Multiple functions of aromatic-carbohydrate interactions in a processive cellulase examined with molecular simulation

Christina M. Payne[1], Yannick J. Bomble[1], Courtney B. Taylor[2], Clare McCabe[2,3], Michael E. Himmel[1], Michael F. Crowley[1], Gregg T. Beckham[4,5]

1. Biosciences Center, National Renewable Energy Laboratory, Golden CO 80401
2. Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville TN 37235
3. Department of Chemistry, Vanderbilt University, Nashville TN 37235
4. National Bioenergy Center, National Renewable Energy Laboratory Golden CO 80401
5. Department of Chemical Engineering, Colorado School of Mines, Golden CO 80401

SUPPLEMENTAL DATA

COMPUTATIONAL METHODS

This section describes the preparation of the systems simulated in this study, the molecular dynamics (MD) simulations, the thermodynamic integration methods, and the error analysis for the thermodynamic integration simulations.

System Preparation and MD Simulations.

CHARMM (1) was used to build, minimize, and equilibrate the protein structure from the wild type crystal structure, 1QK2 (2). The following residues were protonated in all simulations: E107, D170, D221, H331, H340, H414, and E419. The enzyme was minimized for 1,000 steps using steepest descent minization followed by 1,000 steps of conjugate gradient minimization. The enzyme was solvated, and sodium ions were added for electrostatic neutrality. The solvated system was minimized for 1,000 steps with steepest descent minimization holding all but the water and ions fixed, followed by another 1,000 steps of steepest descent minimization holding only the ligand fixed (when present). A final minimization of the entire system was run for 1,000 steps of steepest descent minimization followed by 1,000 steps of Adopted Basis Newton-Raphson minimization.

The resulting minimized systems were equilibrated in the *NVE* ensemble at 300 K for 20 ps followed by another 20 ps in the *NPT* ensemble at 300 K for density equilibration. The parameters given here were used in all the MD simulations. Temperature control was performed using the Nosé-Hoover thermostat at 300 K (3,4). SHAKE was used to fix the distances to hydrogen atoms (5). Non-bonded interactions were truncated with a 10 Å cutoff, and the Particle Mesh Ewald method (6) with a 6th order b-spline, a Gaussian distribution width of 0.320 Å, and a mesh size 80 x 80 x 80 was used to describe the electrostatics. All equilibration simulations used a 2 fs time step. The CHARMM27 force field with the CMAP correction (1,7,8) was used to describe the protein, while glycans and cellulose used the CHARMM35 carbohydrate force field (9-11). Water was modeled using the TIP3P force field (12,13).

The production MD simulations were performed using NAMD (14). The production simulations were performed in the *NVT* ensemble at 300 K for 250 ns using a 1 fs time step for stability in the SHAKE algorithm. Excepting the time step and the use of the Langevin thermostat (15), all simulation parameters described above for the CHARMM equilibration were maintained for the production simulations.

The system size evaluated was approximately 80 Å x 80 Å x 80 Å yielding a system of about 52,000 atoms, of which roughly 46,000 atoms are water. The system dimensions were chosen to ensure that the enzyme, including the glycans, were sufficiently solvated so as not to interact with the periodic image. The catalytically-active systems, referred to here as “bound” systems, included a cellohexaose oligomer in the active site of the enzyme, positioned so as the two product sites were occupied. Systems referred to as “free” contain the enzyme excluding the cellooligomer. *O*-glycosylation of the enzyme was taken from the 1QK2 structure (2), while the *N*-glycosylation was constructed as described by Hui *et al*. (16).

Thermodynamic Integration.
Using a snapshot from the production simulations at 18 ns, TI calculations were performed using NAMD with the dual-topology methodology (17-19). The choice of 18 ns as a snapshot was a function of the completion point of the first run of the MD simulation trajectory. Any point at which the simulation is fully equilibrated may be used as the starting point. This methodology entails the equilibration of a single system containing a "hybrid" residue at the mutation site in which the derivative of the potential energy, *U*, is calculated over the range of a coupling parameter, λ. This hybrid residue contains atoms from both the wild type residue and the mutant, wherein the wild type and mutant atoms do not interact with one another, and interact with the rest of the system via standard bonded and nonbonded interactions scaled by λ from the reactant (wild type) to the product state (mutant) in windows over λ. Electrostatic and van der Waals calculations were decoupled into two separate processes, and soft-core potentials were used to overcome endpoint singularities (20).

The electrostatic and van der Waals calculations included 15 windows ranging from λ values of 0 to 1 for a total of 30 simulations per mutation. The windows were divided as follows: 0.0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.85, 0.9, 0.95, 1.0, using more closely coupled windows near the endpoints to improve accuracy. The electrostatics and van der Waals calculations were equilibrated for 0.5 ns prior to collection of 14.5 ns of TI data.

The change in free energy, ΔG, for each set of simulations was determined using Simpson's Rule with Brun's extension for non-equidistant nodes to integrate $dU/d\lambda$ over $\lambda = 0$ to 1 (21). The electrostatic and van der Waals terms were combined to obtain ΔG for the bound and free states of each point mutation as described in Equation S1, where Mut refers to the mutant and WT refers to the wild type. This scheme is illustrated in the thermodynamic cycle shown in Figure S1.

$$\Delta\Delta G = \left(G_{Mut-bound} - G_{Mut-free}\right) - \left(G_{WT-bound} - G_{WT-free}\right) = \left(G_{Mut-bound} - G_{WT-bound}\right) - \left(G_{Mut-free} - G_{WT-free}\right) \quad \text{(S1)}$$

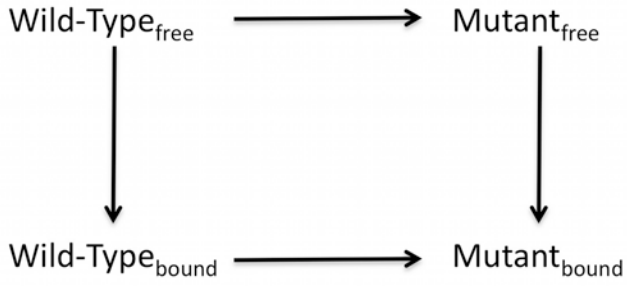

Figure S1. Alchemical cycle for measuring ΔΔG with thermodynamic integration. Free and bound refer to the absence and presence of the ligand, respectively.

Error analysis was performed following the methods of Steinbrecher *et al.* (22) in which the standard deviation, σ, for each window is related to the total window length, t_{MD}, and the autocorrelation time, τ, of $dU/d\lambda$ as given in Equation S2.

$$\Delta\left(\frac{\partial U}{\partial \lambda}\right) = \frac{\sigma}{\sqrt{\left(t_{MD} / \ 2\tau\right)}} \quad \text{(S2)}$$

The total error from λ = 0 to 1 is weighted by the width of each window as in Equation S3.

$$\Delta_{tot} = \sum_i \frac{1}{2}\left(\lambda_{j+1} - \lambda_{j-1}\right)\Delta_i \qquad \text{(S3)}$$

Convergence Analysis.
Simulation analysis included the confirmation of appropriate window overlap and simulation convergence as described by Pohorille *et al*. (20). Figure S2A shows an example of a probability distribution analysis wherein histograms for each simulated window are plotted on the same *x*-axis, visually confirming acceptable overlap of the windows. Figure S2B shows an example plot of autocorrelation function (ACF) versus time. The autocorrelation time for each window was used in the error analysis. The trends seen in this example data set are representative of the thermodynamic data collected for all mutations in this study.

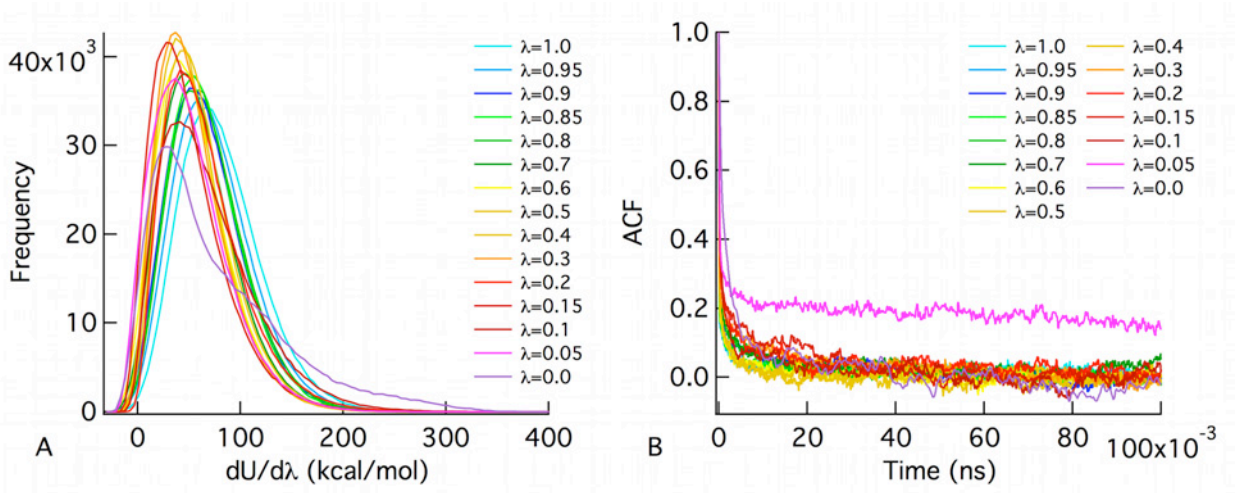

Figure S2. An example data set for window overlap and autocorrelation function analysis. These data are taken from the reactant state in the van der Waals simulation set from the free W272A calculation illustrating (A) window overlap and (B) calculation of the autocorrelation time.

Normal Mode Analysis.
Normal mode analysis was conducted with the molecular mechanics program package NAB (23,24) (now part of AmberTools 1.5) (25,26), using the parameter set parm99SB (27,28), and the pairwise approach of Hawkins *et al.* for the Generalized Born model (29,30). The monomeric Cel6A structure was minimized using the Limited-Memory Broyden–Fletcher–Goldfarb–Shanno Truncated Newton Conjugate minimization technique to obtain an RMS gradient below 1×10^{-8} kcal/mol-Å. This level of convergence is necessary to avoid contamination from translational and rotational modes into true internal modes. Diagonalization of the Hessian matrix was performed using ARPACK routines (31) in combination with Cholesky decomposition and inversion of the Hessian matrix, providing better separation of eigenvalues to enhance convergence. The cross-correlation map in Figure 5 was generated with a modified version of the analysis module PTRAJ.

ADDITIONAL RESULTS

Detailed Thermodynamic Integration Results.

The results shown in Table 1 in the main text are a summary of the information found below in Table S1. Electrostatic and van der Waals interaction energies are combined to obtain the ΔG for a given state of an alchemical mutation (i.e. bound or free). Taking the difference of the bound and free ΔG values gives ΔΔG.

Table S1: Detailed ligand binding free energies and associated binding affinities calculated from TI.

		Bound		Free	
		Energy [kcal/mol]	Error [kcal/mol]	Energy [kcal/mol]	Error [kcal/mol]
W135A	Electrostatics	8.89	0.00	7.24	0.07
	VDW	13.72	0.09	6.01	0.34
	ΔG	22.60	0.09	13.25	0.35
	ΔΔG [kcal/mol]	9.36 +/-0.36			
	K_{wt}/K_{mut}	6.5E+06			
W269A	Electrostatics	5.16	0.06	4.34	0.09
	VDW	0.22	0.48	2.33	0.30
	ΔG	5.38	0.48	6.68	0.32
	ΔΔG [kcal/mol]	-1.30 +/- 0.58			
	K_{wt}/K_{mut}	1.1E-01			
W272A	Electrostatics	5.63	0.06	4.42	0.10
	VDW	3.63	0.30	1.08	0.26
	ΔG	9.26	0.31	5.50	0.27
	ΔΔG [kcal/mol]	3.76 +/- 0.41			
	K_{wt}/K_{mut}	5.5E+02			
W367A	Electrostatics	6.73	0.05	5.90	0.05
	VDW	3.91	0.30	3.14	0.27
	ΔG	10.64	0.30	9.04	0.28
	ΔΔG [kcal/mol]	1.61 +/- 0.41			
	K_{wt}/K_{mut}	1.5E+01			

Molecular Dynamics – Interaction Energies and Hydrogen Bonding.

From the five separate 250 ns MD simulations of the individual mutants and the wild type enzyme, we calculated the interaction energies of the protein with the ligand (Figure S3A) and the hydrogen bonds between the ligand and protein (Figure S3B). Hydrogen bond cutoffs were defined as within 3.4 Å of the donor and acceptor and 60° from linear. Error bars in Figures S3A and S3B represent one standard deviation.

The results in Figure S3A and in Figure S3B, within the confines of standard deviation, show there is no significant change in either interaction energy or average number of hydrogen bonds at a given binding site as a result of mutating a tryptophan to an alanine making conclusions difficult to draw as to the relationship of these dynamical properties in relative ligand binding free energy. It should be noted, however, that the interaction energy shown in Figure S3A is that of the entire protein with the glucose of a particular active site.

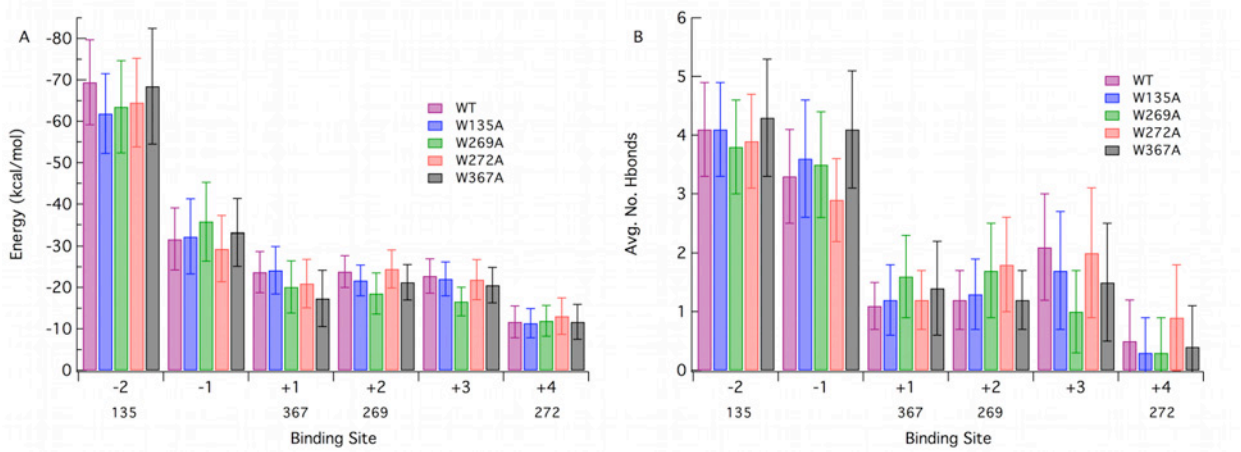

Figure S3. Results from the 250 ns MD simulations of the Cel6A wild type (WT) and the 4 mutants. Labels below the *x*-axis indicate position of the tryptophan to alanine mutations relative to the binding site designation. (A) Interaction energy of the protein with the cello-oligomer per glucose binding site. (B) Average number of hydrogen bonds per glucose binding site.

Molecular Dynamics – Site Interaction Energies.

In addition to calculating the overall interaction energy of the enzyme with the ligand as shown in Figure S3A, we have calculated the interaction energies of a given enzymatic residue interacting with the ligand. Figure S4 shows, for all four mutation locations, the total, electrostatic, and van der Waals interaction energies of the residue with the ligand at a given binding site. For the wild type, this residue is a tryptophan. For all four mutants, this residue is an alanine. Error bars represent 1 standard deviation. For each of the mutants, there is a marked decrease in the total interaction energy, which is expected given the reduced steric hindrance of the alanine residues in comparison to tryptophan. Interestingly, W269 and W367 seem to interact with glucose units in neighboring binding sites to a greater extent than either W135 or W272. Additionally, the W269A and W367A interaction energies with these same binding sites, while reduced in comparison to wild type, remain large in comparison to the reductions observed for W135A and W272A.

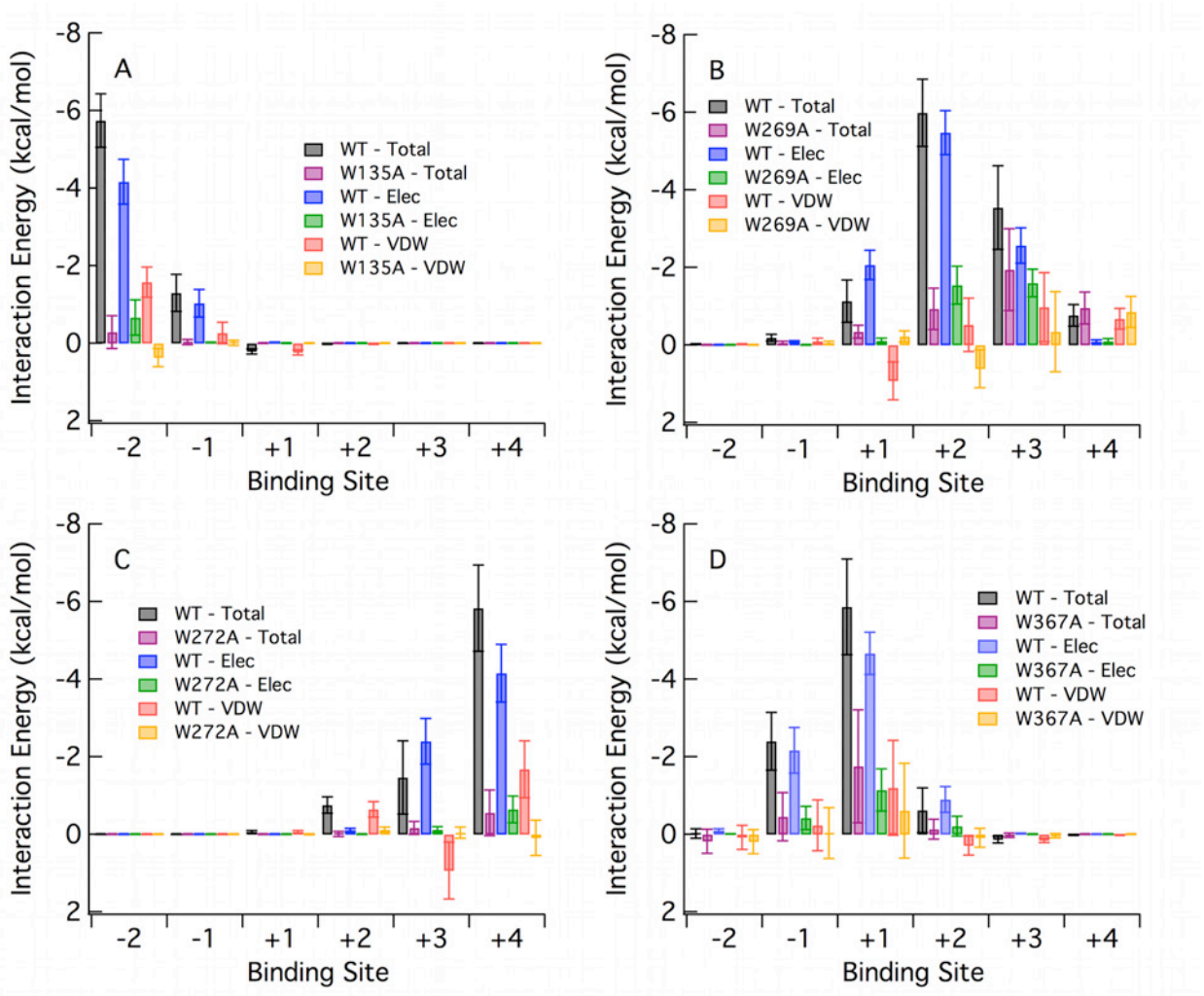

Figure S4. Interaction energies of tryptophan (wild type) and alanine (mutants) with the glucose unit in a given binding site for residues (A) W135/W135A, (B) W269/W269A, (C) W272/W272A, and (D) W367/W367A.

Molecular Dynamics – Root-mean-square deviation (RMSD).
Dynamical properties of the ligand were also evaluated from the five MD simulations including RMSD of the ligand, shown in Figure S5A, and the RMSD of the active site loop residues 172 through 182, Figure S5B. The RMSD values are calculated as compared to the starting crystal structure configuration. While RMSD values are shown starting at 0 ns, this time excludes the minimization steps as discussed above in the methods section, thus the initial deviation from a single starting point at 0 ns.

In the case of the ligand RMSD, Figure S5A, all four mutant RMSD trajectories do not appear to substantially deviate from wild type, most likely due to the inherent confinement within the active site tunnel. The active site loop 172-182 RMSD, Figure S5B, shows some conformational change in the W135A and W272A mutants.

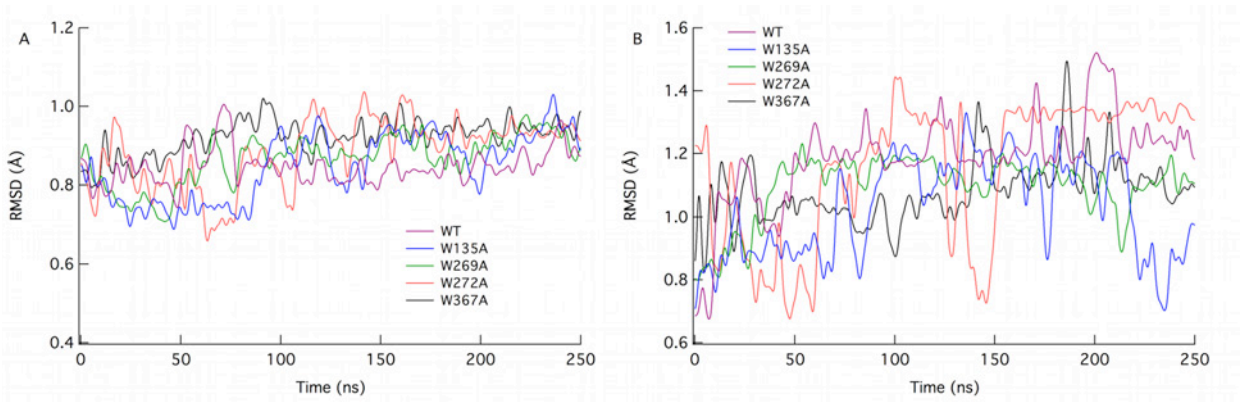

Figure S5. (A) RMSD values of the ligand from MD simulations of the wild type and mutants. (B) RMSD values of the active site loop residues 172 through 182 from MD simulations of the wild type and mutants.

The RMSD of all the residues lining the active site tunnel of Cel6A was also calculated from the wild type MD simulation, as shown below in Figure S6. The intent of this evaluation was to observe any residue conformational changes associated with the -1 glucose ring conformation changes; however, aside from the catalytic acid, D221, the active site residues appear to maintain constant positions or alternate between two positions apparently unrelated to the glucose ring conformation change.

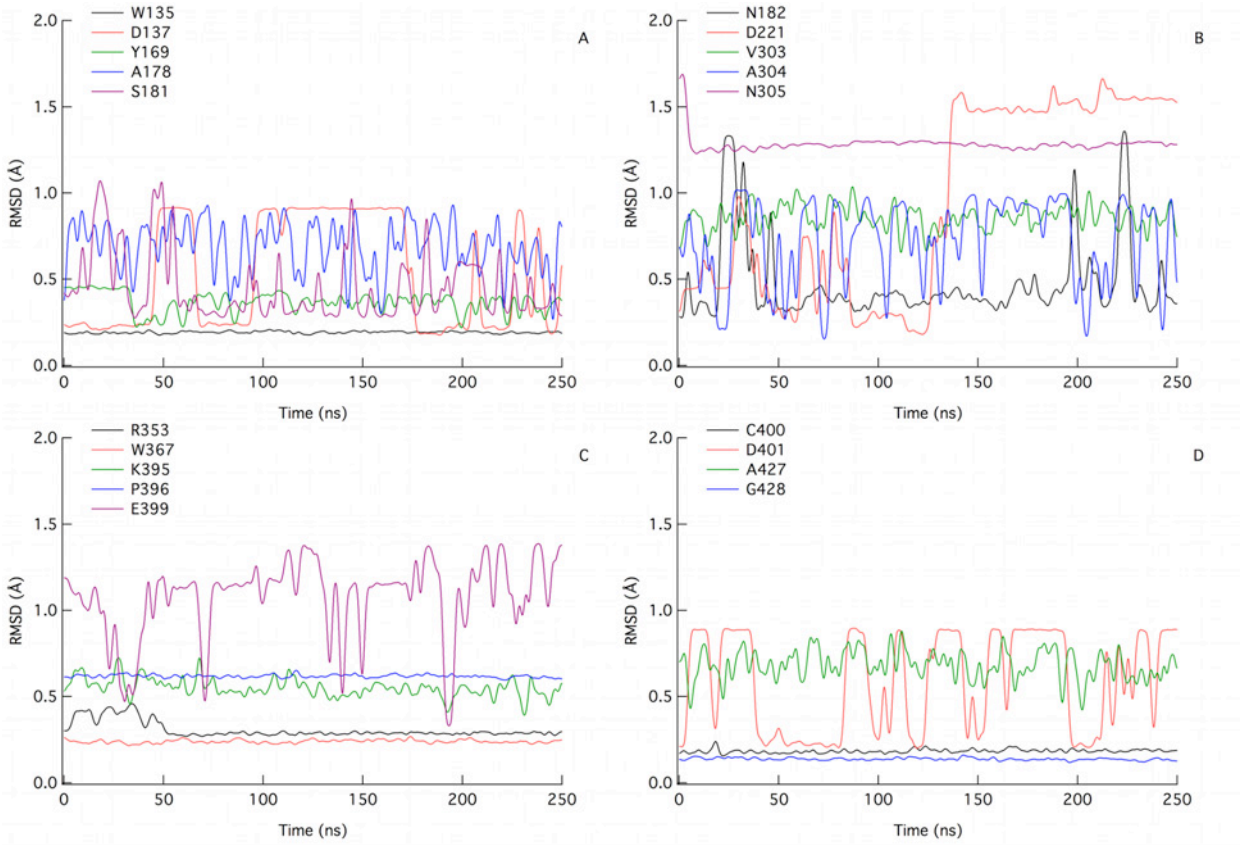

Figure S6. RMSD values of the active site residues from the wild type MD simulation. For clarity, the figure has been divided into four sections – (A), (B), (C), and (D).

Molecular Dynamics – Active Site Residue Dynamics. Additional analysis of the MD simulations included evaluation of key active site residue dynamics through hydrogen bonding and variation in distance from ligand features. Hydrogen bond cutoffs were defined as within 3.4 Å of the donor and acceptor and 60° from linear.

For all four mutants and the wild type enzyme, Figure S7 shows both the distance of the catalytic residue, D221, from the glycosidic oxygen, as measured from the OD2 atom of D221, and the distance of the center of mass of D221 from the center of mass of neighboring residue D175. Figure S7 also shows the hydrogen bonding of D221 to D175 as a function of time, which is primarily between the D221/D175 sidechain atoms. Interesting observations include the virtual lack of hydrogen bonding between D221 and D175 in mutant enzymes W135A, W269A, and W272A, which is related to the proximity of the D221 side chain to the glycosidic oxygen.

Figure S8 illustrates the distance of the center of mass of both Y169 and D221 from the primary alcohol group oxygen of the -1 glucose unit as well as the associated hydrogen-bonding pattern. Hydrogen bonding to the OH group has been shown for D221 as sticks, while Y169 is shown as a marker, so that overlap hydrogen bonding may be observed. Hydrogen bonding of D221 to the primary alcohol group is associated with a conformation change in the -1 glucose ring, from skew to chair as shown in Figure 3 of the main text. Hydrogen bonding of Y169 to the primary alcohol group appears to be independent of the glucose ring conformation.

Hydrogen bonding of residues Y169 and R174 for all four mutants and the wild type is shown in Figure S9. The wild type hydrogen bonding shows a stable pattern of approximately 1 hydrogen bond over the course of the entire simulation. Each of the tryptophan to alanine mutations interrupts this interaction to varying extents. The breaking of this interaction has been said to correspond to a conformational change in active site loop 172-182 residue region (32). While comparison of the hydrogen bonding patterns in Figure S9 to the RMSD of this active site loop in Figure S5B is not conclusive, it does appear that some of the more dramatic fluctuations in this loop result in a breaking of the Y169/R174 interaction.

Figure S10 shows the hydrogen bonding of residues D175 and R174 for the wild type and mutant enzymes as a function of time. Hydrogen bonding of these two residues in the W272A mutant appears to be related to the conformational change in the active site loop 172-182 residue region. However, similar correlation for the W135A mutant is not observed. Rather, the W135A mutant exhibits increased hydrogen bonding of D175 and R174 over wild type.

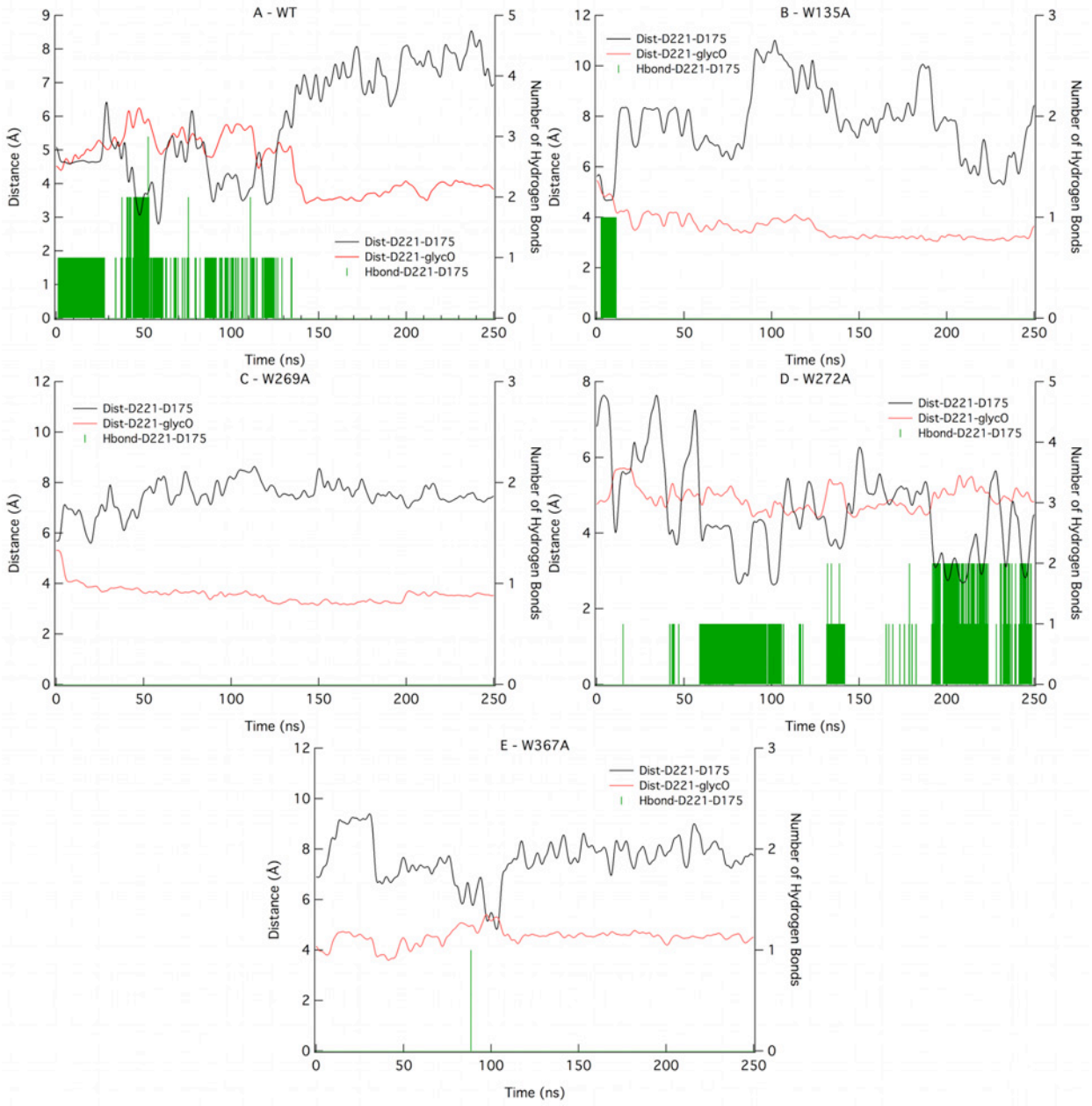

Figure S7. Distance of the catalytic residue, D221, from neighboring D175 and the glycosidic oxygen and the hydrogen bonding pattern of D221 and D175 for (A) WT, (B) W135A, (C) W269A, (D) W272A, and (E) W367A.

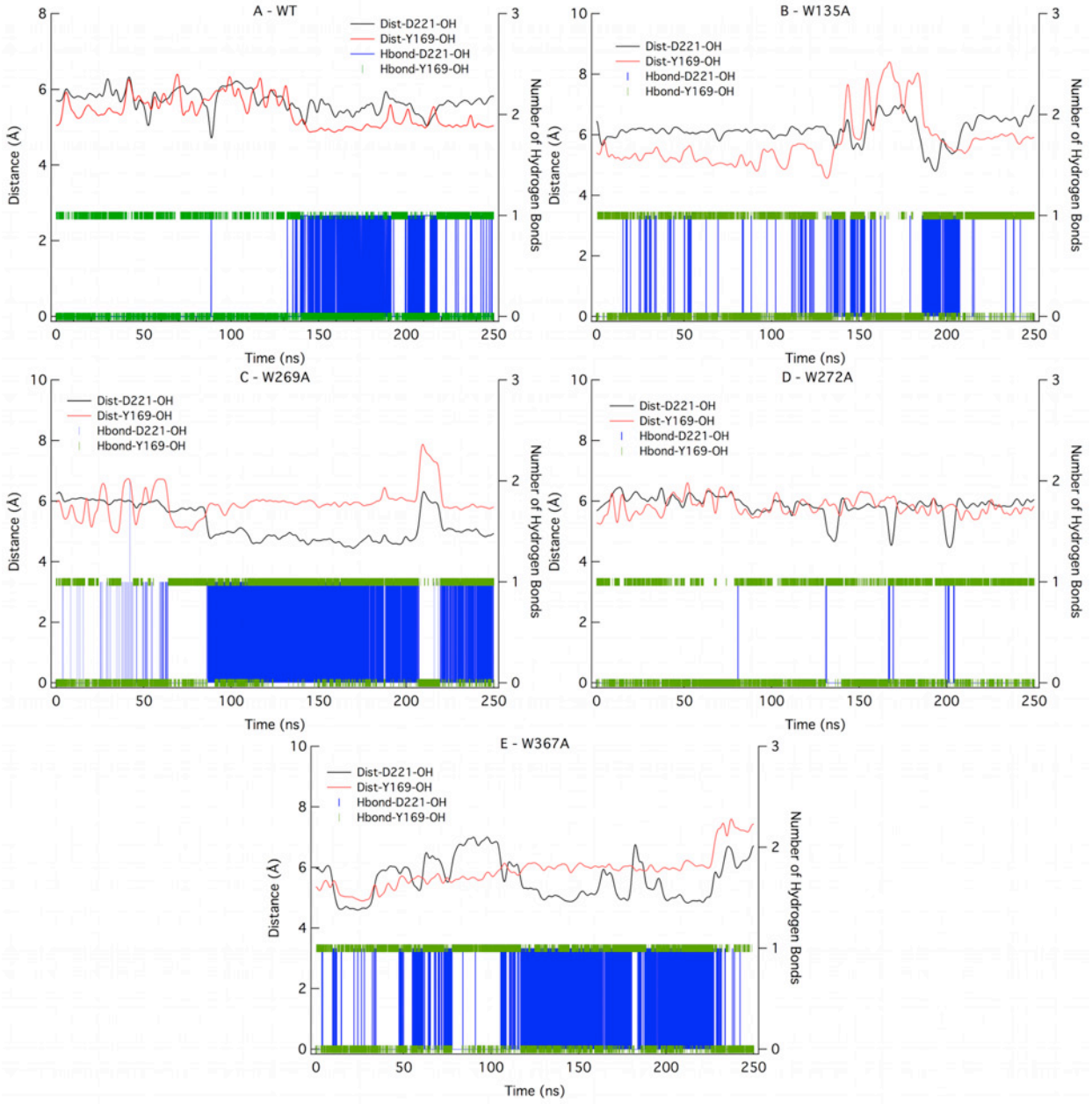

Figure S8. Distance of the catalytic residue, D221, and Y169 primary alcohol oxygen of the -1 glucose unit and the associated hydrogen bonding pattern for (A) WT, (B) W135A, (C) W269A, (D) W272A, and (E) W367A.

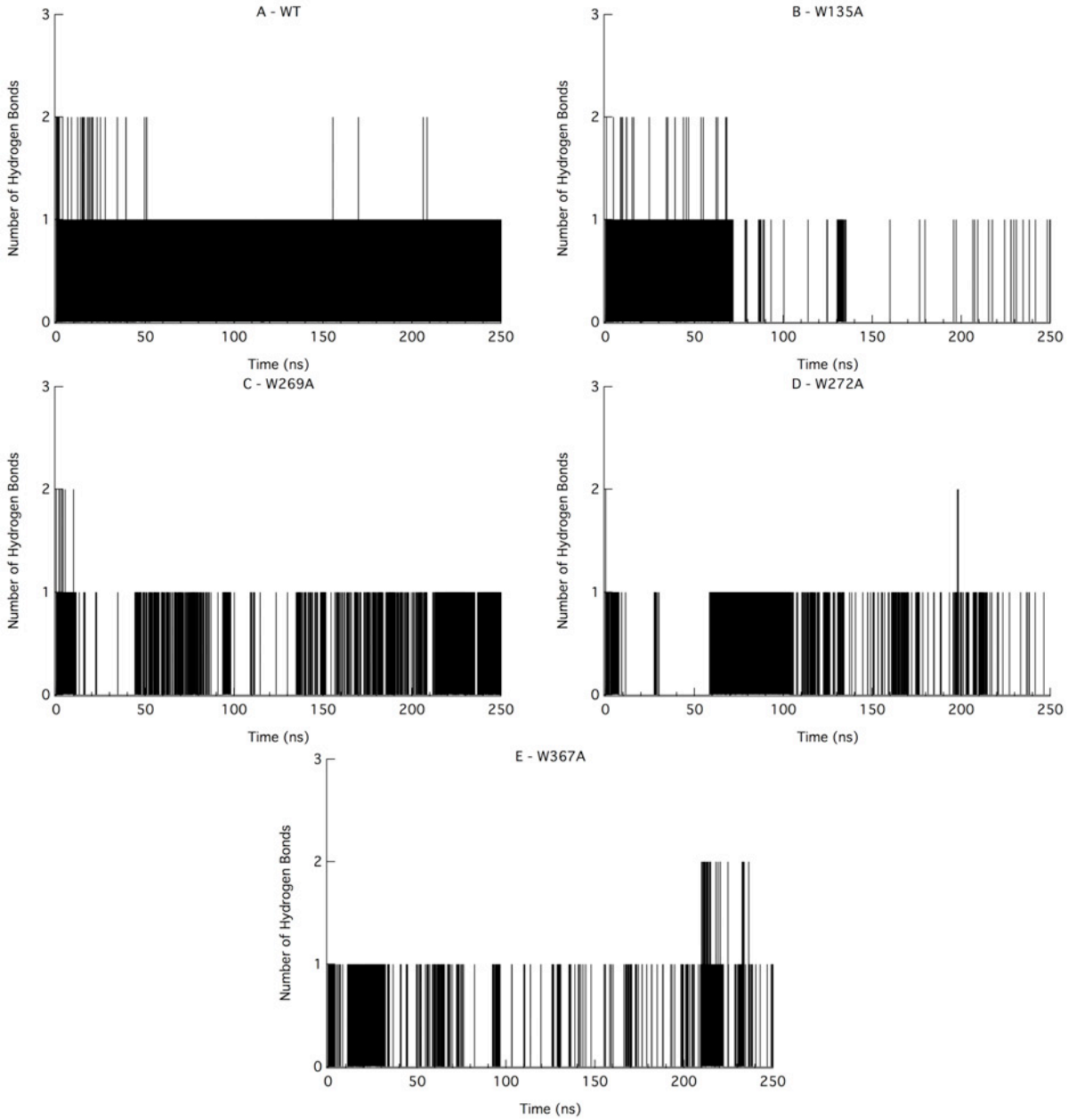

Figure S9. Hydrogen bonding pattern of Y169 and R174 for (A) WT, (B) W135A, (C) W269A, (D) W272A, and (E) W367A.

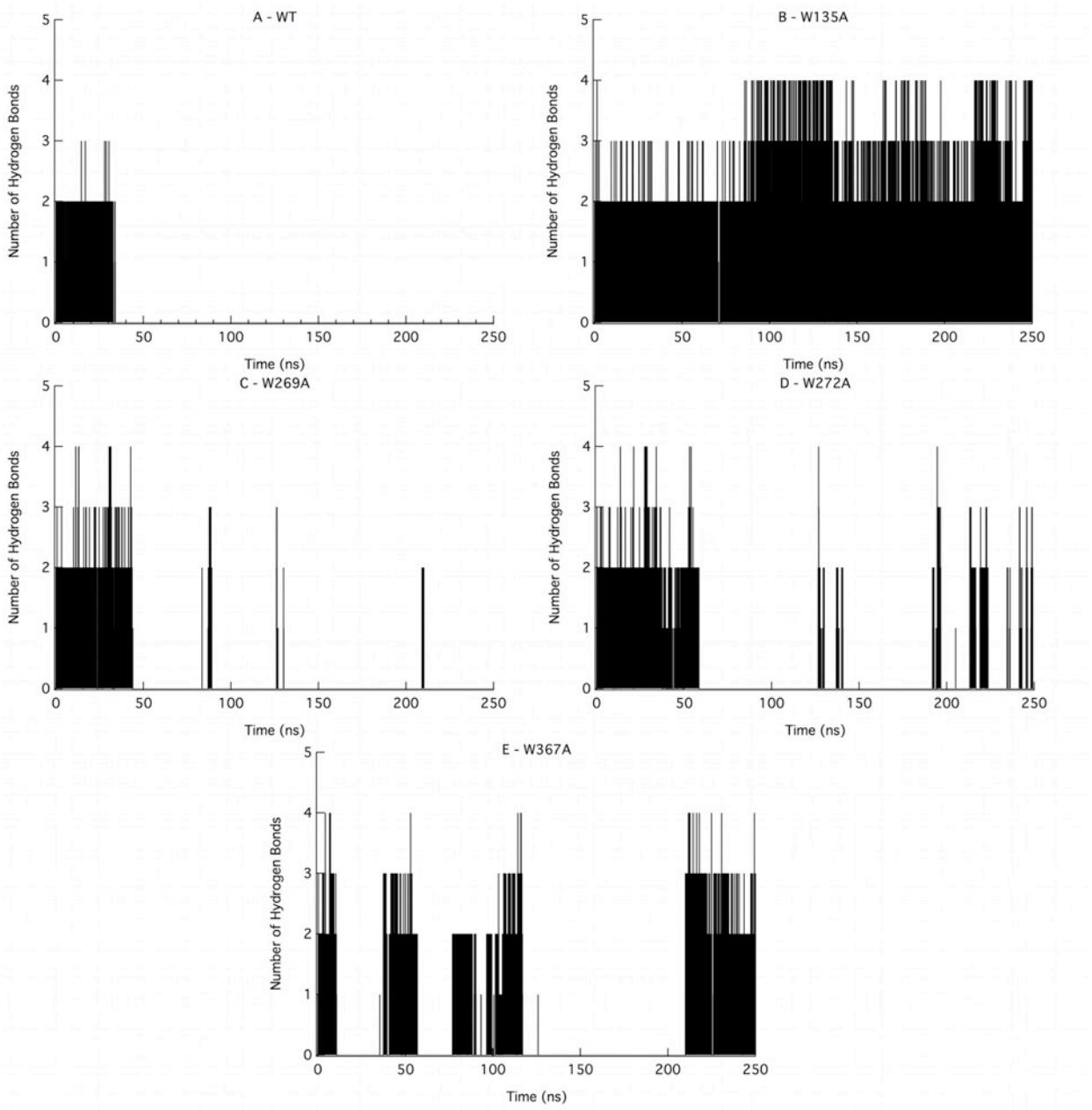

Figure S10. Hydrogen bonding pattern of R174 and D175 for (A) WT, (B) W135A, (C) W269A, (D) W272A, and (E) W367A.

THE EFFECT OF GLYCOSYLATION ON ENZYME DYNAMICS FROM SIMULATION

Prior to building the MD simulations discussed in this work, we set out to determine if native glycosylation patterns, as described by Hui *et al*.,(16) impacted the dynamical and energetic properties typically observed from MD simulations. Simulations set up exactly as described in the methods section were performed for both glycosylated and non-glycosylated Cel6A. The only difference between the two systems was the presence or absence of *N* and *O*-glycans at the surface of the enzyme. Additionally, these simulations were performed in the absence of a ligand, referred to in the Figures as "free", and with a ligand, referred to as "bound", totaling four simulations of 100 ns each.

The RMSF values as a function of residue number for each of these four simulations are shown in Figure S11. Comparing only the glycosylated and non-glycosylated values to each other in both the free and bound states, indicates there is very little difference in fluctuation of individual residues as a result of glycosylation. Similarly, the RMSD values shown in Figure S12 for glycosylated and non-glycosylated Cel6A do not differ enough to conclude any dynamical effects.

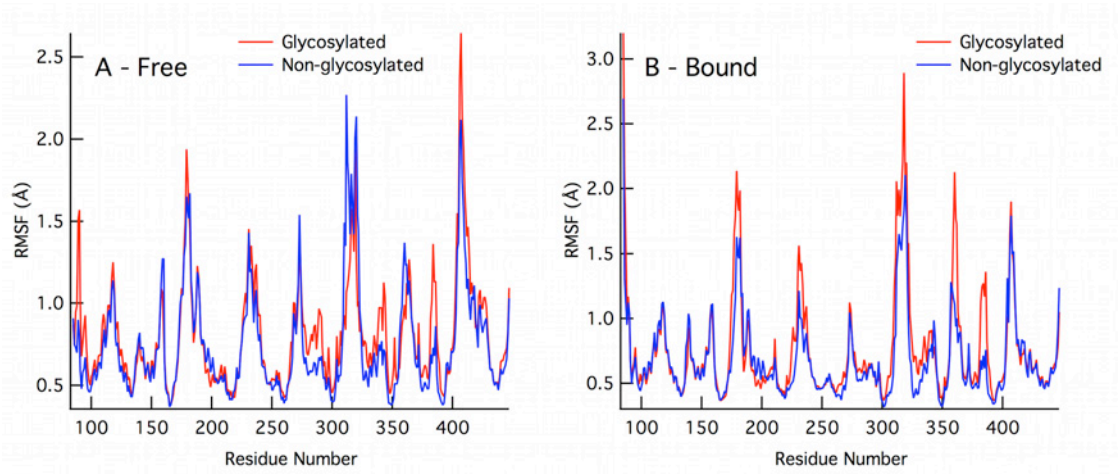

Figure S11. RMSF versus residue number for glycosylated and non-glycosylated Cel6A for both the (A) free enzyme and the (B) bound enzyme.

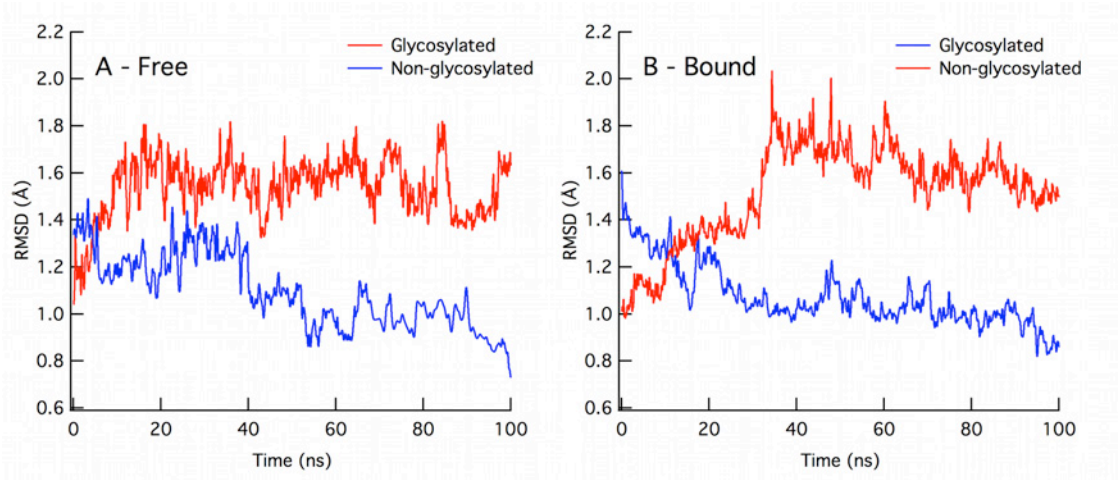

Figure S12. RMSD versus time for glycosylated and non-glycosylated Cel6A for both the (A) free enzyme and the (B) bound enzyme.

Two final comparisons of glycosylated and non-glycosylated properties were calculated including the hydrogen bonding pattern of the protein to the ligand, Figure S13A, and the interaction energy of the

protein with the ligand, Figure S13B. As with the hydrogen bonding evaluation discussed in the main text, hydrogen bond cutoffs were defined as within 3.4 Å of the donor and acceptor and 60° from linear. Error bars shown in Figure S13A and S13B represent 1 standard deviation. Each of these figures show little difference in the protein interaction with the ligand as a result of glycosylation or lack thereof.

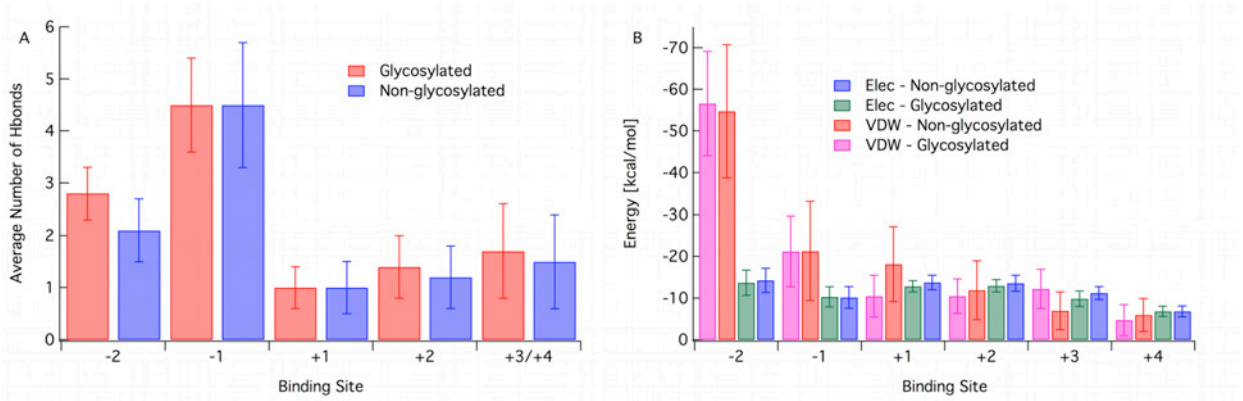

Figure S13. (A) Average number of hydrogen bonds on a per binding site basis for the glycosylated and non-glycosylated bound enzyme. (B) The interaction energy of the protein with the ligand on a binding site basis for the glycosylated and non-glycosylated enzyme.

This evaluation of properties as a function of glycosylation is not intended to be an overarching statement on the effects of native glycosylation on all enzymes, but rather to determine the effects on simulation observables for a very specific case. Despite the apparent conclusion that glycosylation appears to have little effect on the interaction of the enzyme with its ligand, all simulations from which we draw conclusions regarding binding affinity or ligand binding free energy were performed with glycosylation patterns as the enzyme might be found in nature as expressed by *T. reesei*.

REFERENCES

1. Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) *J. Comput. Chem.* **30**, 1545-1614
2. Zou, J.-y., Kleywegt, G. J., Stahlberg, J., Driguez, H., Nerinckx, W., Claeyssens, M., Koivula, A., Teeri, T. T., and Jones, T. A. (1999) *Structure* **7**, 1035-1045
3. Nosé, S., and Klein, M. L. (1983) *Mol. Phys.* **50**, 1055-1076
4. Hoover, W. G. (1985) *Phys. Rev. A* **31**, 1695-1697
5. Ryckaert, J., Ciccotti, G., and Berendsen, H. (1977) *J. Comput. Phys.* **23**
6. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) *J. Chem. Phys.* **103**, 8857
7. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *J. Phys. Chem. B* **102**, 3586-3616
8. Mackerell, A. D., Feig, M., and Brooks, C. L. (2004) *J. Comput. Chem.* **25**, 1400-1415
9. Guvench, O., Hatcher, E., Venable, R. M., Pastor, R. W., and MacKerell, A. D. (2009) *J. Chem. Theor. Comp.* **5**, 2353-2370

10. Guvench, O., Greene, S. N., Kamath, G., Brady, J. W., Venable, R. M., Pastor, R. W., and Mackerell, A. D. (2008) *J. Comput. Chem.* **29**, 2543-2564
11. Guvench, O., Mallajosyula, S. S., Raman, E. P., Hatcher, E. R., Vanommeslaeghe, K., Foster, T. J., Jamison, F. W., and MacKerell, A. D. (2011) *J. Chem. Theory Comput.*, ASAP
12. Jorgensen, W. L., Chandrasekhar, J., and Madura, J. D. (1983) *J. Chem. Phys.* **79**, 926-935
13. Durell, S. R., Brooks, B. R., and Ben-Naim, A. (1994) *J. Phys. Chem.* **98**, 2198-2202
14. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) *J. Comp. Chem.* **26**, 1781-1802
15. Schneider, T., and Stoll, E. (1978) *Phys. Rev. B: Condens. Matter Mater. Phys.* **17**, 1302
16. Hui, J. P. M., White, T. C., and Thibault, P. (2002) *Glycobiology* **12**, 837-849
17. Kirkwood, J. G. (1935) *J. Chem. Phys.* **3**, 300-313
18. Straatsma, T. P., and McCammon, J. A. (1991) *J. Chem. Phys.* **95**, 1175-1188
19. Frenkel, D., and Smit, B. (2002) *Understanding Molecular Simulations: From Algorithms to Applications*, 2nd ed.,
20. Pohorille, A., Jarzynski, C., and Chipot, C. (2010) *J. Phys. Chem. B* **114**, 10235-10253
21. Bruckner, S., and Boresch, S. (2011) *J. Comput. Chem.* **32**, 1320-1333
22. Steinbrecher, T., Mobley, D. L., and Case, D. A. (2007) *J. Chem. Phys.* **127**
23. Macke, T., Svrcek-Seiler, W. A., Brown, R. A., Kolossvary, I., Y.J., B., and Case, D. A. (2011) *NAB* Ver. 6
24. Macke, T. J., and Case, D. A. (1998) Modeling unusual nucleic acid structures. in *Molecular Modeling of Nucleic Acids* (Leontis, N. B., and SantaLucia, J. eds.), Amer Chemical Soc, Washington. pp 379-393
25. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) *J. Comp. Chem.* **26**, 1668-1688
26. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. A., Wang, J., Duke, R. E., *et al.* (2008) *Amber* Ver. 10
27. Junmei, W., Cieplak, P., and Kollman, P. A. (2000) *J. Comput. Chem.* **21**, 1049-10741074
28. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) *Proteins: Struct., Funct., Bioinf.* **65**, 712-725
29. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995) *Chem. Phys. Lett.* **246**, 122-129
30. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996) *J. Phys. Chem.* **100**, 19824-19839
31. Lehoucq, R., Sorensen, D. C., and Yang, C. (1998) Arpack User's Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restored Aroldi Methods. in *SIAM*, Philadelphia, PA
32. Koivula, A., Ruohonen, L., Wohlfahrt, G., Reinikainen, T., Teeri, T. T., Piens, K., Claeyssens, M., Weber, M., Vasella, A., Becker, D., Sinnott, M. L., Zou, J. Y., Kleywegt, G. J., Szardenings, M., Stahlberg, J., and Jones, T. A. (2002) *J. Amer. Chem. Soc.* **124**, 10015-10024