

Supplementary Information

Concerted motions and large scale structural fluctuations of *Trichoderma reesei* Cel7A cellobiohydrolase

Rodrigo L. Silveira and Munir S. Skaf*

Institute of Chemistry, University of Campinas, Campinas, Cx. P. 6154, 13084-862, Sao Paulo, Brazil

*Corresponding author. E-mail: skaf@iqm.unicamp.br

Table S1. Summary of the systems simulated in this work.

System	Number of simulations	Time of each simulation
apo- <i>TrCel7A</i>	3	150 ns
C9- <i>TrCel7A</i>	3	150 ns
Glycosilated apo- <i>TrCel7A</i>	3	150 ns
Glycosilatd C9- <i>TrCel7A</i>	3	150 ns
apo- <i>TrCel7A</i> (aMD)	1	1000 ns
C9- <i>TrCel7A</i> (aMD)	1	1000 ns

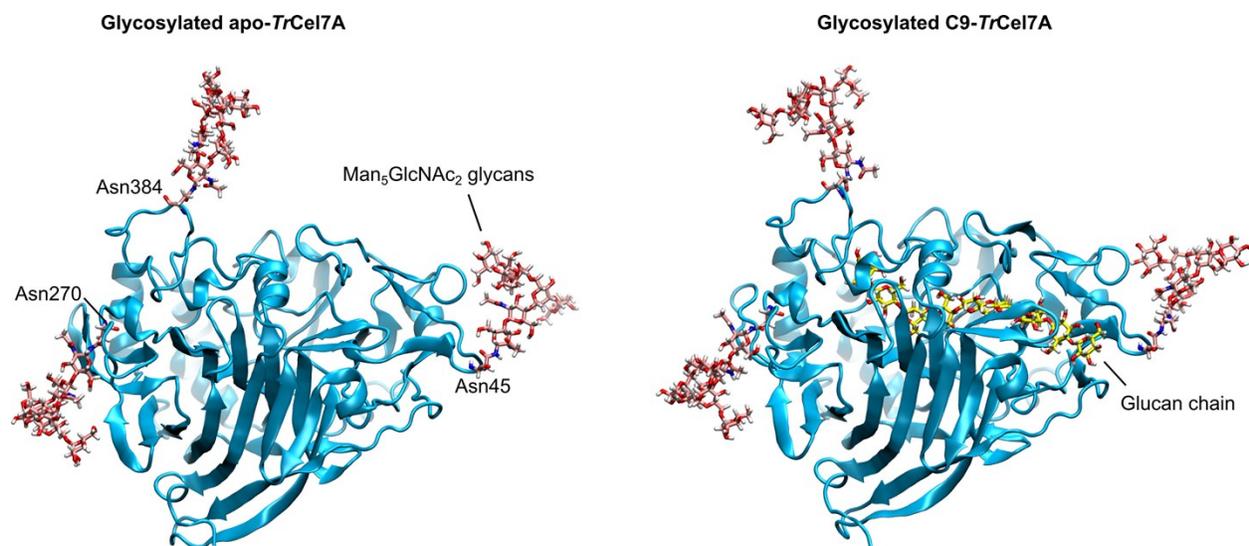
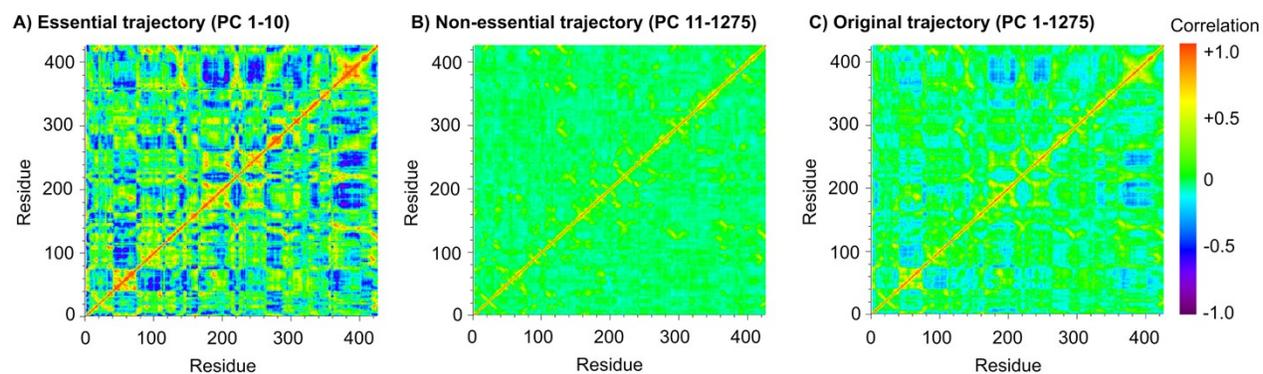


Fig. S1. Equilibrated structure of the glycosylated apo-*Tr*Cel7A and C9-*Tr*Cel7A systems.

apo-TrCel7A



C9-TrCel7A

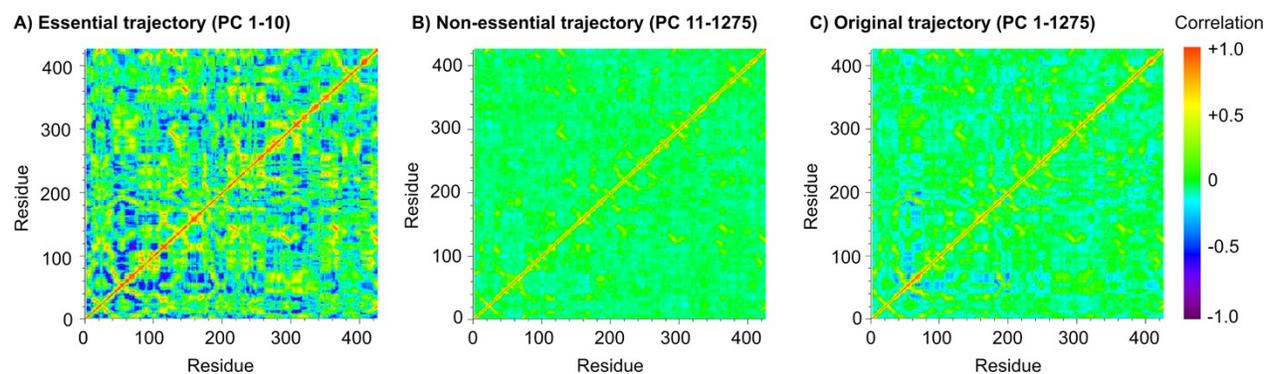


Fig. S3. Comparison of correlation maps of essential, non-essential and original trajectory of apo-TrCel7A (upper panels) and C9-TrCel7A (lower panels).

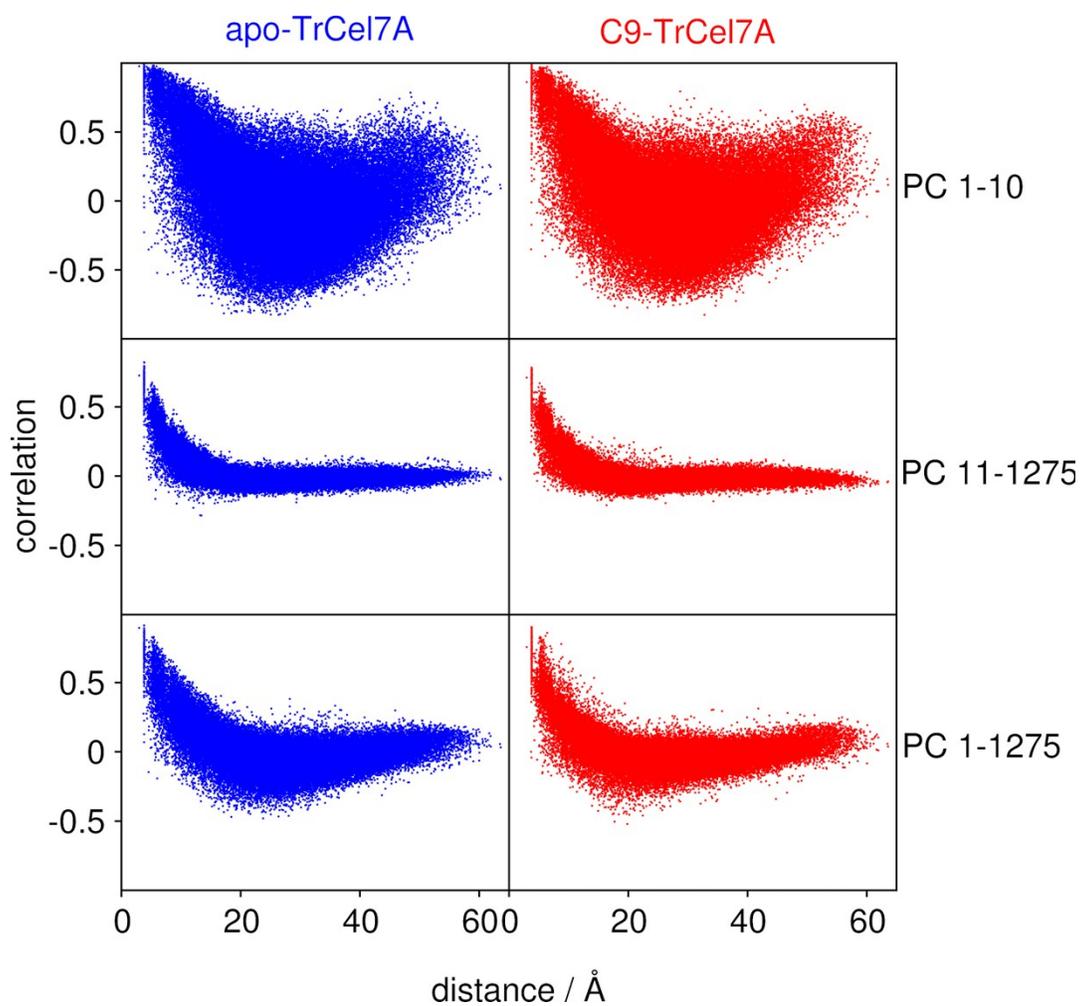
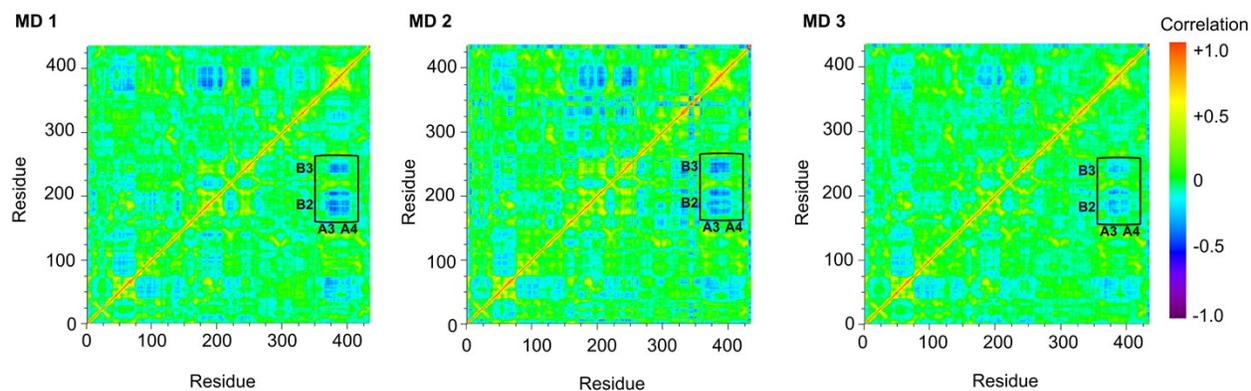


Fig. S4. Correlation vs. distance scatter plots for apo-*TrCel7A* and C9-*TrCel7A*, considering the essential (PC 1-10), non-essential (PC 11-1275) and original (PC 1-1275) trajectories.

apo-TrCel7A (glycosylated)



C9-TrCel7A (glycosylated)

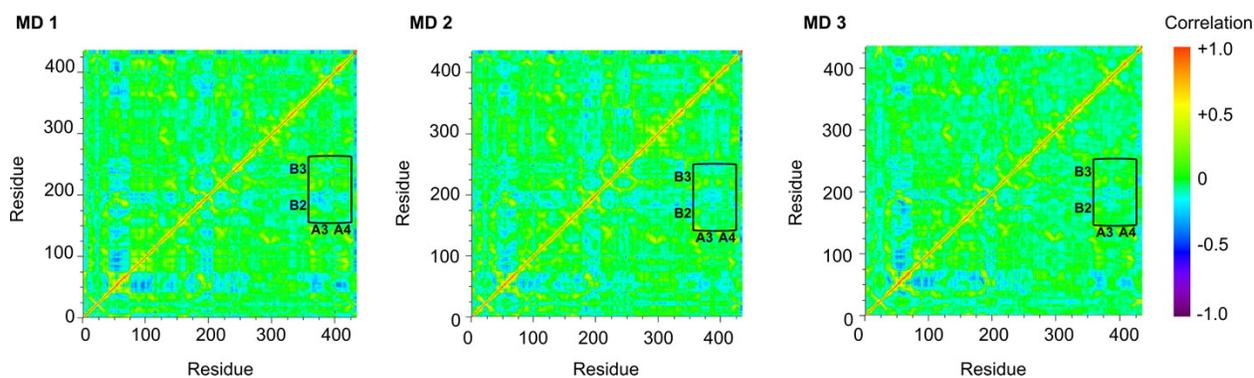


Fig. S5. Correlation maps of glycosylated apo-TrCel7A (upper panels) and C9-TrCel7A (lower panels) in three independent simulations lasting 150 ns. It can be seen that the correlation maps are quite similar for the different simulations of each system. Thus, we consider that the fluctuations have converged in such simulations.

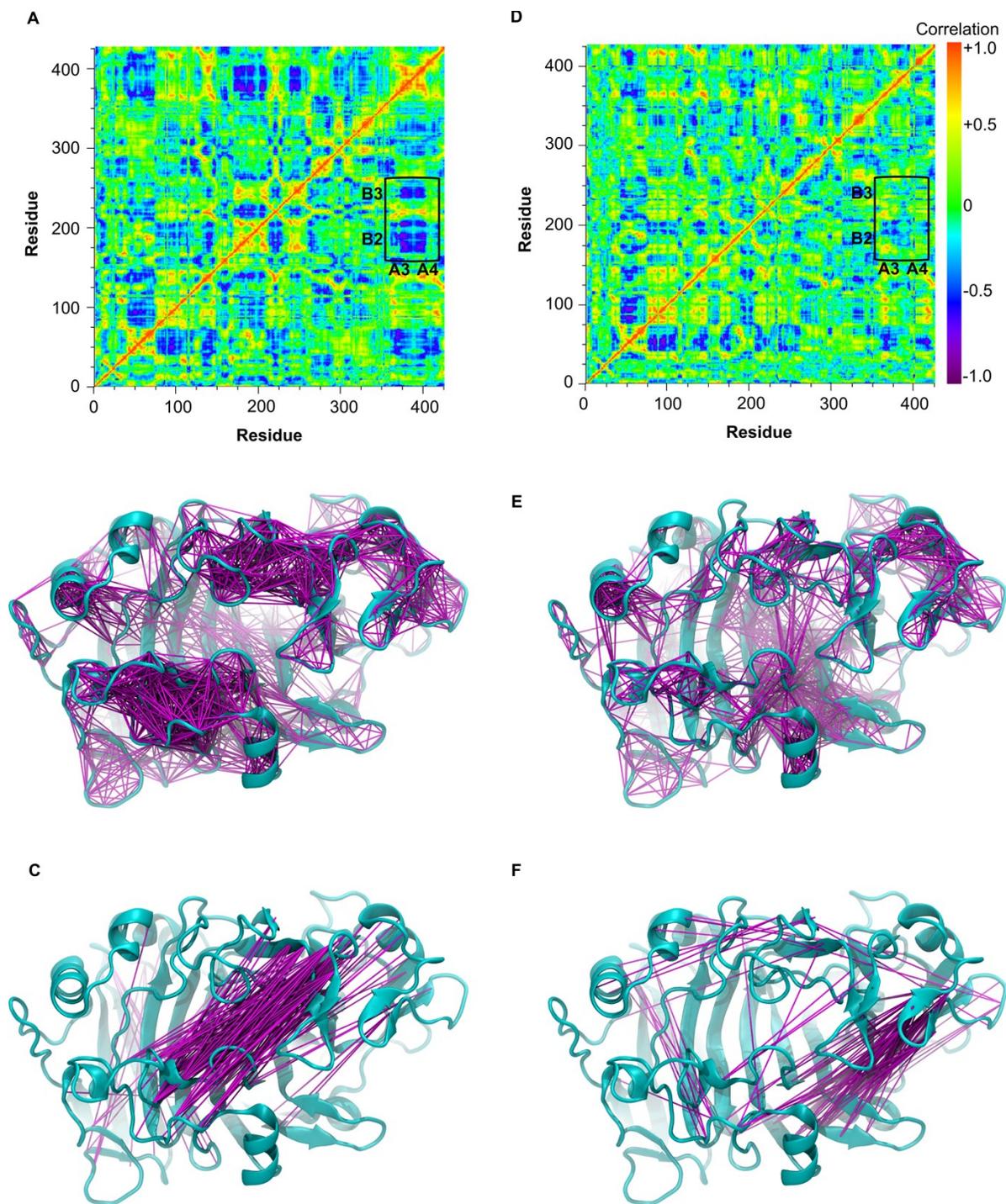


Fig. S6. (a) Essential cross-correlation map for glycosylated apo-*TrCel7A*, with loop-loop negative correlations highlighted. *TrCel7A* structure with pairs of residues that exhibit (b) positive and (c) negative correlations with modulus greater than 0.7 connected by purple sticks. (d), (e) and (f): same as (a), (b) and (c), but for the glycosylated C9-*TrCel7A*.

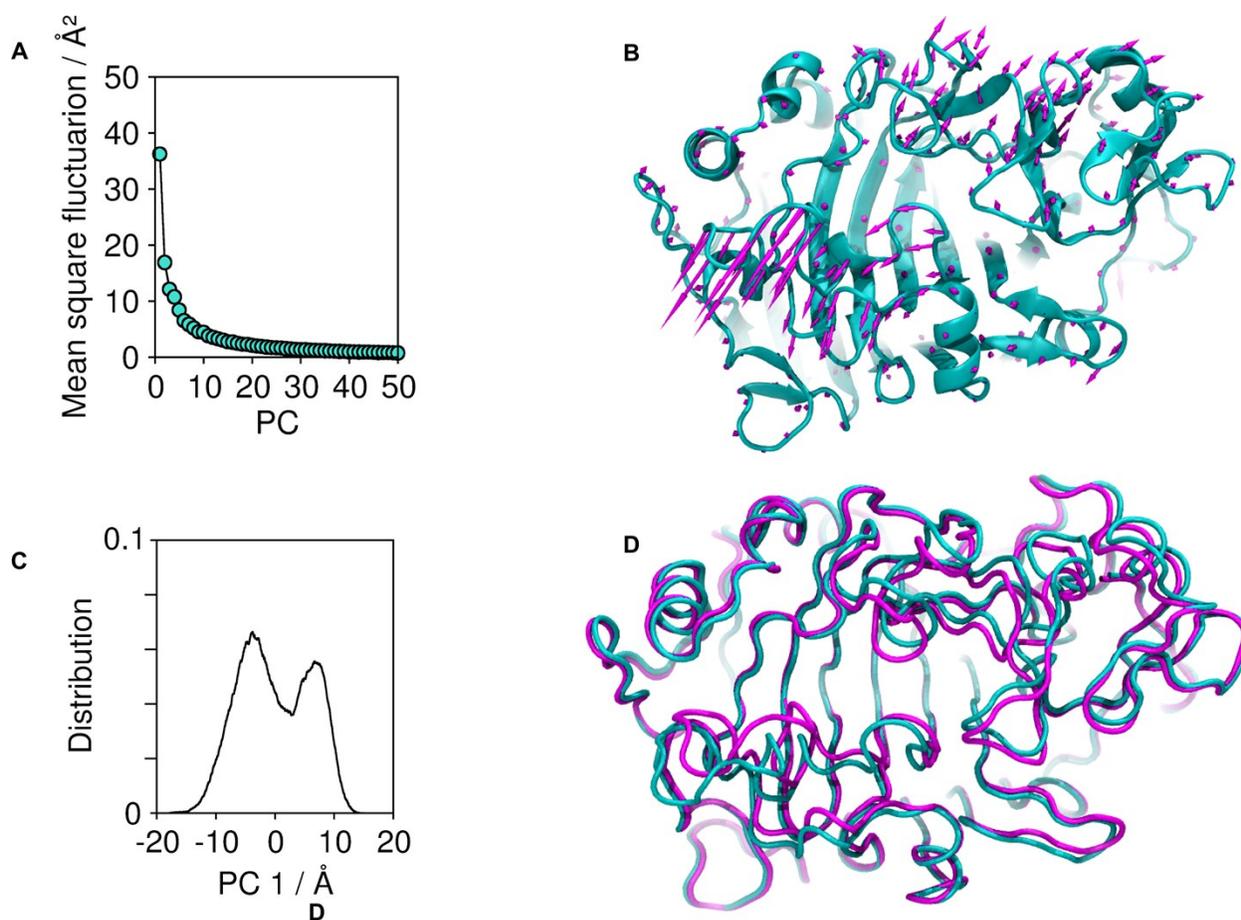


Fig. S7. Essential dynamics of glycosylated apo-TrCel7A. (a) Mean square fluctuations for the first 20 collective modes. (b) Direction of the dominant collective mode of the enzyme. (c) Distribution of the amplitudes of the lowest eigenvector (PC1), showing the existence of interactions that stabilize the open and close states. (d) Superposed structures corresponding to the minimum (cyan) and maximum (purple) amplitudes of the open-close conformational coordinate.

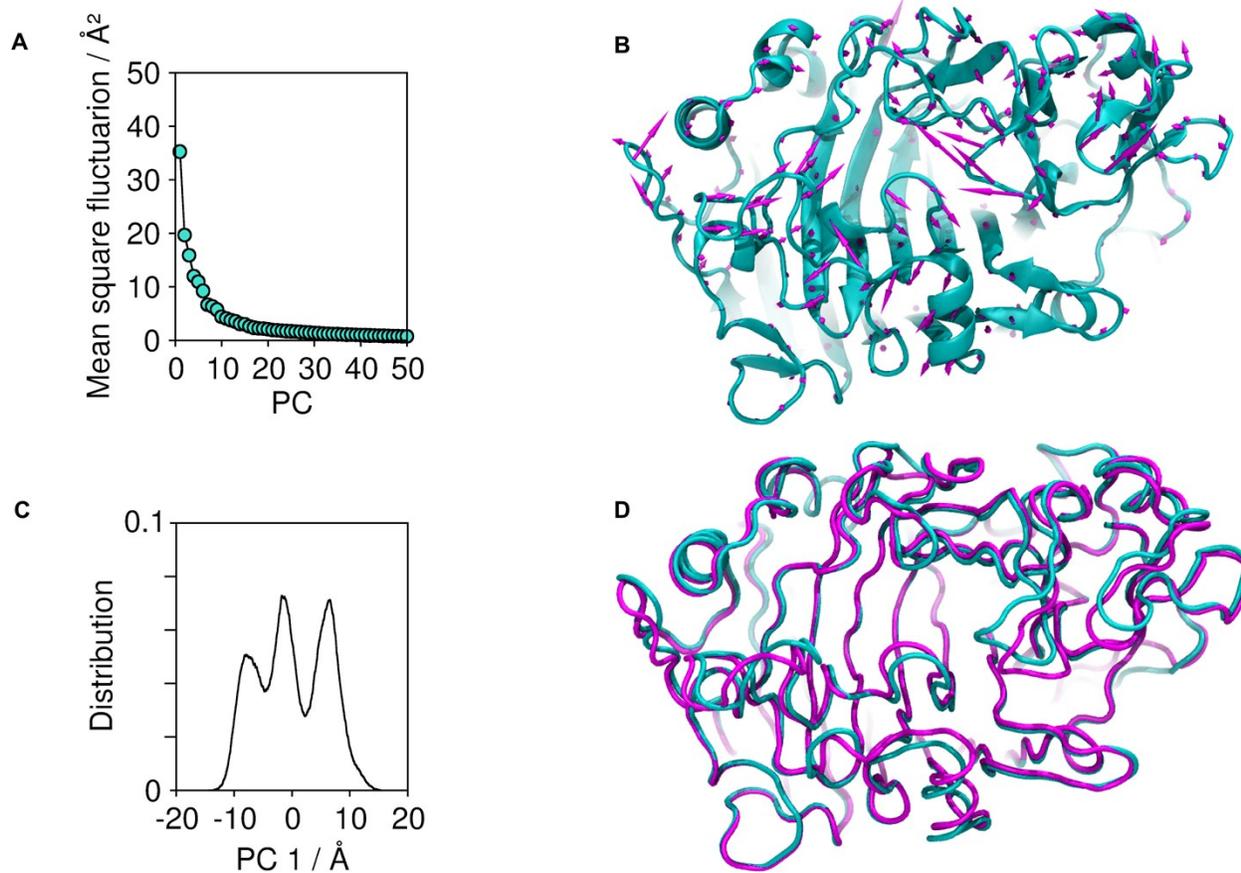


Fig. S8. Essential dynamics of the glycosylated C9-TrCel7A. Panels (a)-(d) are the same as in Fig. S6, but for the liganded C9-TrCel7A structure. The substrate is omitted for clarity.

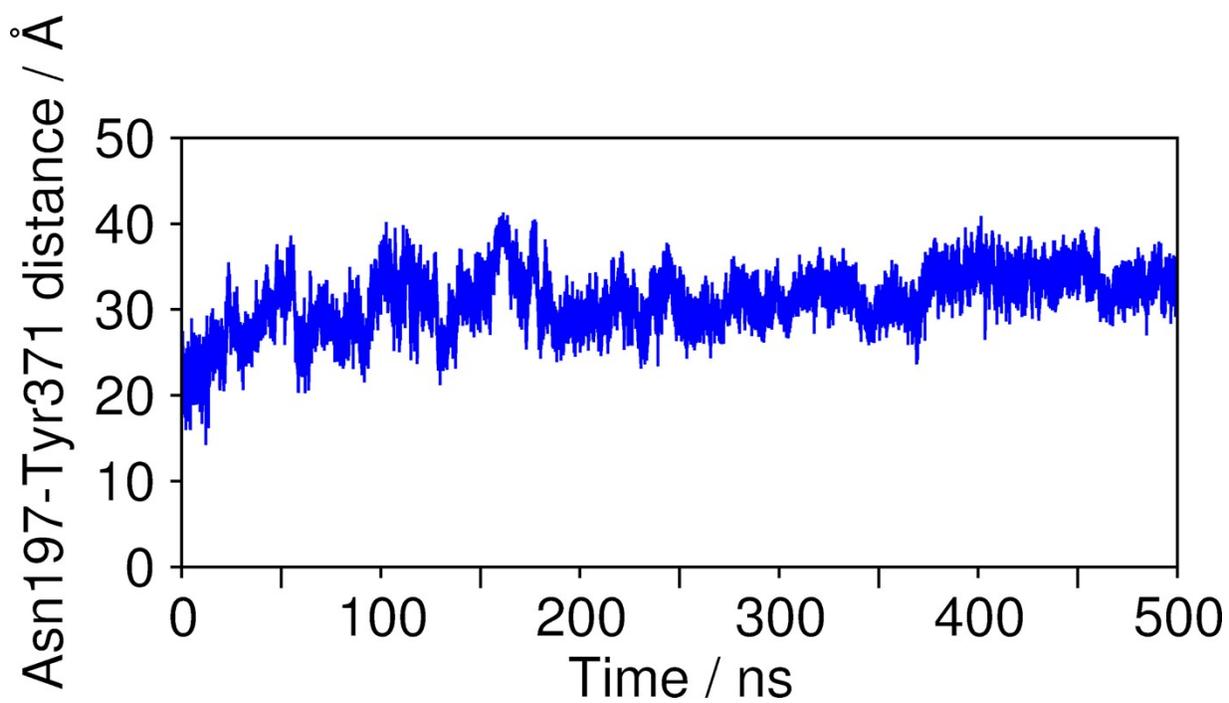


Fig. S9. Time history of the Asn197-Tyr371 distance during a conventional MD simulation after the aMD run. The distance is always greater than 15 Å, indicating that the loops remain open during 500 ns and suggesting that the open state is a stable configuration on such time-scale.

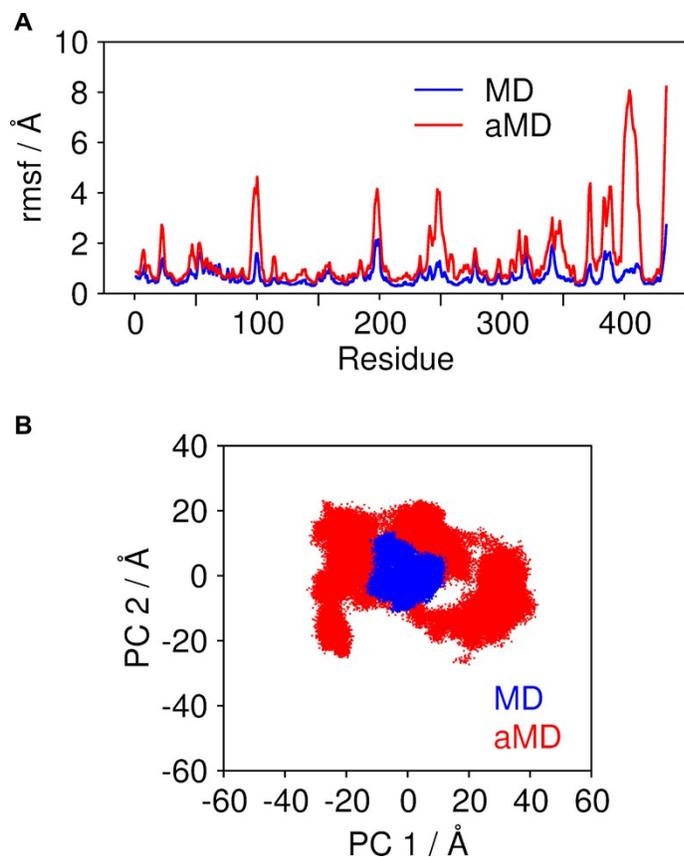


Fig. S10. (a) Root mean square fluctuation of C9-*TrCel7A* during conventional and accelerated MD simulations. All the peaks in the aMD rmsf correspond to loop regions, showing that the boost potential only enhances the conformational sampling of the loops while keeping the protein fold intact. (b) Distribution of the coordinates along the first 2 principal components of PCA obtained from conventional MD and aMD, showing the extensive sampling achieved with aMD. Comparison with Fig. 6 of the main text reveals that the C9 substrate is able to stabilize the *TrCel7A* structure even under the aMD boost potential.