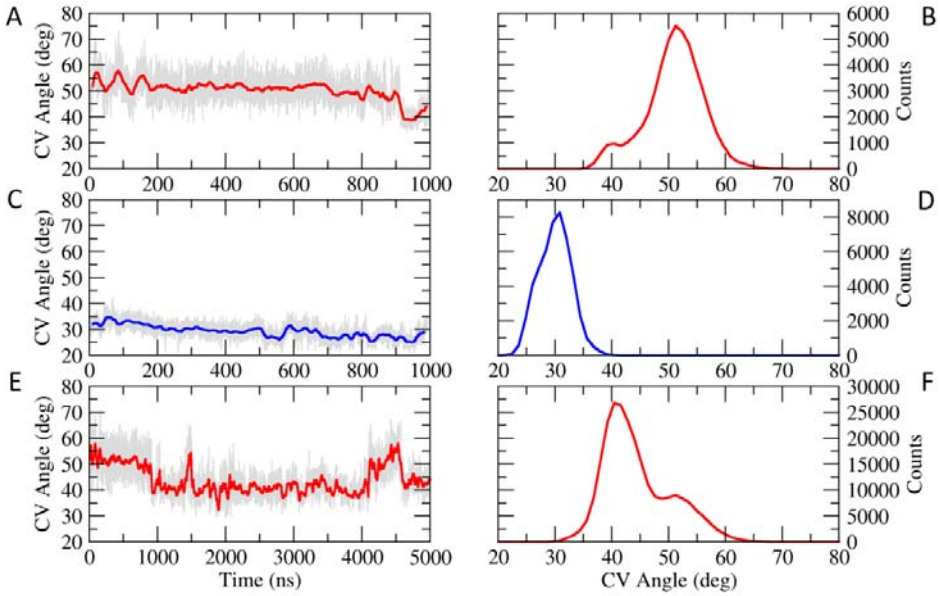


Figure 1. (A) Angles observed during the simulation time for the 1-µs simulation of the APO SaMnaA structure. (B) Histogram of observed angles during the 1-µs simulation of the APO SaMnaA. (C) Angles observed during the simulation of the HOLO SaMnaA structure. (D) Histogram of observed angles in HOLO SaMnaA simulation. (E) Angles observed during the 5-µs simulation of the APO SaMnaA structure. (F) Histogram of the observed angles during the 5-µs APO SaMnaA simulation.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The 1-μs MD trajectory was analyzed using Principal Component Analysis (PCA) [33], and we observed that the first two components described an opening movement of the enzyme domains (first PC) coupled to a torsion between them (second PC), as indicated by the arrows in Figure 2. Confirming the observations made based on the CV angle and the visual inspection of the trajectories, SaMnaA has a low-frequency domain movement towards an opened state in the absence of a ligand and cofactor. At this point, it is tempting to speculate that this opening movement is fast since it was observed in the very beginning of the simulation time.

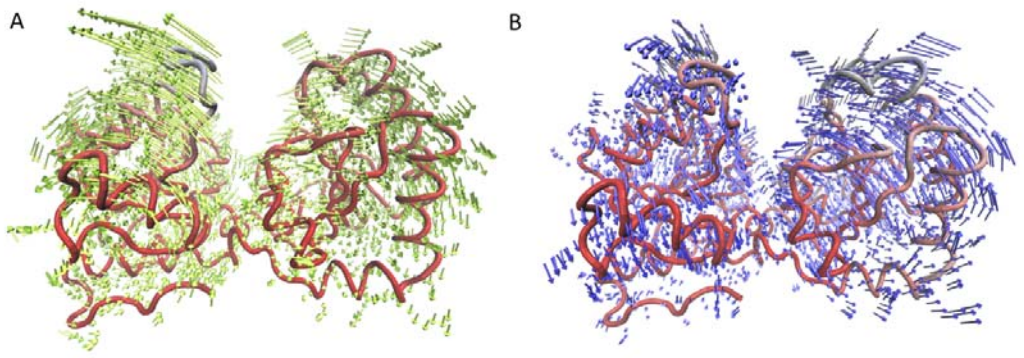


Figure 2. Projections of the first (A) and second (B) components of the PCA analysis of the APO SaMnaA system MD trajectory.

The 5-μs MD trajectory of the APO system was further analyzed using PyEMMA [31]. Within this package, the MD trajectory (reduced to the representation of Cα atoms only) was clustered to define three major macrostates, as shown in Figure 3B. The representative structures show that the structures are going from an intermediate state to the opened one (Figure 3A and 3C). The first time-independent component (TIC), that describes the slowest movement of the system is in agreement with the behavior of the CV angle chosen, which indicates that this was a good choice of the CV indeed (Figure 3A). The Markov-State Modeling of the clusters allowed us to estimate a typical time for the closing movement, i.e., transitioning from the opened to the intermediary state, as 879 ns, and for the opening, i.e., going from the intermediary state to the opened state, as 423 ns. Since the closed state was

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

never sampled in the productive simulation, it was not possible to estimate the typical timescale for the full opening of the enzyme.

In summary, the results obtained from MD simulation for the APO state suggest that the unbound enzyme is found primarily in the opened state, but also sampling an intermediary state (very similar in structure to the opened state). The closed state was never observed, even in longer simulation times, consistently with the predominance of opened conformations observed in crystal structures available in the PDB.

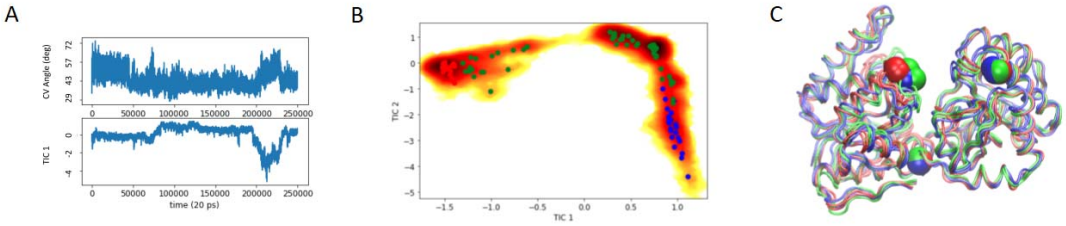


Figure 3. (A) Comparison between the CV angle measured (upper plot) and the TIC 1 from the TICA analysis (lower plot). (B) Heatmap of TIC1 versus TIC2 where the systems stated were coarse-grained, resulting in the separation of three macrostates, shown here as cluster defined by colors red, green and blue. (C) Representative structures of the three states found for the system, with atoms used for CV definition atoms in evidence.

The equilibrium molecular dynamics for the HOLO system showed a markedly different behavior as compared with the APO one (Figure 1C and 1D). Here, the enzyme shows a very well-defined preference for the closed state, with a typical CV angle of about 28-30 degrees. In agreement with crystal structures, the presence of the substrate and of the cofactor drives the equilibrium towards a closed conformation that is maintained during the entire simulation time. The trajectory was also analyzed using PCA (Figure 4) and the two first components represented on the protein structure do not show a clear coordinated movement towards the domain movement for opening, as observed for the unbound SaMnaA.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

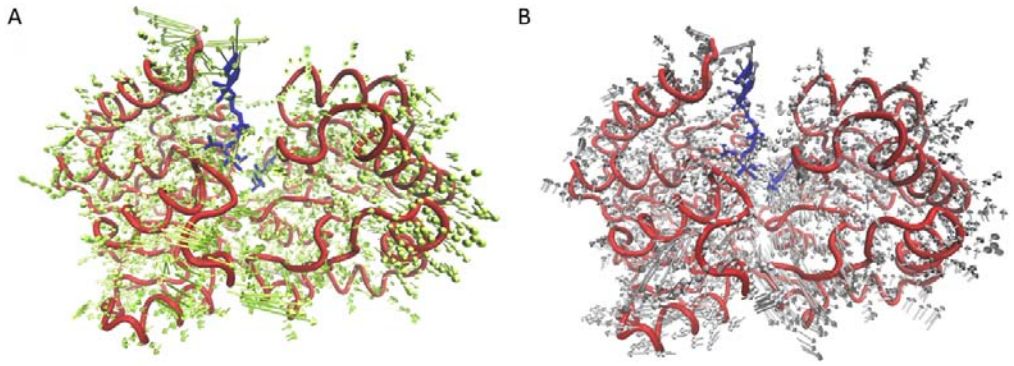


Figure 4. Projections of the first (A) and second (B) components of the PCA analysis for the HOLO SaMnaA system MD trajectory.

For the HOLO SaMnaA, an analysis of the interaction energies involved in protein-UDP and protein-UDP-GlcNac interaction revealed a complex network of polar interactions involving the ligands, the enzyme and solvent atoms. Figure 5A shows a LIGPLOT [36] representation of the interactions involving UDP for a snapshot of the equilibrium simulation. The interactions between UDP and the enzyme mainly involve the residues Glu288, Arg206, Ser282 and Arg11, in addition to several solvent molecules and an interaction between UDP and UDP-GlcNac. The individual contributions of these residues to the total interaction potential energy are shown in Figure 5C, where Arg11, and Arg206 stand out as major contributors due to their interactions with the phosphate groups of UDP. It is also interesting to note the positive (unfavorable) contribution from Glu288 for the UDP interaction (Figure 5C), due to the charge-charge repulsion with the UDP phosphate groups.

For UDP-GlcNac (Figure 5B), the set of interacting residues include Arg206, His40, Arg207, Lys239, His238 and Glu132 in addition to the interactions with Leu68, Gln66, and Met98 through their main chain atoms. Again, as expected, the major contributors to the total interaction potential are the positively charged residues Arg206, Arg207 and Lys239, followed by the histidine residues His40 and His238 (Figure 5D). In conclusion, the interaction with UDP and UDP-GlcNac is maintained mainly through polar residues with the special contribution of positively charged residues in the vicinity of the active site.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

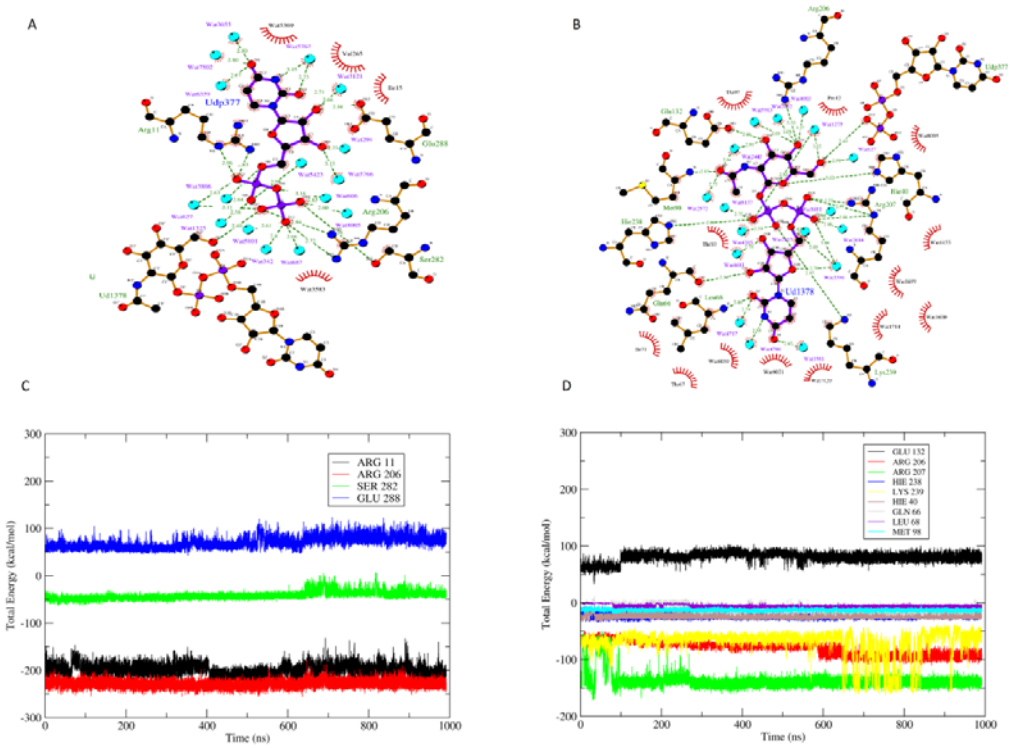


Figure 5. LIGPLOT depiction of the interactions between UDP (A) and UDP-GlcNac (B) with SaMnaA. Total interaction (potential) energy between UDP (C) and UDP-GlcNac (D) and its interacting amino acids are also shown.

Energy Landscape of the Opening for the Unbound Enzyme

In order to better understand the energy landscape associated with the conformational changes coupled to substrate binding, we also run non-equilibrium simulations of the APO and HOLO systems. Although GAMD simulations generate noisier free energy profiles (potential of mean forces) when compared to umbrella sampling techniques, it does not require a pre-chosen CV. Considering this advantage, we used GAMD to probe the energy landscape of the domain movement in SaMnaA.

Three independent simulations of the APO system were performed with three different starting conformations. These conformations are three different coordinates extracted from the equilibrium simulation (0 ns, 50 ns and 100 ns of the productive simulation). Each conformation was simulated for 500 ns. In the accelerated simulations, we observed that in two simulations where the starting conformations with opening angles were close to 55° were reduced to intermediary states with opening angles close to 40-43°. In the

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

third simulation, the enzyme evolved more flexibly, with the opening angle continuously varying between 45 to 65°, as shown in Figure 6A. The histogram of the accelerated simulations indicates a clear preference of the system for the intermediary state (Figure 6B). Interestingly, even with the boost potential, the APO enzyme never visits the closed state, suggesting a high energy barrier to reach this conformation for the unbound enzyme.

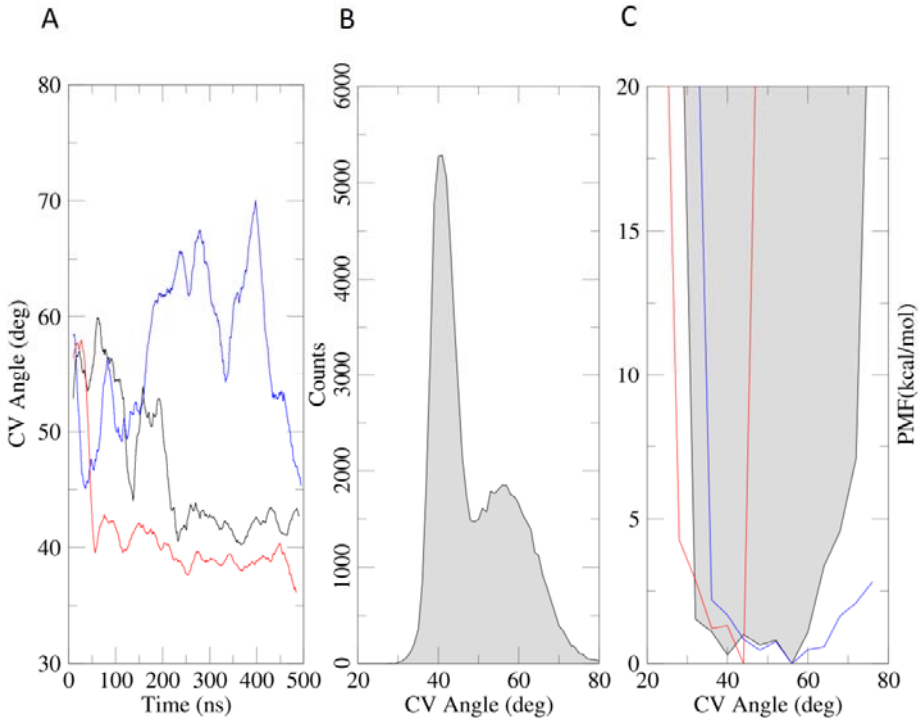


Figure 6. GAMD simulation for APO SaMnaA structure. The results from three simulations are shown: GAMD-APO-I in black curve, GAMD-APO-II in red curve and GAMD-APO-III in blue curve. (A) Observed CV angle versus simulation time. (B) Distribution of observed CV angles during the simulation. C) PMF computed after reweighting.

Potential of mean force (PMF) profiles were computed from the GAMD simulations and are shown in Figure 6C. Two slightly different profiles can be traced from the computed PMFs: (i) in all cases, the energies increase rapidly when the opening angle is below 35°. The simulation GAMD-APO-I (black curves in Figures 6A and 6C) shows a low energy for the conformations encompassing opening angles between 35 to 55 degrees, GAMD-APO-II (red curves in Figures 6A and 6C) shows low energies associated with opening angles between 35 to 45 degrees and GAMD-APO-III (blue curves) show low energies associated with opening angles between 35 to 70 degrees. All simulations agree that the intermediary state is a low energy

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

state for the apo system, and simulations I and III also suggest that the opened state (CV angle > 50°) can be thermally accessed by the unbound enzyme. Although the PMFs computed using GAMD data are coarse, the overall results agree with the 5-µs simulation data, suggesting that, once unbound, a putative equilibrium between conformations with opening angles between 40-45 degrees (intermediate) and 50-60 degrees (opened conformation) are expected to be found in solution.

Similarly, three independent GAMD simulations for the HOLO SaMnaA were conducted and analyzed. The simulations also showed that the bound protein tends to stay in a closed conformation, with an opening angle of about 30 degrees, as shown in Figure 7A. Here, two simulations remained the entire simulation time in the closed conformation (GAMD-HOLO-II and GAMD-HOLO-III, shown in Figure 7A in red and blue curves, respectively) while the other simulation (GAMD-HOLO-I, shown in black curve in Figure 7A) showed a continuously opening angle, starting with about 26 degrees at the beginning of the GAMD simulation and reaching about 48 degrees at the end of the simulation. Interestingly, in this simulation, the ligand (UDP-GlcNac) was observed to unbind from the enzyme at a simulation time of 175 ns, as indicated in the binding energies and RMSD computed from this trajectory (Figures 7D and 7E).

The PMFs calculated for the GAMD simulations of the HOLO system (Figure 7C) indicated a clear minimum in the free energy associated with an opening angle of about 25 degrees. For the simulations GAMD-HOLO-I, where the unbinding event was observed, a shift in the minimum towards the intermediary state was observed, consistent with minimum expected for an APO enzyme. The unbinding of the UDP-GlcNac causes a major shift in the UDP position inside the active site, forcing it to lose some of the interactions with the protein. Altogether, the histogram of the opening angle over the three GAMD simulations clearly indicates a preference for the closed state for the bound system, consistently with the equilibrium simulations. In addition, the unbinding event observed due to the acceleration applied in our non-equilibrium simulation clearly suggests that the opening event is coupled with the absence of the ligand.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

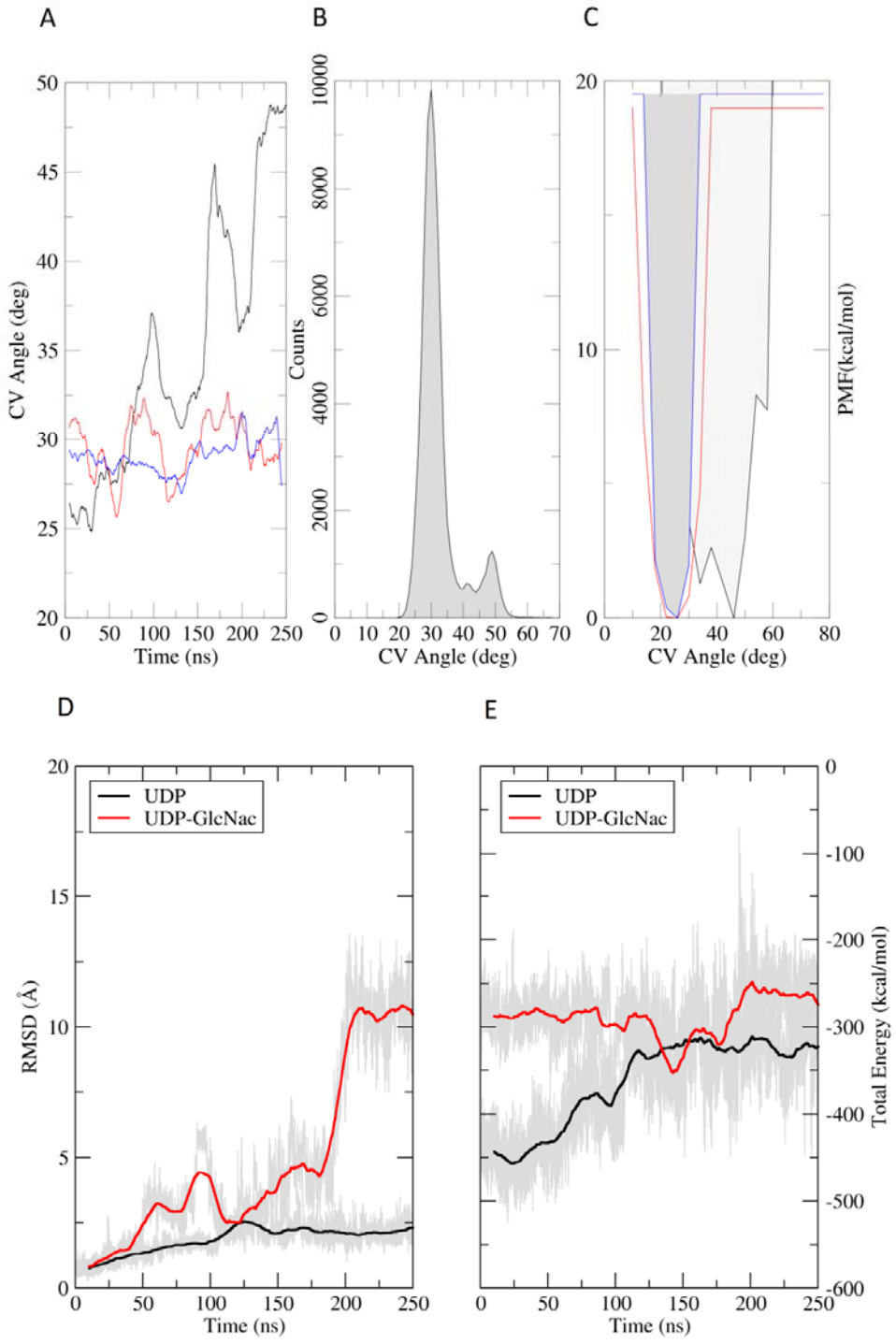


Figure 7. GAMD simulation data for the HOLO SaMnaA. The results from the three independent simulations are shown: GAMD-HOLO-I in black curve, GAMD-HOLO-II in red and GAMD-HOLO-III in blue curve. (A) Observed CV angles during the simulation time. (B) Histogram of the observed angles during the simulation. (C) Computed PMFs from the energy reweighting. (D) RMSD calculated for the UDP (black) and UDP-GlcNac (red). (E) Total interaction energies computed for UDP (black) and UDP-GlcNac (red) with the protein.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The coarse description of the energy profile associated with the conformational state of SaMnaA may be associated with the poor sampling of the transition between the states, which is rarely observed in our simulations. In this way, a finer energetic description can be obtained for enhanced sampling techniques where the reaction coordinate or collective variable is systematically sampled over the simulation. Looking for this finer description of the opening/closing movement in SaMnaA, we explored the enzyme dynamics with Adaptatively-Biased MD (ABMD).

Three independent ABMD simulations were conducted for the APO SaMnaA, with the same starting coordinates as chosen for GAMD. Each simulation explored the opening angle as a CV with a range of 10-100 degrees and resolution of 1 degree in 250 ns of ABMD simulations. The results obtained, shown in Figure 8, indicate that the CV angles in the range of 40-60° lie in the minimum of the free energy landscape, with energy barriers of about 10 kcal/mol (~16 kT) to reach a closed state, defined as a CV angle of about 25 degrees. The first simulation (ABMD-APO-I) showed an interesting broad minimum covering the entire region defined as the intermediary and opened state, while simulations ABMD-APO-II and ABMD-APO-III showed more defined minimum at about 55°, consistent with an opened state. Together, these simulations indicate a clear energy minimum covering opened or intermediary-to-opened state for the unbound enzyme.

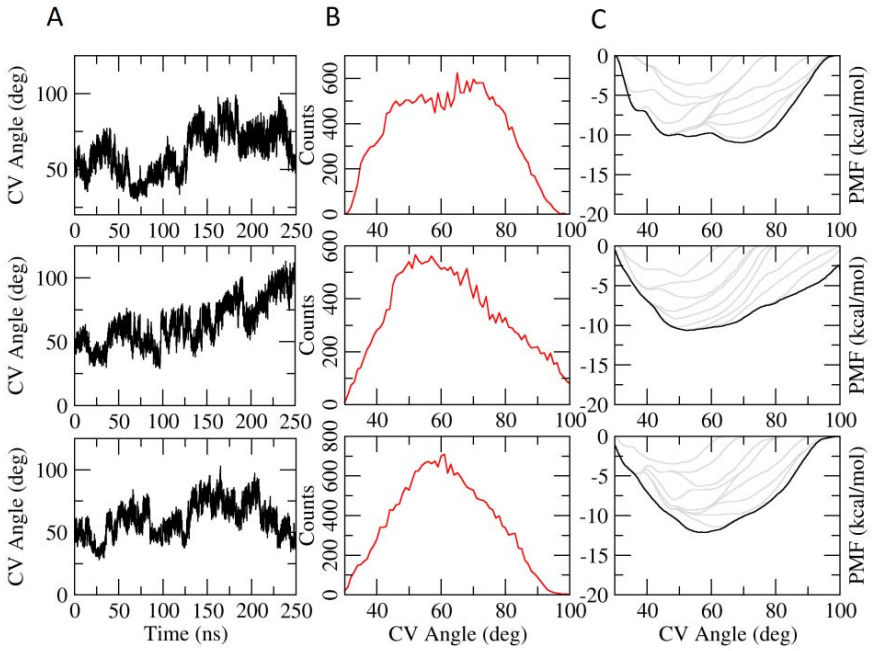


Figure 8. ABMD of the APO system for simulations ABMD-APO-I (upper row), ABMD-APO-II (middle row) and ABMD-APO-III (bottow row). (A) Time evolution of the CV angle; (B) CV angle distribution; C) Potential of mean

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

force. The gray lines indicate the time evolution (and convergence) of the PMF in 10 step ABMD simulations of 25 ns each.

For the HOLO system, ABMD simulations accessed a wide range of CV angles, including the opened state. Again, once forced towards higher CV angles, the substrate was dissociated from the enzyme structure, resulting in the unbound substrate and the opened enzyme. The correlation between unbinding and opening becomes clear when Figure 9 is compared with Figure S1 (supplementary material), that shows the interaction energy between the enzyme and substrate. The simulation data indicate that the presence of the ligand is crucial for holding the closed conformation stable and, once the conformation is destabilized (by external forces, in this case), the ligand is spontaneoulsy dissociated from the enzyme.

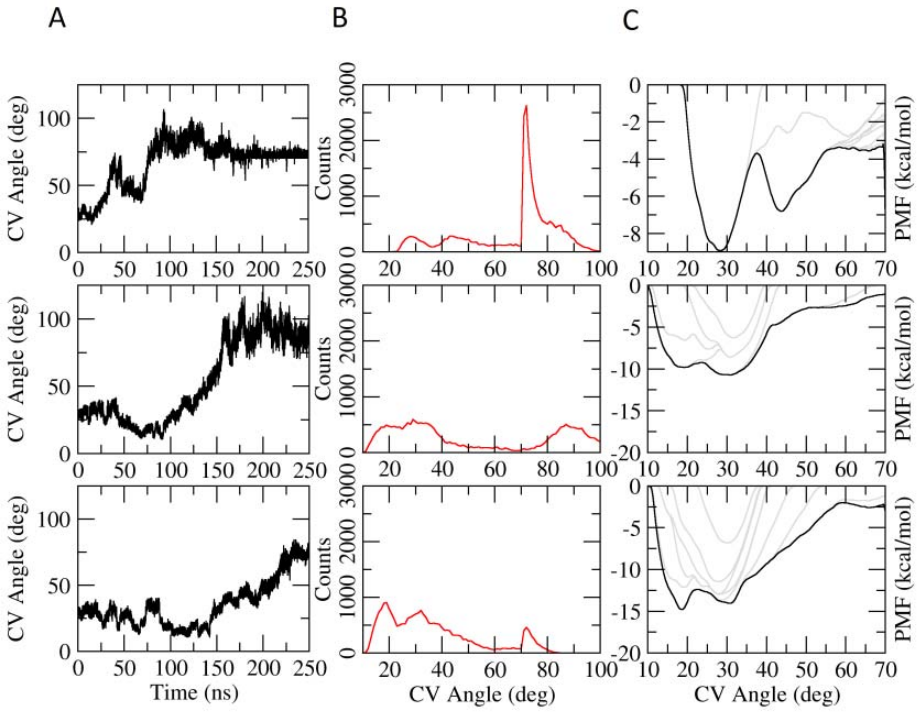


Figure 9. ABMD of the HOLO system for simulations ABMD-HOLO-I (upper row), ABMD-HOLO-II (middle row) and ABMD-HOLO-III (bottom row). (A) CV angle evolution during the simulation. (B) Histogram of CV angle. (C) Potential of mean force computed for the ABMD simulation data. The gray lines indicate the time evolution (and convergence) of the PMF in 10 step ABMD simulations of 25 ns each.

An important question remains unanswered: what is the mechanism involved in ligand binding and enzyme domain movement? Or, what is the nature of the driving force for the

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

domain movement observed in the presence of a substrate? A visual analysis of the enzyme structure reveals an extensive network of positively charged amino acids in the vicinity of the substrate pocket. These residues stabilize the interaction with the phosphate groups of the UDP-GlcNac, thus making the electrostatic contribution of the binding energy as the largest contribution for binding. In the other hand, in the absence of a substrate, or product, the interaction of these positively charged residues is very unfavorable and could be alleviated if the residues are moved apart in a domain movement. Thus, we hypothesized a major driving force for this large-scale movement may be the electrostatic energy associated with the interaction of positively charged residues.

In order to evaluate this hypothesis, the APO and HOLO structures were analyzed using the Delphi Force web server [32] to assess the magnitude of electrostatic forces between the N-terminal and C-terminal domains of SaMnaA. For the HOLO enzyme, the ligands were set as part of the C-terminal domain. The computed net forces are represented as arrows in protein structure in Figure 10. Here we observe that, in the absence of the substrate/cofactor, a net force vector is observed as pointing towards the direction of the domain movement (orange arrow in Figure 10). In the presence of the substrate, the net force is shifted to the twisting movement between the domains (also observed in the equilibrium simulation for the HOLO system). So, the analysis of the electrostatic forces provides a piece of evidence favoring the electrostatic term as a major driving force for the domain movement.

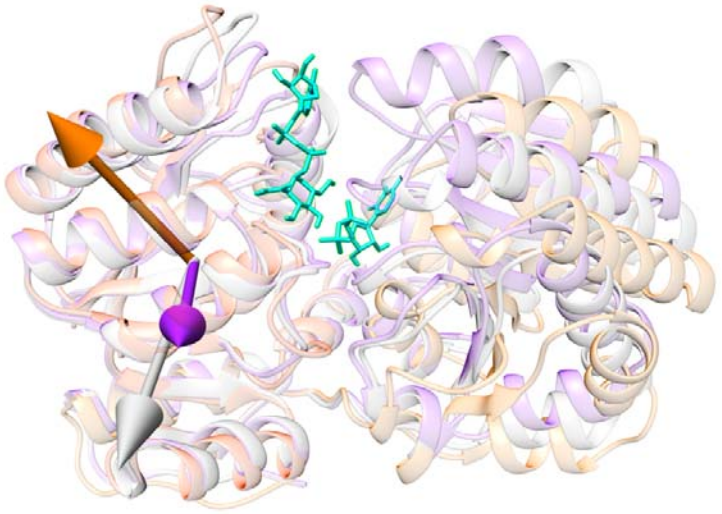


Figure 10 – The sum of the resulting interaction electrostatic forces between the C-terminal and N-terminal domains for the open (orange), intermediary (gray) and closed (purple) states of the protein, with ligands in the active site (cyan).

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

In this work, we used over 10 µs of simulation data to evaluate the mechanisms involved in the structural rearrangement of *S. aureus* UDP-N-acetylglucosamine 2-epimerase as a function of the binding of the substrate (Table 1). Our equilibrium and non-equilibrium simulation data showed that the enzyme is remarkably stabilized in a closed conformation when bound to substrate and cofactor, while the unbinding event leads to a change in the enzyme equilibrium structure towards a more opened structure. The analysis of the internal forces for the opened and closed conformation strongly suggested that the electrostatic term may be a major driving force for the enzyme opening, alleviating the unfavorable interactions between positively charged residues in the absence of the electron-rich substrate.

This shift in the energy landscape, caused by the change in the environment (substrate in the active site) agrees with the previously proposed concept of a dynamic energy landscape [13], which is an important feature for the protein function since the macromolecule can adjust its population accordingly to changes in the environment, such as pH shifts or the presence of substrates. Kumar and coworkers [13] proposed two main types of dynamic energy landscapes. The Type I is characterized by an environment change that causes a shift in the energy minimum of the energy landscape, thus shifting the preferential conformation of the protein. Type II, on the other hand, describes the situation where the change in the environment causes a change of the roughness in the surface, where the conformations that were trapped in an energy minimum now have a lower barrier to overcome and change into another conformation. Within these definitions, the dynamic landscape observed for SaMnaA seems to be closer to type I, that is usually observed in enzymes that are allosterically regulated, or in multimolecular complexes.

SaMnaA may be an important target for the development of drug candidates to fight resistant infections. The molecular simulations and previous data show that the ligands may enter the active site when the protein is in the open/intermediate conformation, being closed by the interactions made between the molecules and the protein. Also, positively charged residues are involved in the most energetically favorable interactions. This indicates that the open conformation should be targeted and the electrostatic forces should be taken into account when screening for new drug candidates.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conclusion

In this work, we explored the dynamics of the *S. aureus* MnaA that plays an important role in the biosynthesis of the wall teichoic acid and in the cellular resistance to β-lactam antibiotics. The protein has an open/close movement, important for the maintenance of the substrate in the active site during the catalysis. This movement is mediated by the ligands, which interact with positive amino acids of the protein, shifting the energy profile for the system. The dynamic energy landscape for the systems APO and HOLO suggest that there is a population of preferential states that shifts with the change in the environment. Finally, a structure-based mechanism is proposed, providing a rationale for the domain movement observed for SaMnaA.

Acknowledgements

We thank the fruitful contributions made by prof. Milton T. Sonoda (*in memorium*). The authors thank the financial support provided by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) through grants 2017/18173-0, 2015/26722-8 and 2015/13684-0 and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), through grant # 303165/2018-9. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. McMillan K, Moore SC, McAuley CM, Fegan N, Fox EM. Characterization of Staphylococcus aureus isolates from raw milk sources in Victoria, Australia. BMC Microbiol. BioMed Central; 2016;16: 169. doi:10.1186/s12866-016-0789-1

2. Peacock SJ, Paterson GK. Mechanisms of Methicillin Resistance in Staphylococcus aureus. Annu Rev Biochem. 2015;84: 577–601. doi:10.1146/annurev-biochem-060614-034516

3. L.M. W, A.K. W, B. L, M.A. D, J. P, A.J. K, et al. Antimicrobial-Resistant Pathogens Associated with Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014. Infect Control Hosp Epidemiol. Centers for Disease Control; 2016;37: 1288–1301. doi:10.1086/668770

4. Farha MA, Leung A, Sewell EW, D'Elia MA, Allison SE, Ejim L, et al. Inhibition of WTA synthesis blocks the cooperative action of pbps and sensitizes MRSA to??-lactams. ACS Chem Biol. 2013;8: 226–233. doi:10.1021/cb300413m

5. Lowy FD. *Staphylococcus aureus* Infections. N Engl J Med. Massachusetts Medical Society ; 1998;339: 520–532. doi:10.1056/NEJM199808203390806

6. Sewell EWC, Brown ED. Taking aim at wall teichoic acid synthesis: new biology and new leads for antibiotics. J Antibiot (Tokyo). Nature Publishing Group; 2014;67: 43–51. doi:10.1038/ja.2013.100

7. de Kraker MEA, Davey PG, Grundmann H, BURDEN study group on behalf of the B study. Mortality and hospital stay associated with resistant Staphylococcus aureus and Escherichia coli bacteremia: estimating the burden of antibiotic resistance in Europe. PLoS Med. Public Library of Science; 2011;8: e1001104. doi:10.1371/journal.pmed.1001104

8. Sobhanifar S, Worrall LJ, King DT, Wasney GA, Baumann L, Gale RT, et al. Structure and Mechanism of Staphylococcus aureus TarS, the Wall Teichoic Acid β-glycosyltransferase Involved in Methicillin Resistance. PLOS Pathog. 2016;12: e1006067. doi:10.1371/journal.ppat.1006067

9. Roemer T, Schneider T, Pinho MG. Auxiliary factors: A chink in the armor of MRSA

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

resistance to β-lactam antibiotics. Curr Opin Microbiol. 2013;16: 538–548. doi:10.1016/j.mib.2013.06.012

10. Brown S, Xia G, Luhachack LG, Campbell J, Meredith TC, Chen C, et al. Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. Proc Natl Acad Sci U S A. 2012;109: 1–6. doi:10.1073/pnas.1209126109

11. Mann PA, Müller A, Wolff KA, Fischmann T, Wang H, Reed P, et al. Chemical Genetic Analysis and Functional Characterization of Staphylococcal Wall Teichoic Acid 2-Epimerases Reveals Unconventional Antibiotic Drug Targets. PLOS Pathog. 2016;12: e1005585. doi:10.1371/journal.ppat.1005585

12. Velloso LM, Bhaskaran SS, Schuch R, Fischetti VA, Stebbins CE. A structural basis for the allosteric regulation of non-hydrolysing UDP-GlcNAc 2-epimerases. EMBO Rep. 2008;9: 199–205. doi:10.1038/sj.embor.7401154

13. Kumar S, Ma B, Tsai C, Sinha N, Nussinov R. Folding and binding cascades: Dynamic landscapes and population shifts. Protein Sci. 2000;9: 10–19.

14. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput. 2015;11: 3696–3713. doi:10.1021/acs.jctc.5b00255

15. Anandakrishnan R, Aguilar B, Onufriev A V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res. Oxford University Press; 2012;40: W537–W541. doi:10.1093/nar/gks375

16. Bashford D, Karplus M. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry. American Chemical Society; 1990;29: 10219–10225. doi:10.1021/bi00496a010

17. Price DJ, Brooks CL. A modified TIP3P water potential for simulation with Ewald summation. J Chem Phys. 2004;121: 10096–10103. doi:10.1063/1.1808117

18. Rassolov VA, Ratner MA, Pople JA, Redfern PC, Curtiss LA. 6-31G* basis set for third-row atoms. J Comput Chem. Wiley-Blackwell; 2001;22: 976–984. doi:10.1002/jcc.1058

19. Comell WD, Cieplak P, Bayly CI, Kollman PA. Application of RESP Charges To Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. J Am

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Chem Soc. 1993;115: 9620–9631.

20. Wang J, Wang W, Kollman PA, Case DA. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations Correspondence to. Journal of Chemical Information and Computer Sciences.

21. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem. Wiley-Blackwell; 2004;25: 1157–1174. doi:10.1002/jcc.20035

22. Šali A, Blundell TL. Comparative Protein Modelling by Satisfaction of Spatial Restraints. J Mol Biol. 1993;234: 779–815. doi:10.1006/jmbi.1993.1626

23. Webb B, Sali A. Comparative protein structure modeling using MODELLER. Curr Protoc Bioinforma. 2016; doi:10.1002/cpbi.3

24. Case DA, Cheatham III TE, Darden T, Gohlke H, Luo R, Merz Jr. KM, et al. The Amber Biomolecular Simulation Programs. J Comput Chem. 2005;26: 1668–1688.

25. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. WIREs Comput Mol Sci. 2013;3: 198–210. doi:10.1002/wcms.1121

26. Miao Y, Feher VA, McCammon JA. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. J Chem Theory Comput. 2015;11: 3584–3595. doi:10.1021/acs.jctc.5b00436

27. Miao Y, Feher VA, McCammon JA. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. J Chem Theory Comput. 2015;11: 3584–3595. doi:10.1021/acs.jctc.5b00436

28. Miao Y, Sinko W, Pierce L, Bucher D, Walker RC, Mccammon JA. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation - Journal of Chemical Theory and Computation (ACS Publications). J Chem Theory Comput. 2014;10: 2677–2689.

29. Babin V, Karpusenka V, Moradi M, Roland C, Sagui C. Adaptively biased molecular dynamics: An umbrella sampling method with a time-dependent potential. Int J Quantum Chem. 2009;109: 3666–3678. doi:10.1002/qua.22413

30. Barducci A, Bussi G, Parrinello M. Well-tempered metadynamics: A smoothly

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

converging and tunable free-energy method. Phys Rev Lett. 2008;100: 1–4. doi:10.1103/PhysRevLett.100.020603

31. Scherer MK, Trendelkamp-Schroer B, Paul F, Peez-Hernadez G, Hoffmann M, Plattner N, et al. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. doi:10.1021/acs.jctc.5b00743

32. Li L, Jia Z, Peng Y, Chakravorty A, Sun L, Alexov E, et al. DelPhiForce web server: electrostatic forces and energy calculations and visualization. Valencia A, editor. Bioinformatics. Oxford University Press; 2017;33: 3661–3663. doi:10.1093/bioinformatics/btx495

33. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput. 2013;9: 3084–3095. doi:10.1021/ct400341p

34. Nascimento AS. AmberEnergy++. In: AmberEnergy [Internet]. 2018 [cited 13 Dec 2018]. Available: https://github.com/alessandronascimento/amberenergy

35. Chen S-C, Huang C-H, Shin Yang C, Liu J-S, Kuan S-M, Chen Y. Crystal structures of the archaeal UDP-GlcNAc 2-epimerase from Methanocaldococcus jannaschii reveal a conformational change induced by UDP-GlcNAc. Proteins Struct Funct Bioinforma. Wiley-Blackwell; 2014;82: 1519–1526. doi:10.1002/prot.24516

36. Laskowski RA, Swindells MB. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. J Chem Inf Model. American Chemical Society; 2011;51: 2778–2786. doi:10.1021/ci200227u

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure Captions

Figure 1. (A) Angles observed during the simulation time for the 1-µs simulation of the APO SaMnaA structure. (B) Histogram of observed angles during the 1-µs simulation of the APO SaMnaA. (C) Angles observed during the simulation of the HOLO SaMnaA structure. (D) Histogram of observed angles in HOLO SaMnaA simulation. (E) Angles observed during the 5-µs simulation of the APO SaMnaA structure. (F) Histogram of the observed angles during the 5-µs APO SaMnaA simulation.

Figure 2. Projections of the first (A) and second (B) components of the PCA analysis of the APO SaMnaA system MD trajectory.

Figure 3. (A) Comparison between the CV angle measured (upper plot) and the TIC 1 from the TICA analysis (lower plot). (B) Heatmap of TIC1 versus TIC2 where the systems stated were coarse-grained, resulting in the separation of three macrostates, shown here as cluster defined by colors red, green and blue. (C) Representative structures of the three states found for the system, with atoms used for CV definition atoms in evidence.

Figure 4. Projections of the first (A) and second (B) components of the PCA analysis for the HOLO SaMnaA system MD trajectory.

Figure 5. LIGPLOT depiction of the interactions between UDP (A) and UDP-GlcNac (B) with SaMnaA. Total interaction (potential) energy between UDP (C) and UDP-GlcNac (D) and its interacting amino acids are also shown.

Figure 6. GAMD simulation for APO SaMnaA structure. The results from three simulations are shown: GAMD-APO-I in black curve, GAMD-APO-II in red curve and GAMD-APO-III in blue curve. (A) Observed CV angle *versus* simulation time. (B) Distribution of observed CV angles during the simulation. C) PMF computed after reweighting.

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 7. GAMD simulation data for the HOLO SaMnaA. The results from the three independent simulations are shown: GAMD-HOLO-I in black curve, GAMD-HOLO-II in red and GAMD-HOLO-III in blue curve. (A) Observed CV angles during the simulation time. (B) Histogram of the observed angles during the simulation. (C) Computed PMFs from the energy reweighting. (D) RMSD calculated for the UDP (black) and UDP-GlcNac (red). (E) Total interaction energies computed for UDP (black) and UDP-GlcNac (red) with the protein.

Figure 8. ABMD of the APO system for simulations ABMD-APO-I (upper row), ABMD-APO-II (middle row) and ABMD-APO-III (bottow row). (A) Time evolution of the CV angle; (B) CV angle distribution; C) Potential of mean force. The gray lines indicate the time evolution (and convergence) of the PMF in 10 step ABMD simulations of 25 ns each.

Figure 9. ABMD of the HOLO system for simulations ABMD-HOLO-I (upper row), ABMD-HOLO-II (middle row) and ABMD-HOLO-III (bottom row). (A) CV angle evolution during the simulation. (B) Histogram of CV angle. (C) Potential of mean force computed for the ABMD simulation data. The gray lines indicate the time evolution (and convergence) of the PMF in 10 step ABMD simulations of 25 ns each.

Figure 10. The sum of the resulting interaction electrostatic forces between the C-terminal and N-terminal domains for the open (orange), intermediary (gray) and closed (purple) states of the protein, with ligands in the active site (cyan).

bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tables

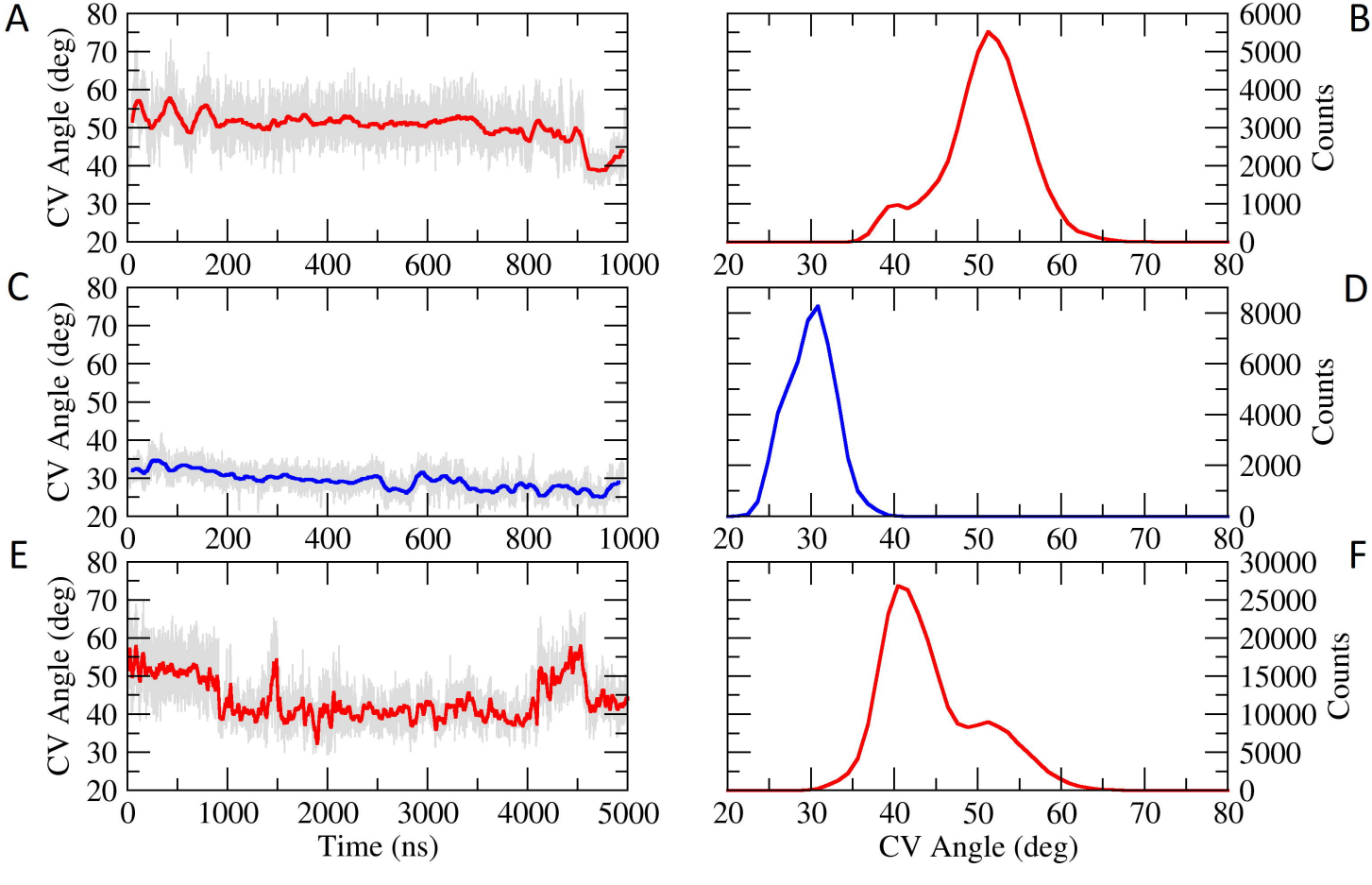
Table 3. Simulations used in this work. The simulation methods and times used are shown.

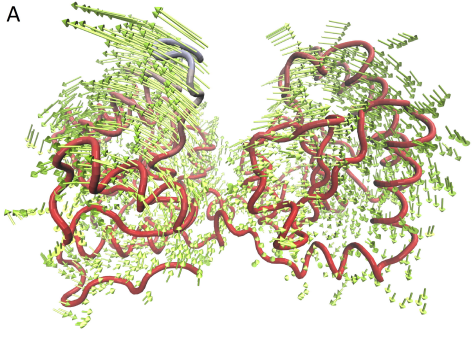
Method	Protein System	Time (ns)
Equilibrium	APO	1,000 + 4,000
Equilibrium	HOLO	1,000
GAMD	APO	3 x 500
GAMD	HOLO	3 x 250
ABMD	APO	3 x 250
ABMD	HOLO	3 x 250

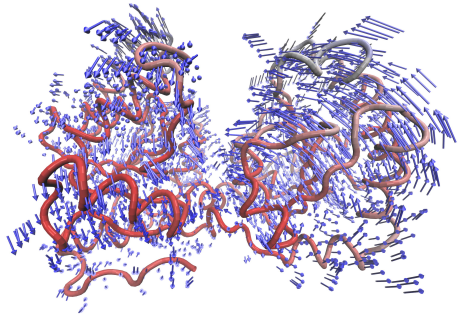
bioRxiv preprint doi: https://doi.org/10.1101/535138; this version posted January 30, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

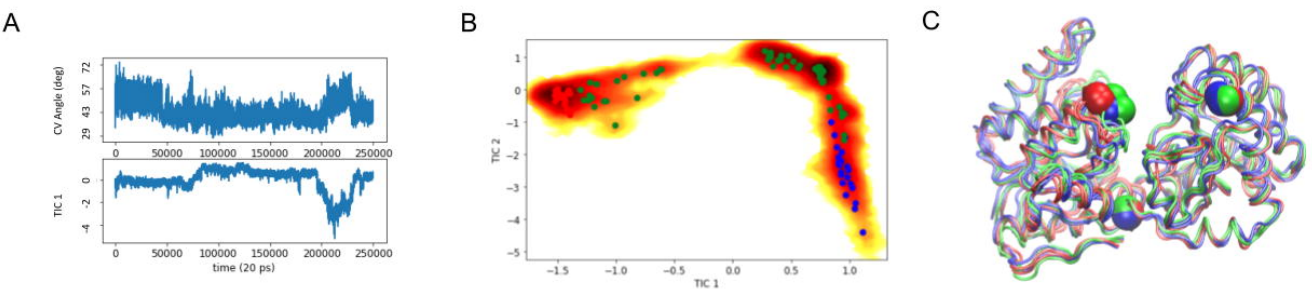
Table 4 – Opening state of known crystal structures of UDP-GlcNac 2-epimerases. *chain without ligands

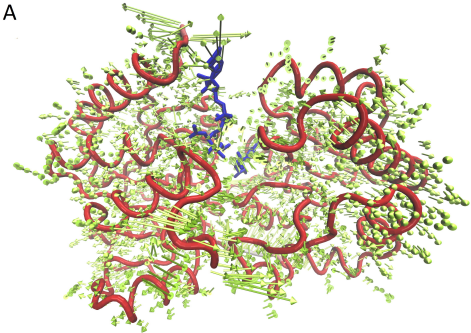
PDB ID	Organism	Ligands	Oligomeric state	CV Angle Chain A (deg)	CV Angle Chain B (deg)
		Enzymes in the unbound state			
4NEQ	*M. jannaschii*	-	Monomer	53	-
1V4V	*T. thermophilus*	-	Dimer	53	40
3OT5	*L. monocytogenes*	-	Dimer	47	33
1O6C	*B. subtilis*	-	Dimer	53	58
4HWG	*R. bellii*	-	Monomer	25	-
3DZC	*V. cholerae*	-	Dimer	48	46
		Enzymes bound to UDP and UDP-GlcNac			
1VGV	*E. coli*	UDP-GlcNac	Dimer	40*	25
3BEO	*E. coli*	UDP, UDP-GlcNac	Dimer	30	30
4FKZ	*B. subtilis*	UDP, UDP-GlcNac	Dimer	25	25
4NES	*M. jannaschii*	UDP, UDP-GlcNac	Monomer	29	-
5DLD	*B. vietnamiensis*	UDP, UDP-GlcNac	Monomer	24	-
		Enzyme bound to UDP only			
5ENZ	*S. aureus*	UDP	Dimer	31	26
1F6D	*E. coli*	UDP	Dimer	40	27

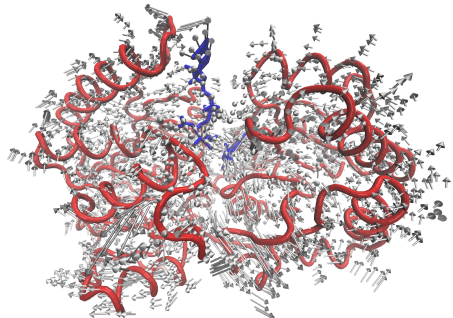

A
B
C
D
E
F
CV Angle (deg)
Counts
Time (ns)
CV Angle (deg)


A

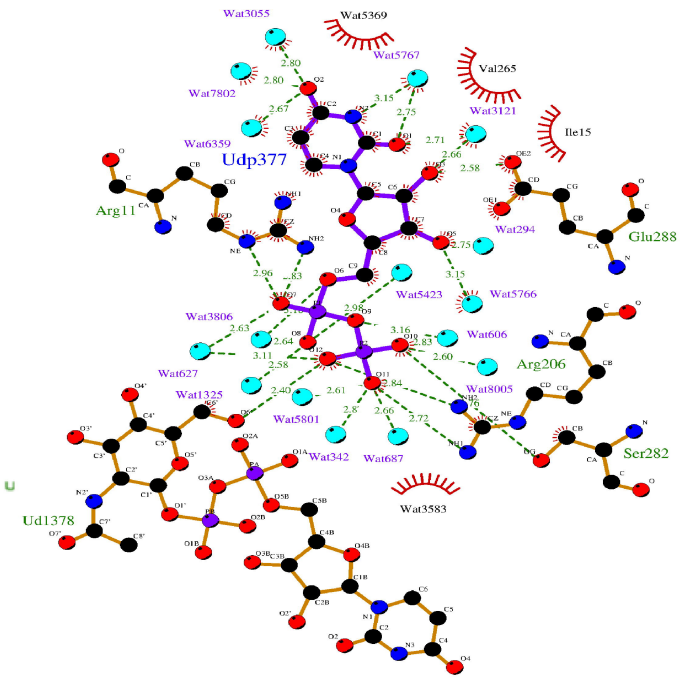

B


A
CV Angle (deg)
72
57
43
29
0
50000
100000
150000
200000
250000
TIC 1
0
-2
-4
time (20 ps)
B
TIC 2
1
0
−1
−2
−3
−4
−5
−1.5
−1.0
−0.5
0.0
0.5
1.0
TIC 1
C

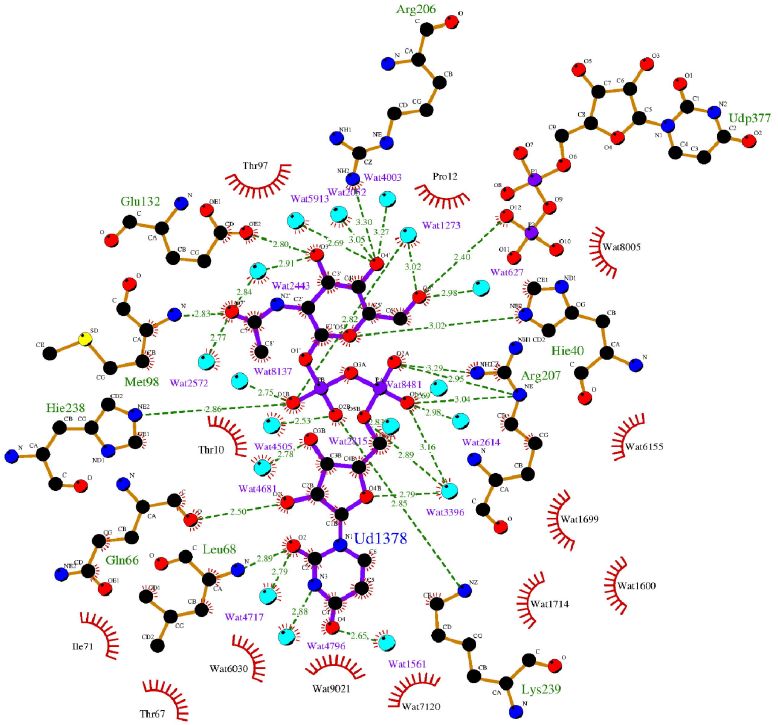

A


B

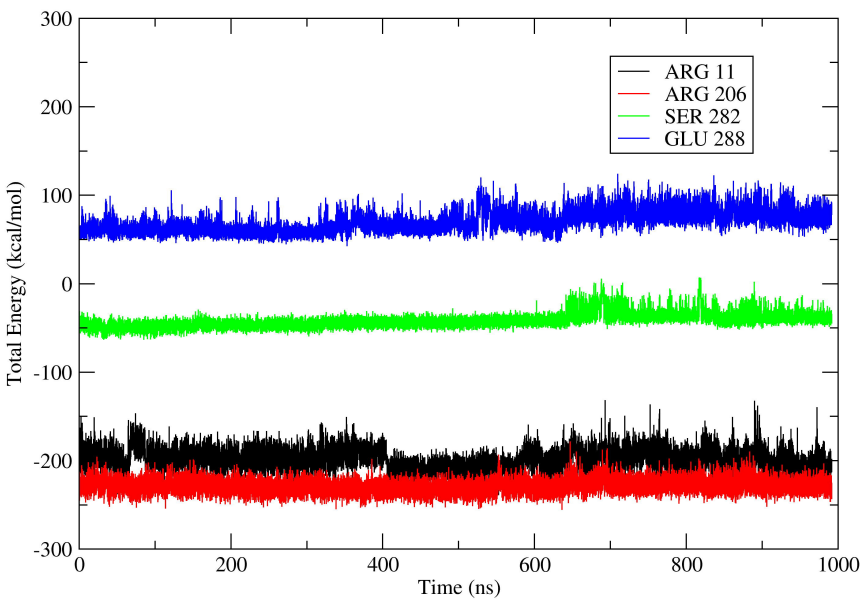
A


Udp377
Ud1378
Arg11
Glu288
Arg206
Ser282
Ile15
Val265
Wat3055
Wat5369
Wat5767
Wat7802
Wat6359
Wat3121
Wat294
Wat5423
Wat5766
Wat3806
Wat606
Wat8005
Wat627
Wat1325
Wat5801
Wat342
Wat687
Wat3583

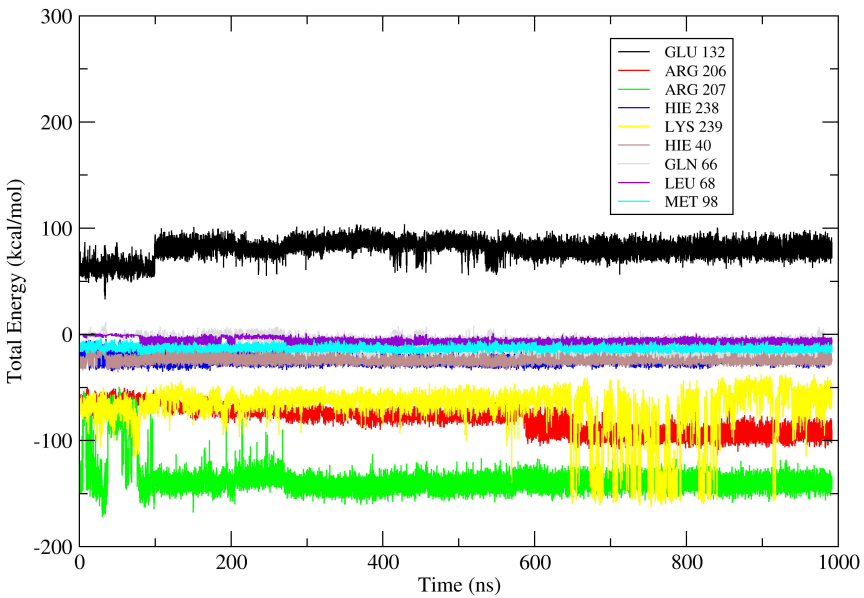
B

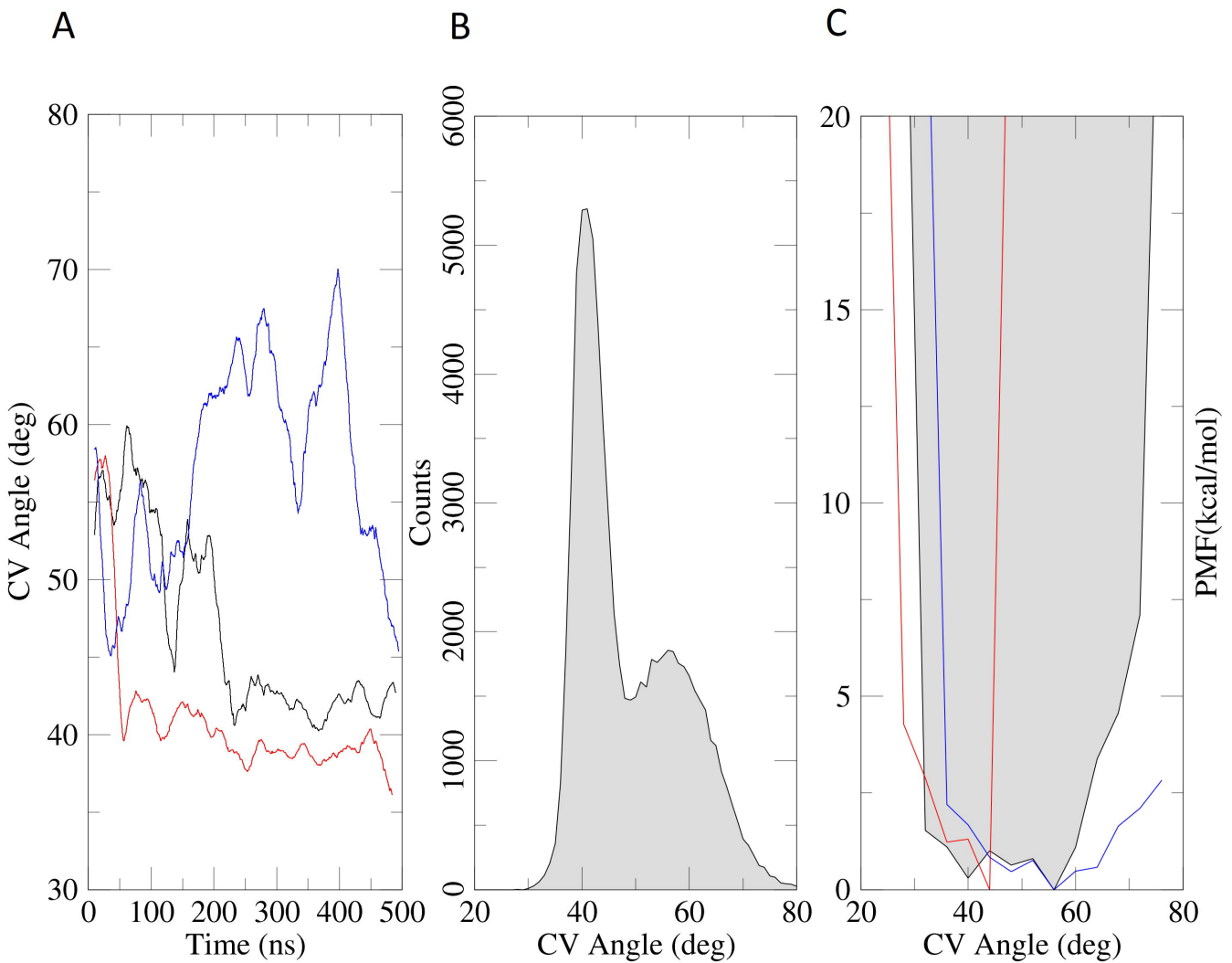

Arg206
Udp377
Thr97
Pro12
Glu132
Wat4003
Wat5913
Wat1273
Wat8005
Wat627
Wat2443
Hie40
Arg207
Met98
Wat2572
Wat8137
Wat8481
Hie238
Thr10
Wat4505
Wat2155
Wat2614
Wat6155
Wat4681
Wat3396
Wat1699
Ud1378
Leu68
Gln66
Wat1600
Wat1714
Wat4717
Ile71
Wat4796
Wat6030
Wat9021
Wat1561
Wat7120
Lys239
Thr67

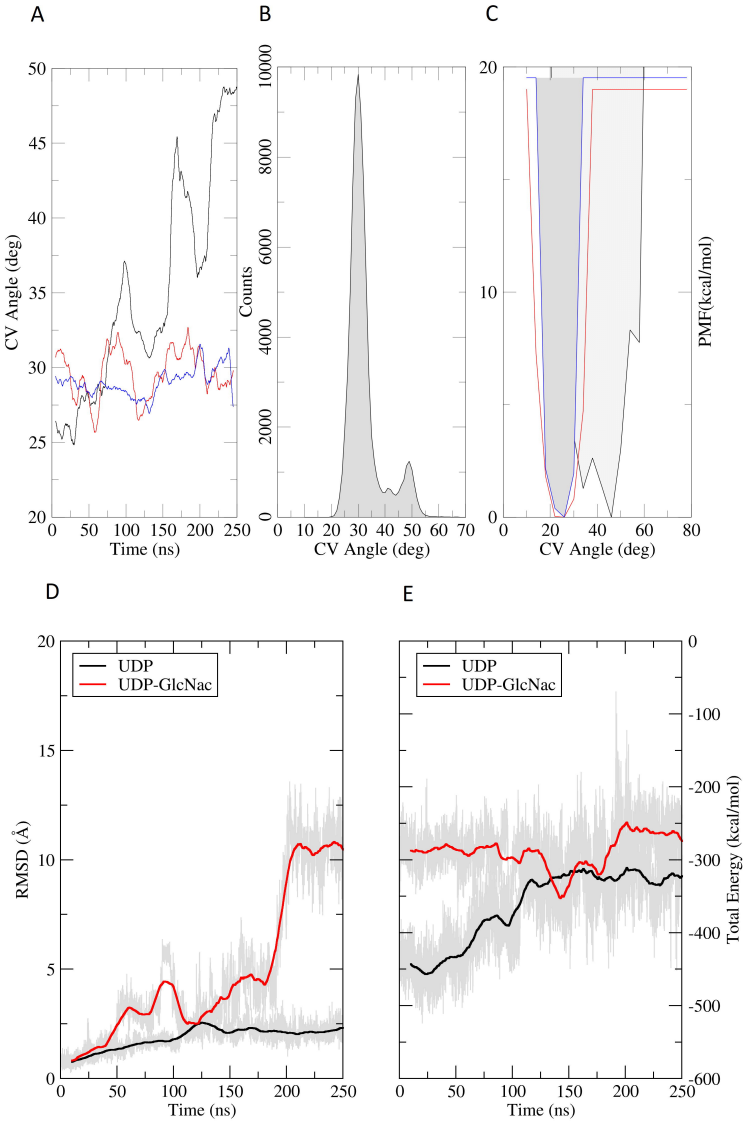
C

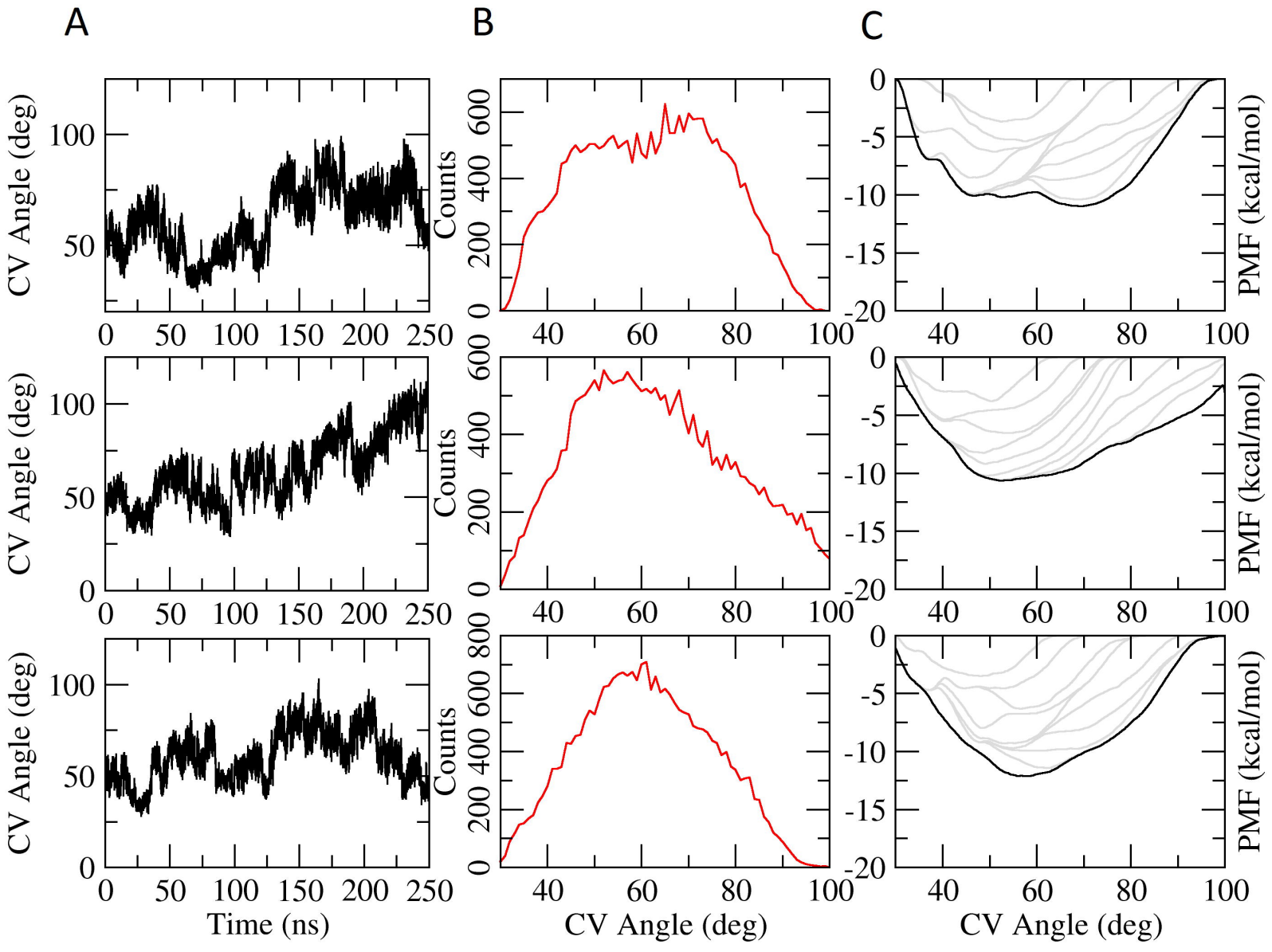

ARG 11
ARG 206
SER 282
GLU 288
Total Energy (kcal/mol)
Time (ns)
300
200
100
0
-100
-200
-300
0
200
400
600
800
1000

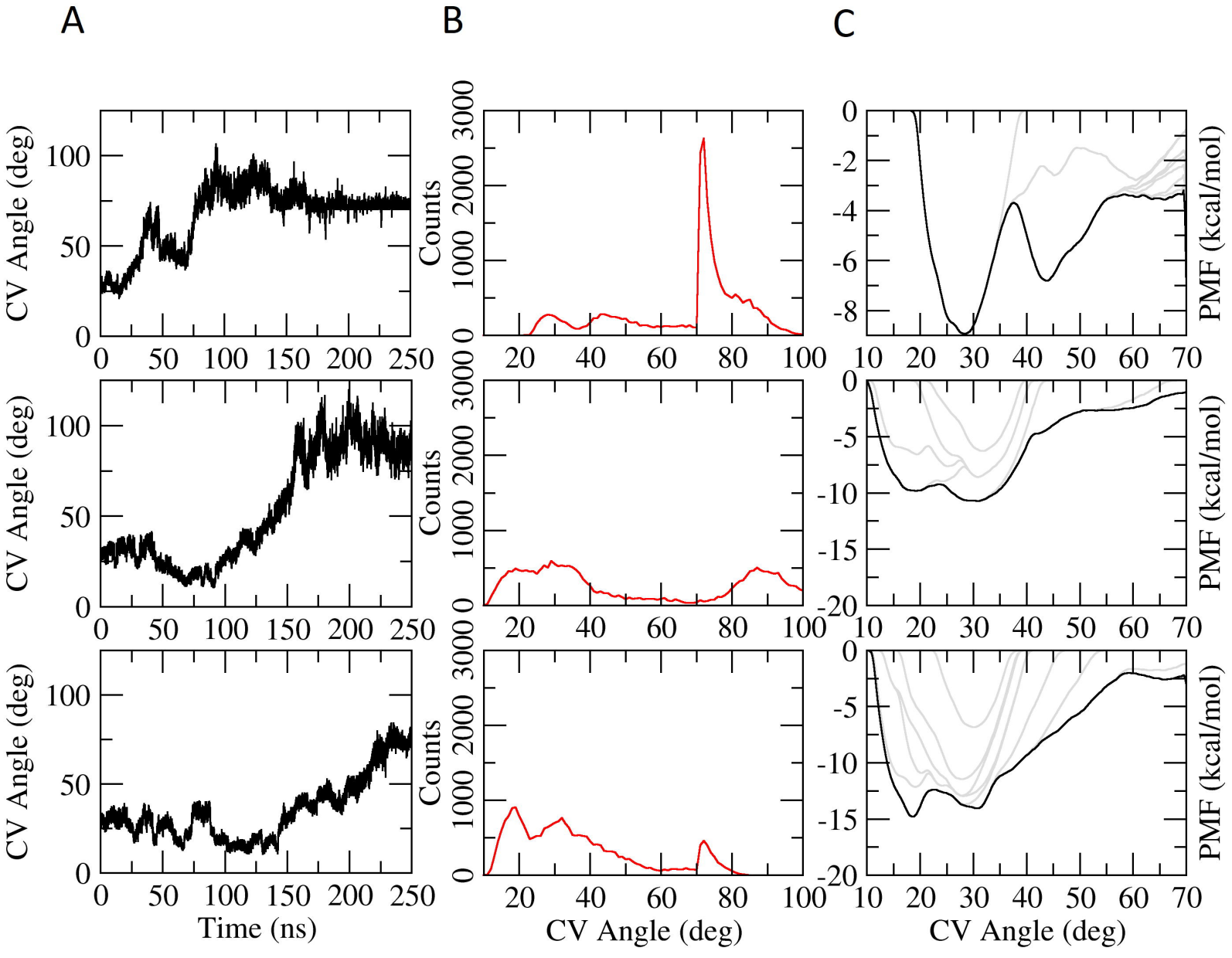
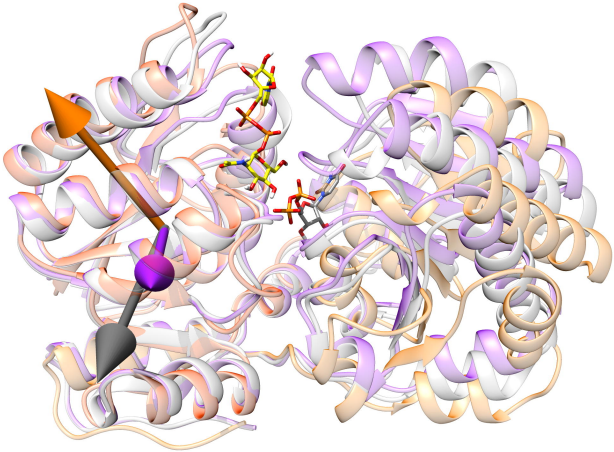
D


GLU 132
ARG 206
ARG 207
HIE 238
LYS 239
HIE 40
GLN 66
LEU 68
MET 98
Total Energy (kcal/mol)
Time (ns)
300
200
100
0
-100
-200
0
200
400
600
800
1000


A
B
C
CV Angle (deg)
Time (ns)
Counts
CV Angle (deg)
PMF(kcal/mol)
CV Angle (deg)


A
CV Angle (deg)
Time (ns)
B
Counts
CV Angle (deg)
C
PMF(kcal/mol)
CV Angle (deg)
D
UDP
UDP-GlcNac
RMSD (Å)
Time (ns)
E
UDP
UDP-GlcNac
Total Energy (kcal/mol)
Time (ns)


A
B
C
CV Angle (deg)
Counts
PMF (kcal/mol)
Time (ns)
CV Angle (deg)
CV Angle (deg)


A
B
C
CV Angle (deg)
Counts
PMF (kcal/mol)
Time (ns)
CV Angle (deg)
CV Angle (deg)