Structural, biochemical, and computational characterization of the glycoside hydrolase family 7 cellobiohydrolase of the tree-killing fungus *Heterobasidion irregulare*

Majid Haddad Momeni[1,‡], Christina M. Payne[2,3,‡], Henrik Hansson[1], Nils Egil Mikkelsen[1], Jesper Svedberg[1], Åke Engström[4], Mats Sandgren[1], Gregg T. Beckham[5,6,*], and Jerry Ståhlberg[1,*]

1. Department of Molecular Biology, Swedish University of Agricultural Sciences, POB 590, SE-751 24 Uppsala, Sweden
2. Biosciences Center, National Renewable Energy Laboratory, Golden CO 80401, United States
3. Department of Chemical and Materials Engineering, University of Kentucky, Lexington KY 40506, United States
4. Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden.
5. National Bioenergy Center, National Renewable Energy Laboratory, Golden CO 80401, United States
6. Department of Chemical Engineering, Colorado School of Mines, Golden CO 80401, United States

‡ These authors contributed equally to this work.
* Corresponding Author E-mail: jerry.stahlberg@molbio.slu.se; gregg.beckham@nrel.gov

The Supplemental Data contains information regarding detailed experimental methods for protein identification, purification, and crystallization; information regarding the construction and execution of the molecular dynamics simulations; and additional data related to the observations from the molecular dynamics simulations.

Cultivation, Protein Identification, and Protein Purification Methods

The *H. irregulare* strain TC-32-1 that was recently sequenced (1) was maintained on Hagem agar (2) at room temperature. The fungus was cultivated at ambient temperature (20-25°C) using Kremer and Woods (3) minimal medium [$(NH_4)_2HPO_4$ 2.6 g, 2,2-dimethylsuccinic acid 2.2 g, KH_2PO_4 1.1 g, urea 0.6 g, $MgSO_4{\cdot}7H_2O$ 0.5 g, $FeSO_4{\cdot}H_2O$ 10 mg, CaCl2·$2H_2O$ 74 mg, $ZnSO_4{\cdot}7H_2O$ 6 mg, $MnSO_4{\cdot}4H_2O$ 5 mg, $CoCl_2{\cdot}6H_2O$ 1 mg, thiamin 0.1 mg, in 1 L of distilled water, adjusted to pH 5.0], with microcrystalline cellulose, milled spruce heart wood, or Aspen sawdust as carbon sources. Protein separation was performed on the Äkta Explorer chromatography system, using columns and chromatography media from GE Healthcare (Uppsala, Sweden).

Protein identification was performed with broth from a static culture on spruce wood powder. Two small squares (0.5 cm X 0.5 cm) cut from an agar plate overgrown with *H. irregulare* TC-32-1 were used to inoculate 200 mL medium with 1% glucose as carbon source. After one week incubation, the preculture was mixed and transferred equally to six 2.8 L Fernbach flasks containing 150 mL medium and 15 g spruce powder, 10 g aspen sawdust or 10 g Avicel cellulose (Whatman CF-11) (two of each). After three weeks, 400 mL of medium was added to each flask, each of which was then subjected to shaking at 80 rpm for an additional 2 days before harvest. The cultures were filtered through ~1 µm Whatman GF-B glass fiber filters and centrifuged for 30 min at 16,000 x g (4°C). The supernatants were sterile filtered through 500 mL bottle-top vacuum filtration units (0.2 um PES; FiltroPur BT-50; Sarstedt, USA). The spruce culture filtrate was concentrated by tangential-flow filtration (Vivaflow 50; Sartorius, France) and the buffer was exchanged to 10 mM sodium acetate, pH 5.0, by repeated concentration and dilution in a Vivaspin 20 centrifugation unit (5,000 MWCO; Sartorius, France). The concentrated protein solution (25 mL) was passed through an 8 mL Source 30S cation-exchange column equilibrated with 10 mM sodium acetate, pH 5.0. The flow-through fraction was then diluted three times with 25 mL BisTris-HCl, pH 6.5, applied to an 8 mL Source 30Q anion-exchange column and fractionated with a salt gradient up to 1 M NaCl in the same buffer. Samples from selected fractions were separated by SDS-

PAGE in a homogeneous gel (15%) using a Mini-PROTEAN Tetra Cell (BIO-RAD) and stained with Colloidal Blue Staining Kit (Invitrogen) according to the manufacturer's instructions.

The proteins in cutout gel bands were alkylated by iodoacetamide, digested with trypsin (Promega modified trypsin) and the tryptic peptides were analyzed by MALDI-TOF using a Bruker Ultraflex TOF/TOF instrument (4). A peptide mass fingerprint search with Mascot (Matrix Science; http://matrixscience.com) against predicted proteins in the *H. irregulare* genome database yielded the best hit (score 96; six peptides ranging from 1,075 to 2,864 Da) with protein ID 38802, annotated as "candidate [reducing end-acting] cellobiohydrolase" of glycoside hydrolase family GH7 (http://genome.jgi-psf.org/Hetan2/Hetan2.home.html) (5). A signal peptide cleavage site is predicted between residues 18 and 19 of the gene product, giving a secreted protein of 440 amino acids. The ProtParam tool at the ExPASy Proteomics Server (http://expasy.org/tools/protparam.html) was used to calculate the peptide molecular weight (46,978.5 Da), a theoretical isoelectric point (pI 4.51) and a molar extinction coefficient (ε 60,485 $M^{-1}cm^{-1}$).

The *Hir*Cel7A enzyme was purified from a static culture with Aspen sawdust as a carbon source. Precultures as above were mixed and divided equally to six 2.8 L Fernbach flasks containing 90 mL medium and 15 g Aspen sawdust (or Spruce powder or Whatman CF-11 cellulose, two of each). After four weeks, 500 mL 10 mM sodium acetate, pH 5.0, was added to each flask, and they were shaken for 2 days at 80 rpm before harvest. Filtration, concentration and buffer exchange was conducted, as above, but using 20 mM Bis-Tris-HCl, pH 6.5, and the final volume was 14 mL.

Purification of *Hir*Cel7A was performed by anion-exchange (two steps) and size-exclusion chromatography. The protein sample was fractionated on a DEAE-Sepharose column (50 mL) with a linear gradient of 0-1 M NaCl in 20 mM Bis-Tris, pH 6.5. Fractions containing *Hir*Cel7A, as judged by SDS-PAGE (8-25 % Phast gel; GE Healthcare) and activity measurements on p-nitrophenyl-β-lactoside (pNP-Lac; Sigma) were pooled together and fractionated on a Source 30Q column (10 mL) with a linear gradient of 0-1 M NaCl in 20 mM Bis-Tris, pH 6.5. Fractions containing *Hir*Cel7A, judged by SDS-PAGE and activity on Avicel cellulose, were pooled and concentrated and applied to a final size-exclusion purification step on a Superdex-200 column (~120 mL volume), equilibrated and eluted with 10 mM sodium acetate, pH 5.0. The purified protein was concentrated to 9.0 mg/mL, sterile filtered (0.2 µm) and stored in the freezer at -20°C. The protein concentration was determined by UV absorbance at 280 nm in a Nanodrop spectrophotometer (ThermoFisher Scientific, Wilmington, DE 19810 USA).

Enzyme activity assays on fractions from the DEAE-Sepharose fractionation were done by mixing 30 µL of fractions with 120 µL pNP-Lac in 0.1 M sodium acetate, pH 5.0. After 1 hour incubation at 30°C, the reaction was stopped by addition of 150 µL 0.5 M sodium carbonate, and released p-nitrophenol was measured spectrophotometrically at 405 nm. Enzyme activity assays of the Source 30Q fractionation were done with Avicel PH-101 cellulose (Fluka, Sigma-Aldrich) as substrate and estimation of solubilized reducing sugar with p-hydroxybenzoic acid hydrazide (PABAH; (6)) essentially as described in reference (7). 10 µL of fractions were mixed with 140 µL 1% Avicel cellulose suspension in 10 mM sodium acetate, pH 5.0, and incubated at 37°C, 220 rpm for 2 hours, then filtered in a 96-well vacuum filtration unit (1 µm glass-fiber; Porvair Sciences). 100 µL filtrate was mixed with 200 µL PABAH reagent (0.1 M PABAH, 0.2 M NaK tartrate, 0.5 M NaOH; freshly made), heated for 10 min in a boiling water bath and cooled down on ice, prior to absorbance measurements at 405 nm in a microtiter plate reader.

Protein Crystallization Methods

The initial search for crystallization conditions was done using the JCSG+ Suite sparse matrix screen (Qiagen) by sitting drop vapor diffusion in MRC-2 96-well plates (Hampton Research) using Oryx-8 crystallization robot (Douglas Instruments). 0.25 µL of 4.5 mg/mL *Hir*Cel7A protein in 10 mM sodium

acetate, pH 5.0, was mixed with 0.25 µL of each precipitant solution and equilibrated against 80 µL of the precipitant. Crystals appeared in four conditions, in drops: (A9) 0.2 M ammonium chloride, pH 6.3, 20% PEG 3350; (C1) 0.2 M Na chloride, 0.1 M phosphate-citrate, pH 4.2, 20% PEG 8000; (G2) 0.02 M $MgCl_2$, 0.1 M HEPES, pH 7.5, 22% polyacrylic acid 5100 Na salt; (H8) 0.8 M Na chloride, 0.1 M Bis-Tris, pH 5.5, 25% PEG 3350.

The crystals used for diffraction were grown in hanging-drop vapor diffusion experiments (8) at 20°C against 1 mL precipitant solution in 24-well crystallization plates. For the crystallization of the apo protein, 1 µL *Hir*Cel7A enzyme (8 mg/mL) was mixed with 1 µL of precipitant (20 mM $MgCl_2$, 0.1 M HEPES, pH 7.5, 22% polyacrylic acid 5100 Na salt). Crystals appeared within 24 hours without seeding, and could be flash frozen in liquid nitrogen without prior cryo-protectant soaking.

Crystals for ligand soaking were obtained with 0.8 µL *Hir*Cel7A protein (4 mg/mL) and 0.8 µL precipitant (20 mM $MgCl_2$, 0.1 M HEPES, pH 7.7, 15-22 % polyethylene glycol 3350) and microseeding after 24 h of equilibration. A drop with previously obtained crystals was mixed with 200 µL of precipitant and the solution was vortexed 1 min and then centrifuged at 13,000 x g for 1 min. Microseeds were transferred from the supernatant to the surface of the new crystallization drops using an acupuncture needle. Ligand solution was prepared by mixing 2 µL of 3 mM thio-β-D-xylopentaoside (SX5) in water with 4 µL of precipitant (20 mM $MgCl_2$, 0.1 M HEPES, pH 7.7, 22 % polyethylene glycol 3350) and equilibrated for 72 h against precipitant. Thereafter, 1 µL was added to drops with *Hir*Cel7A crystals. After 48 h soaking with ligand (~0.8 mM), single crystals were transferred with ready-made cryo loops and soaked for a few seconds in cryoprotectant solution (mother liquor with 22.5 % PEG 3350 + 10 % glycerol) prior to being flash frozen in liquid nitrogen.

MD Simulation Methods

MD simulations of three GH7 enzymes were constructed and performed, both with and without bound ligands. The three enzymes examined here are the *H. irregulare* Cel7A (*Hir*Cel7A) structure determined in this study, the *Hypocrea jecorina* Cel7A enzyme (*Hje*Cel7A), and the *Phanerochaete chrysosporium* Cel7D enzyme (*Pch*Cel7D). Docking of the ligand, system size, and protonation has been discussed in the main text.

Each system was initially setup, minimized, heated, and density equilibrated using CHARMM (9). The CHARMM force field with CMAP correction was used to describe the enzymes (9-11), while the modified TIP3P model was used to represent the water molecules (12,13), and the CHARMM C36 force field described the carbohydrates (14-16). The minimization procedure following setup as described in the manuscript included: 1,000 steps of Steepest Descent Minimization (SDM) with the protein and ligand (where applicable) fixed, 1,000 steps of SDM with the ligand fixed (where applicable), and 1,000 steps of SDM with no fixed atoms. After minimization, the systems were then heated from 100 K to 300 K over 20 ps with a temperature ramp of 50 K every 4 ps. Following heating, the system densities were equilibrated in the *NPT* ensemble with a Nosé-Hoover thermostat and barostat for 100 ps (17,18). The density-equilibrated systems were then used as the starting point to conduct production simulations.

NAMD was used for the 250-ns production simulations (19). For each simulation (including the equilibration in CHARMM), the non-bonded cutoff distance was 10 Å, the SHAKE algorithm (20) was used to fix the bond distances to hydrogen atoms, and the Particle Mesh Ewald method was used to treat long range electrostatics with a grid size of 80 x 80 x 80 and a Gaussian distribution width of 0.32 (21). All simulations used a 2 fs time step.

Residue cross-correlation maps for *Hir*Cel7A, *Hje*Cel7A, and *Pch*Cel7D were generated using a modified version of the AmberTools 1.5 (22,23) analysis module PTRAJ. The normal mode analysis was conducted using NAB (24,25), which is also part of the AmberTools package. The minimization was conducted using

the Limited-Memory Broyden-Fletcher-Goldfarb-Shanno Truncated Newton Conjugate technique with a 10^{-8} kcal/mol-Å tolerance for the RMS gradient. The parm99SB parameter set (26,27) and the Generalized Born model (28,29) with the Hawkins *et al.* pairwise approach were used for the minimization. ARPACK routines (30) with Cholesky decomposition and inversion were used for diagonalization of the Hessian matrix.

Xylose and HEPES binding in the crystal structure

Figure S1 shows the position of the HEPES and xylose residues at the -2 and +1 subsites, respectively in the SX5 structure of *Hir*Cel7A.

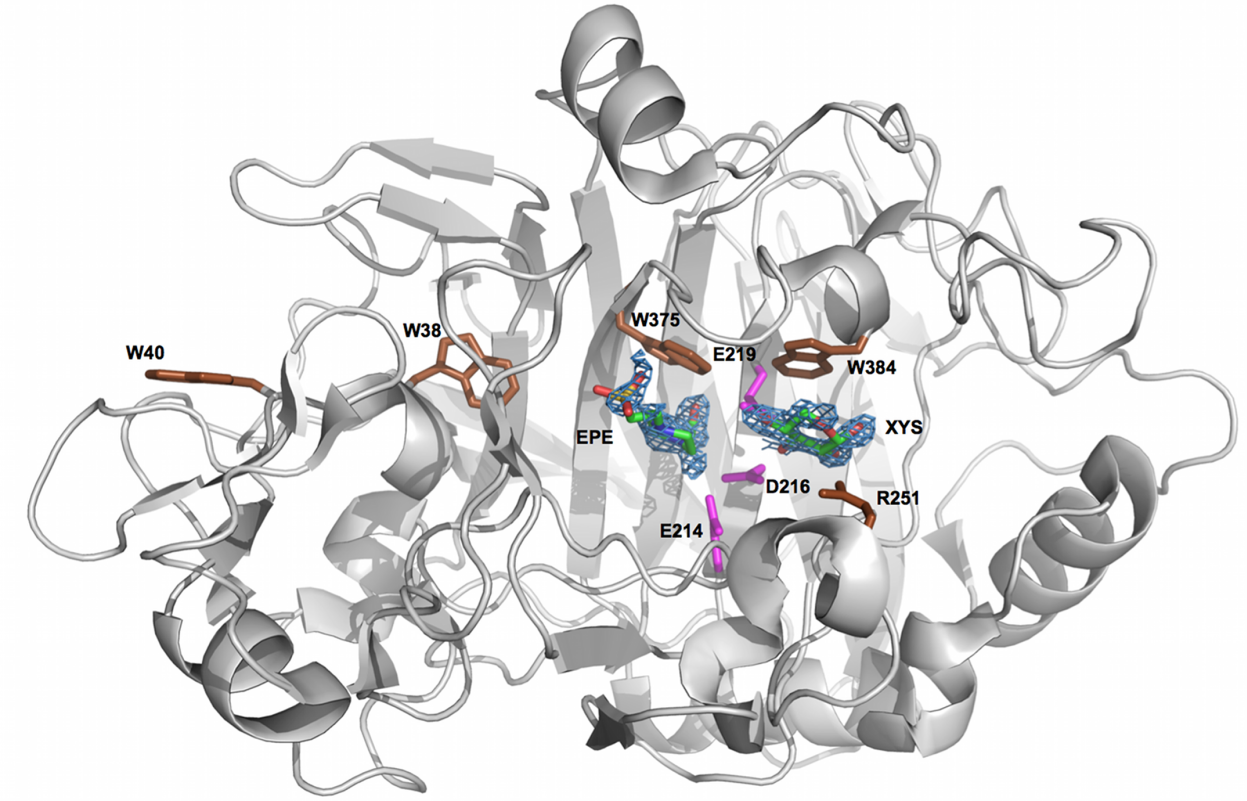

Figure S1. Overall structure of the *Hir*Cel7A with the HEPES (EPE) and xylose (XYS) ligands shown in stick representation, and the final $2F_o$-F_c electron density map contoured at 0.47 $e^-/Å^3$ around the ligands.

Additional Results from the Molecular Dynamics Simulations

Root Mean Square Deviation (RMSD) and Fluctuation (RMSF)

Simulations of *Hir*Cel7A were performed both with and without a ligand. For each of these simulations, we calculated the RMSD and RMSF, Figure S2A and S2B. The RMSDs of the *Hir*Cel7A structure with and without a ligand, *Hje*Cel7A, and *Pch*Cel7D are not substantially different and each enzyme is stable on the timescale of the MD simulation conducted in this study. Comparing the *Hir*Cel7A simulations with and without a ligand bound in the active site, the RMSF as a function of residue is similar across all residues except in the vicinity of Loop B2, as discussed in the main text and below. This loop is likely involved in processivity and thus may be expected to deviate upon docking of a ligand within the active site.

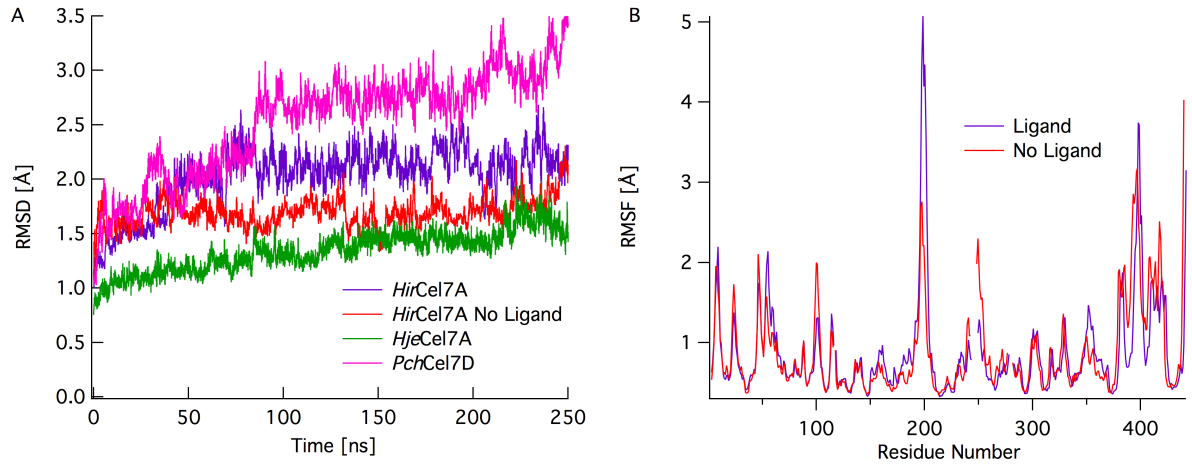

Figure S2. Comparison of the (A) RMSD over the 250 ns simulations of *Hir*Cel7A with and without a ligand bound in the active site, *Hje*Cel7A, and *Pch*Cel7D and the (B) RMSF as a function of the residue number for *Hir*Cel7A with and without a ligand bound in the active site.

Tyr101-Ligand Interactions

The distance of residues Trp40 and Tyr101 of *Hir*Cel7A from the glucose bound in the -7 subsite of the active site tunnel over the 250-ns simulation is shown in Figure S3. The distances were determined from the center of mass of the heavy atoms of the protein residues and the center of mass of the ring atoms of the glucose. The distance of these residues from the glucose is generally just over 4 Å. Occasional changes in conformation of the protein residues, as shown in the snapshots in the main text, result in short-lived spikes in the relative distance of these molecules. The longest-lived period of conformational change of these residues occurs around 175 ns for approximately 5 ns before the residues return to the preferred conformation. This 5 ns period corresponds to the movement of Tyr101 such that it resides temporarily over the -6 binding subsite.

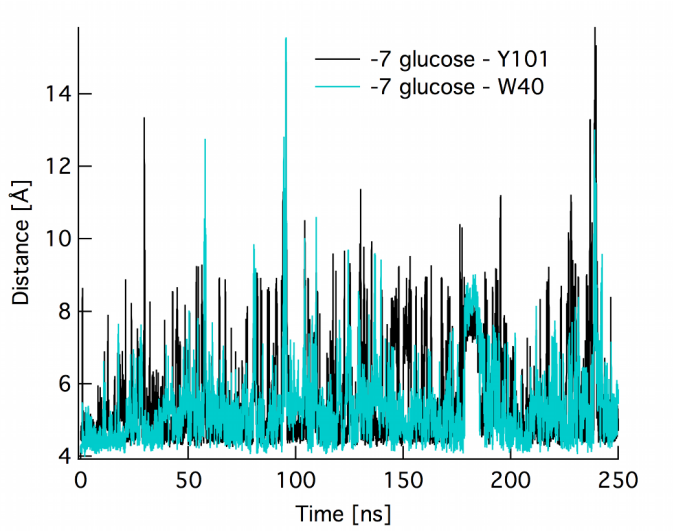

Figure S3. Distance of residues Trp40 and Tyr101 from the -7 subsite glucose over the 250-ns simulation of *Hir*Cel7A.

Loop Dynamics and Loop-Loop Contacts

Figure S4 illustrates loop dynamics from simulations using two different analysis methods. In panels S4A, S4B, and S4C, the RMSD of the side chain heavy atoms have been calculated as compared to the original oriented initial crystal structures. The loops examined, as labeled at the top of each panel, are identified by residue number in the manuscript. Loop A1 is shortest in *Pch*Cel7D and deviates little from its initial position as a result, as shown in Figure S4A. In both *Hir*Cel7A and *Hje*Cel7A, two unique states are observed with rapid conversion between the states. Loop B2 of *Pch*Cel7D appears to visit three states (Figure S4B), moving away from its initial position after approximately 75 ns. Loop B2 of *Hir*Cel7A deviates from its initial position after 175 ns. In *Hje*Cel7A on the other hand, Loop B2 is stable in the original position throughout the 250 ns simulation. In Figure S4C, Loop B3 in *Hir*Cel7A maintains the same relative position over the course of the simulation. In *Hje*Cel7A, Loop B3 moves away from the crystal structure position briefly, and returns to its original position for the majority of the simulation. Loop B3 of *Pch*Cel7D undergoes a large conformational change relative to the starting structure.

Figures S4D, S4E, and S4F illustrate the measured minimum distance between Loops A3 and B3, Loops B2 and A3, and Loops B2 and B3, respectively. This distance is determined on a per-frame basis between the center of mass of the each of the two loop selections. As illustrated in Figures S4D, S4E, and S4F, Loop B2 relative to both Loop A3 and B3 and Loop A3 to B3 are quite stable in *Hje*Cel7A, while both *Pch*Cel7D and *Hir*Cel7A exhibit a great deal more loop movement in these regions. This is additionally illustrated as a histogram of the observed loop states relative to other loops as described below.

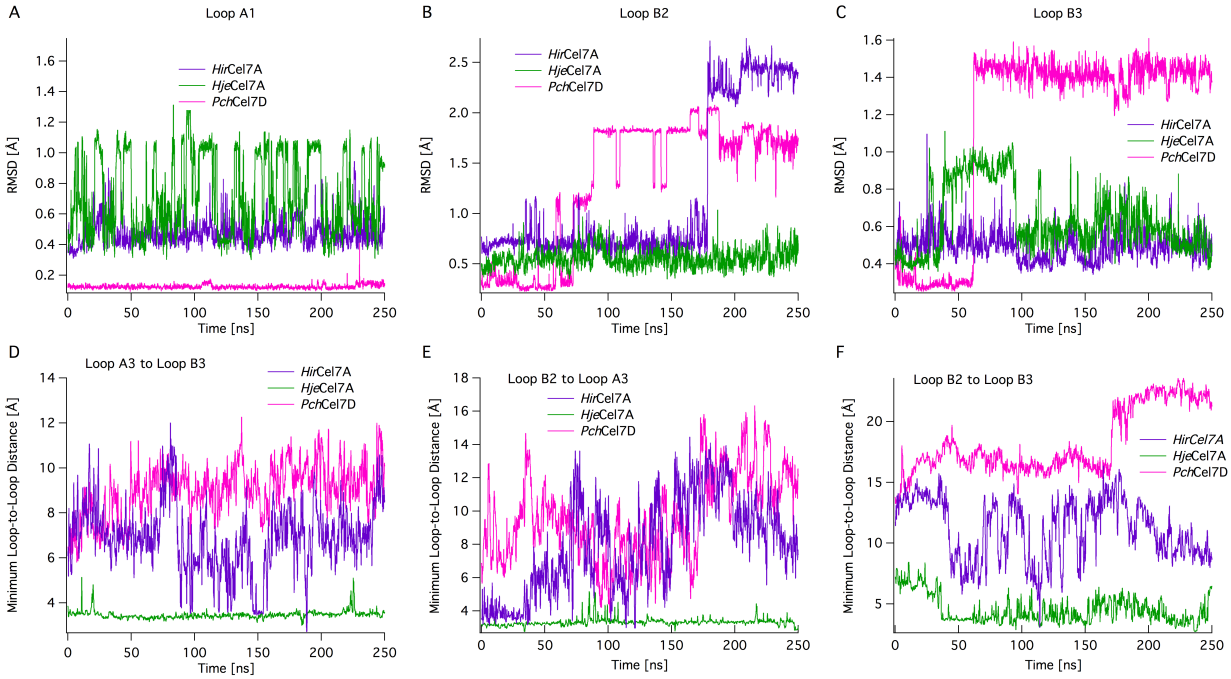

Figure S4. RMSD calculations of (A) Loop A1, (B) Loop B2, and (C) Loop B3 for *Hir*Cel7A, *Hje*Cel7A, and *Pch*Cel7D. The RMSD has been calculated oriented to the backbone atoms of the crystal structures so as to observe any conformational changes in the side chain atoms of the loops. (D) Distance of Loop A3 from neighboring Loop B3, as measured by center of mass. (E) Distance of Loop B2 from neighboring Loop A3, as measured by center of mass. (F) Distance of Loop B2 from neighboring Loop B3. These loops – A3, B2, and B3 – correspond to those hypothesized to participate in endo/exo initiation events.

For three sets of defined loops, as labeled in the manuscript, histograms of the minimum distance between the center of mass of each loop selection have been determined. The sets include Loop A1 to Loop B2 in Figure S5A, Loop B2 to Loop B3 in Figure S5B, and Loop B2 to Loop A3 in Figure S5C. The most

striking observation from each of the three sets of histograms is the apparent single state of the *Hje*Cel7A active site loops as illustrated by a single peak for each set. Both *Pch*Cel7D and *Hir*Cel7A active site loops included here more broadly sample conformational space relative to *Hje*Cel7A.

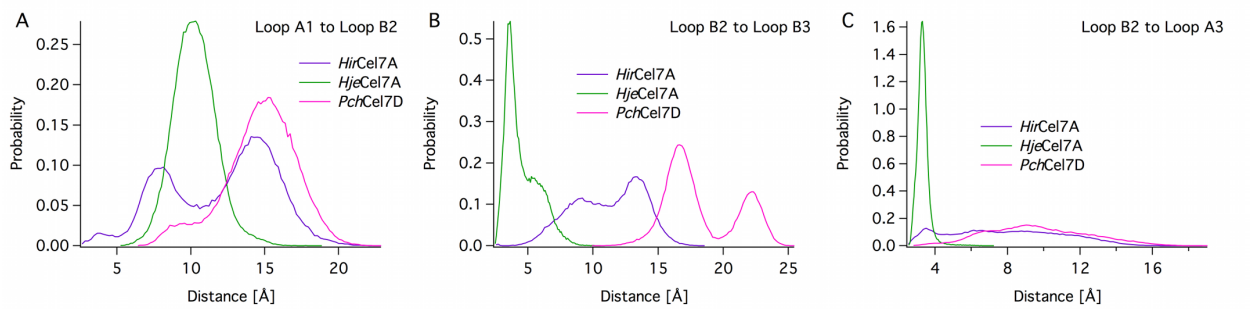

Figure S5. Histograms of the distances of the center of mass from Figure S4 of (A) Loop A1 to B2, (B) Loop B2 to B3, and (C) Loop B2 to A3 for *Hir*Cel7A, *Hje*Cel7A, and *Pch*Cel7D. These Loops correspond to those hypothesized to participate in endo/exo initiation events.

Cross-correlation Maps

Residue cross-correlation maps were generated for each of the three ligand bound simulations – *Hir*Cel7A, *Hje*Cel7A, and *Pch*Cel7D – as shown below in Figure S6. The correlation value scales for each of the three images have been normalized so as to allow comparison of the three enzymes as well as the residue-residue correlations. The general correlation patterns across the three Family 7 GHs are similar, as would be expected for enzymes with similar functionality.

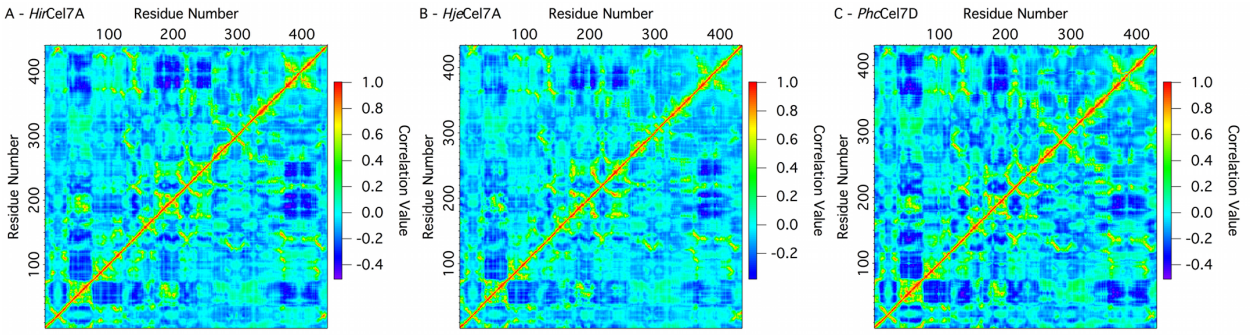

Figure S6. Residue cross-correlation maps from normal mode analysis (first 100 modes) of (A) *Hir*Cel7A, (B) *Hje*Cel7A, and (C) *Pch*Cel7D. Positive correlation values (red) represent positive correlation of residues, and negative correlation values (dark blue) represent negatively correlated motions.

References

1. Olson, Å., Aerts, A., Asiegbu, F., Belbahri, L., Bouzid, O., Broberg, A., Canback, B., Coutinho, P. M., Cullen, D., Dalman, K., *et al.* (2012) Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. *New Phytol.* **194**, 1001-1013
2. Stenlid, J. (1985) Population-structure of *Heterobasidion annosum* as determined by somatic incompatibility, sexual incompatibility, and isoenzyme patterns. *Can. J. Bot. - Rev. Can. Bot* **63**, 2268-2273
3. Kremer , S. M., Wood, P.M. . (1992) Evidence that cellobiose oxidase from *Phanerochaete chrysosporium* is primarily an Fe(III) reductase. Kinetic comparison with neutrophil NADPH oxidase and yeast flavocytochrome b2. . *Eur. J. Biochem.*, 205, 133–138

4. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Anal. Chem.* **68**, 850-858
5. Grigoriev, I. V., Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., Kuo, A., Minovitsky, S., Nikitin, R., Ohm, R. A.*, et al.* (2011) The Genome Portal Of The Department Of Energy Joint Genome Institute. *Nucleic Acids Res.*, 1-7
6. Lever, M. (1972) A new reaction for colorimetric determination of carbohydrates. *Anal. Biochem.* **47**, 273-279
7. Hori, C., Igarashi, K., Katayama, A., and Samejima, M. (2011) Effects of xylan and starch on secretome of the basidiomycete Phanerochaete chrysosporium grown on cellulose. *FEMS Microbiol. Lett.* **321**, 14-23
8. McPherson, A. (1982) Preparation and Analysis of Protein Crystals, Krieger Pub Co
9. 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.*, et al.* (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545-1614
10. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S.*, et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586-3616
11. Mackerell, A. D., Feig, M., and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comp. Chem.* **25**, 1400-1415
12. Durell, S. R., Brooks, B. R., and Ben-Naim, A. (1994) Solvent-induced forces between 2 hydrophilic groups. *J. Phys. Chem.* **98**, 2198-2202
13. Jorgensen, W. L., Chandrasekhar, J., and Madura, J. D. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935
14. Guvench, O., Greene, S. N., Kamath, G., Brady, J. W., Venable, R. M., Pastor, R. W., and Mackerell, A. D. (2008) Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.* **29**, 2543-2564
15. Guvench, O., Hatcher, E., Venable, R. M., Pastor, R. W., and MacKerell, A. D. (2009) CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J. Chem. Theory Comput.* **5**, 2353-2370
16. Guvench, O., Mallajosyula, S. S., Raman, E. P., Hatcher, E., Vanommeslaeghe, K., Foster, T. J., Jamison, F. W., and MacKerell, A. D. (2011) CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J. Chem. Theory Comput.* **7**, 3162-3180
17. Hoover, W. G. (1985) Canonical dynamics - equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695-1697
18. Nosé, S., and Klein, M. L. (1983) Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **50**, 1055-1076
19. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comp. Chem.* **26**, 1781-1802
20. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical-integration of cartesian equations of motion of a system with constraints - Molecular-dynamics of N-alkanes. *J. Comput. Phys.* **23**, 327-341
21. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh ewald method. *J. Chem. Phys.* **103**, 8577-8593

22. 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) The Amber biomolecular simulation programs. *J. Comp. Chem.* **26**, 1668-1688
23. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. A., Wang, J., Duke, R. E., and al., e. (2008) Amber 10
24. Macke, T., Svrcek-Seiler, W. A., Brown, R. A., Kolossvary, I., Y.J., B., and Case, D. A. (2011) NAB 6
25. 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
26. Junmei, W., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049-10741074
27. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **65**, 712-725
28. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995) Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett.* **246**, 122-129
29. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996) Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem.* **100**, 19824-19839
30. 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