

**Supporting Information for “Aromatic-mediated Carbohydrate Recognition in Processive
Serratia marcescens Chitinases”**

Suvamay Jana,^{1,†} Anne Grethe Hamre,^{2,†} Patricia Wildberger,² Matilde Mengkrog Holen,²
Vincent G. H. Eijsink,² Gregg T. Beckham,³ Morten Sørli,^{2,*} and Christina M. Payne^{1,*}

¹*Department of Chemical and Materials Engineering, University of Kentucky, Lexington, KY,
USA*

²*Department of Chemistry, Biotechnology, and Food Science, Norwegian University of Life
Sciences, Ås, Norway*

³*National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, USA*

† These authors contributed equally to this work.

* To whom correspondence should be addressed. E-mail: morten.sorlie@nmbu.no, Tel.: +47-
67232562 and Fax: +47-64965901 or christy.payne@uky.edu, Tel.: + 859-257-2902 and Fax: +
859-323-1929

1. Computational Methods

Initial System Preparation

The crystal structures from PDB entries 1EHN and 1E6N were used as input for construction of the ChiA and ChiB models, respectively.¹⁻² In total, 10 separate molecular dynamics (MD) simulations were constructed: ChiA and ChiB wild-type with and without a bound (GlcNAc)₆ ligand, and the chitinase variants ChiA-W167A, ChiA-W275A, ChiA-F396A,

ChiB-W97A, ChiB-W220A, and ChiB-F190A with a bound (GlcNAc)₆ ligand state.

The ChiA wild-type simulations were constructed from the 1EHN entry in the PDB (ChiA E315Q complexed with chitooctase) by reversing the E315Q mutation. The entire catalytic domain and fused *N*-terminal chitin binding domain ChiA structure was used in these models, as cleaving the chitin binding domain from the catalytic domain adversely impacted the stability of the catalytic domain fold during the simulations. Aromatic to alanine variants, W167A, W275A, and F396A, were constructed by simply removing the aromatic residue side chain atoms. In the simulations where the (GlcNAc)₆ ligand was bound to ChiA, the pyranose rings bound to the -4 through +2 subsites of 1EHN were retained, and the remaining two pyranose rings (in the solvent exposed -5 and -6 sites) were deleted. The ligand-free ChiA models were constructed by removing all the 1EHN pyranose rings from the cleft.

The wild-type ChiB model was constructed from the 1E6N entry in the PDB (ChiB E144Q in complex with chitopentaose) by reversing the E144Q mutation. Here, we cleaved the chitin binding domain (up to Leu-448) from the ChiB catalytic domain to minimize computational expense. The stability of the catalytic domain in the absence of the chitin binding domain was maintained over the course of simulations. Again, ChiB variants, W97A, W220A, and F190A, were constructed from the wild-type by removing the aromatic side chain atoms. The (GlcNAc)₆ ligand bound in the -3 to +3 subsites of the ChiB cleft was modeled from the ligands of two different chitinase structures. The protein backbones of PDBs 1E6N and 1OGG were aligned in PyMOL.³⁻⁴ The pyranose ring in the -2 through +3 binding sites were retained from the 1E6N structure, while the -3 binding site sugar originated from the 1OGG structure (ChiB D142N in complex with allosamidin).

In both ChiA and ChiB ligand-bound simulations, the catalytically competent active site

conformation was constructed by manually rotating the *N*-acetyl group of the -1 site pyranose ring and the side chains of the catalytic residues (Asp-313 and Glu-315 in ChiA and Asp-142 and Glu-144 in ChiB) in PyMOL to ensure that the catalytic residues and the -1 sugar reflected the distorted Michaelis complex of Family 18 chitinases.² The manual rotation of the -1 pyranose and catalytic residues was followed by additional stepwise minimization (100 steps of steepest descent minimization followed by 100 steps of adopted basis Newton Raphson minimization) to make sure the rotation did not adversely affect stability of the protein and the remaining 5 pyranose rings.

CHARMM was used to build and solvate the wild-type and aromatic variants of ChiA and ChiB.⁵ The input protonation states of the systems were determined using H++ at pH 6 and internal and external dielectric constants of 10 and 80, respectively.⁶⁻⁸ The disulfide bonds, between Cys115-Cys120 and Cys195-Cys218 for ChiA, and Cys328-Cys331 for ChiB, were specified based on crystal structures. The constructed systems were solvated with TIP3P water molecules, and sodium counterions were used to make the systems charge neutral. For ChiA, the final system size was approximately 120 Å x 120 Å x 120 Å, totaling roughly 175,000 atoms. The ChiB final system was 80 Å x 80 Å x 80 Å, for a total of approximately 52,000 atoms.

MD Simulations Protocols

The solvated systems were minimized, heated, and equilibrated using CHARMM. The minimization of the systems was conducted in a step-wise manner. First, the water molecules were minimized for 10,000 steps with the protein and ligand (if present) held rigid. The harmonic restraints on the ligand were then released, and the protein and water molecules were minimized for 10,000 steps. Finally, all restraints were removed, and the entire system was minimized for 10,000 steps. The minimized system was heated from 100 K to 300 K in steps of 4

ps with 50 K increments. The system density was equilibrated in the *NPT* ensemble at 1 atm and 300K with the Nosé-Hoover thermostat and barostat for 100 ps using a 2-fs time step.⁹⁻¹⁰ For all MD simulations conducted as part of this study, the CHARMM force field with the CMAP correction was used to describe the protein.^{5,11-12} The chitin oligomer was described with the CHARMM C36 carbohydrate force field,¹³⁻¹⁴ and water was described with the modified TIP3P model.¹⁵⁻¹⁶

The equilibrated systems were then simulated for 250 ns in the *NVT* ensemble using NAMD.¹⁷ A 2-fs time step was used in the velocity Verlet integration scheme. The Langevin thermostat was used for temperature control in the 250-ns MD simulations.¹⁸ The SHAKE algorithm was used to fix all hydrogen distances for computational efficiency.¹⁹ Non-bonded interactions used the following cutoffs: a non-bonded cutoff distance of 10 Å, a switching distance of 9 Å, and a non-bonded pair list distance of 12 Å. The Particle Mesh Ewald method was used to calculate the long-range electrostatics.²⁰ The PME grid spacing was 1 Å, and a 6th order b-spline and Gaussian distribution width of 0.320 Å was used.

Thermodynamic Integration (TI) Protocols

An equilibrated 25-ns snapshot from MD simulations was used as the input coordinates for the TI calculations. We used the dual-topology method with NAMD to determine the relative change in free energy.²¹⁻²³ The electrostatic and van der Waals calculations were decoupled in separate processes, each of which included 15 λ -windows ranging from λ values of 0 to 1 (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 and 1.0. More closely coupled windows were selected near the endpoints to improve the accuracy of the calculations.²⁴ The electrostatics and van der Waals calculations were equilibrated (0.5 ns) before collecting 14.5 ns of TI data. The change in

free energy, ΔG , for each set of simulations was evaluated using cubic spline Gaussian quadrature numerical method to integrate $dU/d\lambda$ over $\lambda = 0$ to 1. The protocol described by Steinbrecher et al. was followed to determine associated error.²⁵

Convergence Analysis

The simulations comprising the TI calculations were evaluated for appropriate window overlap and convergence as described by Pohorille et al.²⁴ For each van der Waals and electrostatic calculation of each variant evaluated, we plotted histograms of $dU/d\lambda$ for each λ window (Figure S1A). Window overlap was considered sufficient when at least 25% of neighboring windows overlapped. The autocorrelation function (ACF) of each dataset was also determined, as required for error analysis. An example is provided in Figure S1B. The convergence analysis illustrated in Figure S1 for the ChiB F190A system is representative of the trends observed for the thermodynamic data collected in this study.

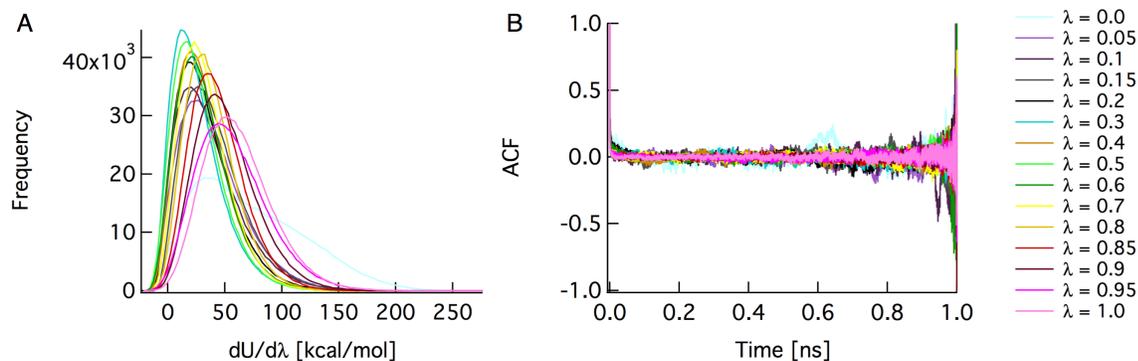
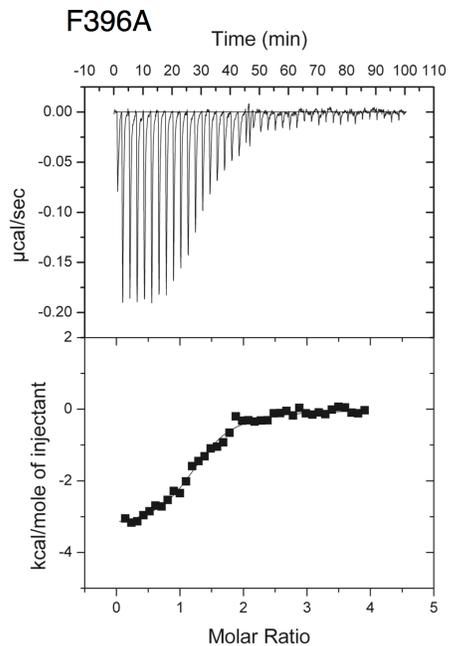
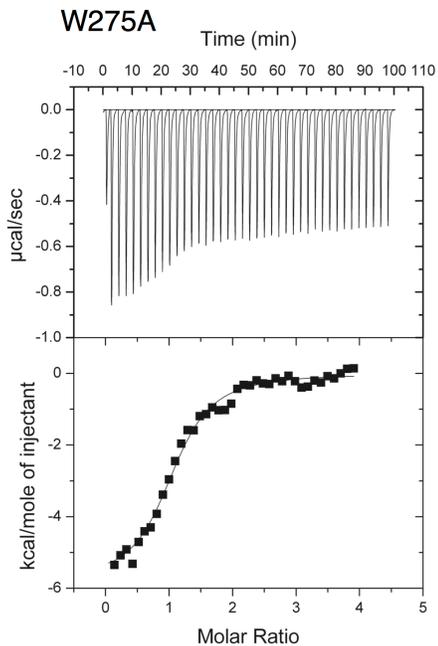
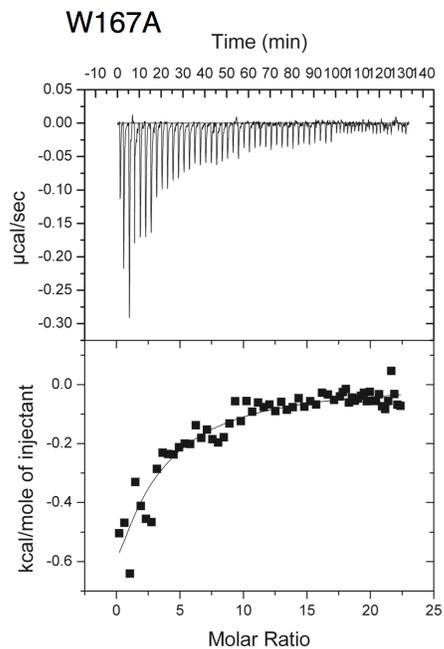
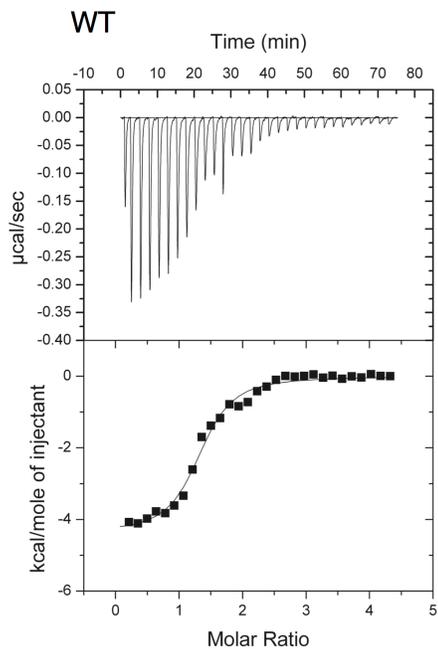


Figure S1. TI convergence assessment example from the reactant state of the van der Waals calculations for ChiB F190A. (A) Overlap of the $dU/d\lambda$ histograms determines if selected windows appropriately sample the change in potential energy. (B) The autocorrelation function (ACF) is used to determine the point at which data is no longer correlated, as required for error analysis.

2. Binding isotherms from isothermal titration calorimetry

A

ChiA



B

ChiB

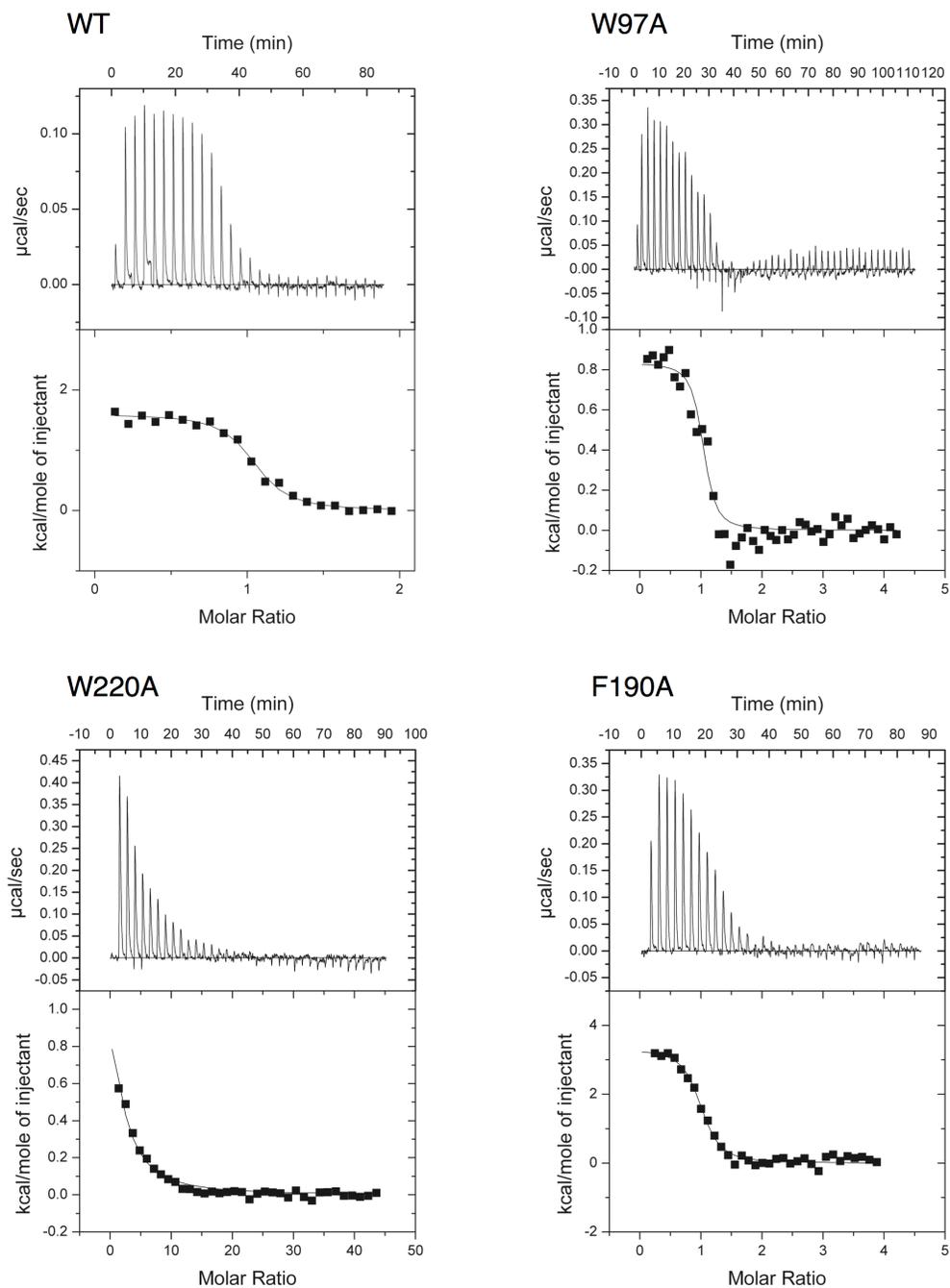


Figure S2. Binding isotherms with theoretical fits for (A) ChiA wild-type²⁶ and its aromatic variants at $t = 30^\circ\text{C}$ and $\text{pH} = 6.0$ as well as for (B) ChiB wild-type²⁷ and its aromatic variants at $t = 20^\circ\text{C}$ and $\text{pH} = 6.0$. In each case, the ligand was $(\text{GlcNAc})_6$.

3. Experimental determination of [(GlcNAc)₂]/[GlcNAc] for ChiA

Apparent processivity (P^{app}) is defined as the average number of consecutive catalytic cycles performed per initiated processive run along the crystalline substrate. It has previously been shown that the [(GlcNAc)₂]/[GlcNAc] ratio can be used as a measure of an enzyme's apparent processive ability.²⁸⁻²⁹ However, this approach has several pitfalls, including the assumption of the exclusive formation of odd numbered oligosaccharides from the first cleavage. This assumption may not hold in the case of all chitinases, as different enzymes may have different preferences for the orientation of the chain end relative to the polymer surface or a different probability of endo-mode initiation.²⁸ Additionally, the [(GlcNAc)₂]/[GlcNAc] ratio as a measure of processive ability requires 50% / 50% occupancy of the -3 to +3 and +2 to -4 or -2 to +4 subsites, which we know may not absolutely be true and likely changes with mutations in the active site.^{28, 30} Moreover, P^{app} tends to decrease as the substrate is consumed, most likely because the substrate becomes enriched with recalcitrant regions where there are fewer obstacle-free paths for processive enzyme attachment.^{29, 31-32} It is thus important to assess processivity during the early stages of the reaction.

Unlike P^{app} , which is dependent on the substrate, intrinsic processivity (P^{intr}) is the average number of successive catalytic events before dissociation of an “ideal” substrate and is thus the upper limit of P^{app} . P^{intr} was recently determined for ChiA-WT, ChiA-W167A, and ChiA-W275A on α -chitin, showing the wild type to have an approximately 3.5 and 1.5 fold higher intrinsic processivity than ChiA-W167A and ChiA-W275A, respectively. It was also shown that these three enzymes have a relatively high probability of endo-mode initiation with the values for the two mutants being even higher than that for the wild type.³³

Here, we have determined from initial degradation of β -chitin [(GlcNAc)₂]/[GlcNAc]

ratios of 27.2 ± 1.8 , 36.9 ± 4.9 , and 15.3 ± 0.6 for ChiA-W167A, ChiA-W275A, and ChiA-F396A, respectively. The value for ChiA-WT has previously been determined to be 30.1 ± 1.5 .²⁹ Taken at face value, this means that ChiA-W167A and W275A show a similar, initial processive ability as the wild type, while ChiA-F396A is less processive. However, as described above, we cannot extricate the effects of substrate occupancy and endo-initiated activity from the $[(\text{GlcNAc})_2]/[\text{GlcNAc}]$ values, and thus, $[(\text{GlcNAc})_2]/[\text{GlcNAc}]$ ratios are imperfect measures of processive ability in ChiA. Accordingly, we are unable to determine the processive ability of ChiA-F396A on β -chitin using $[(\text{GlcNAc})_2]/[\text{GlcNAc}]$ ratios.

4. Additional Results from MD simulations and TI calculations

Active site solvation

Active site solvation has been determined from MD simulations for wild-type ChiA, ChiB, and the aromatic variants for comparison to the entropic energy changes determined by ITC. To determine active site solvation, we counted the number of water molecules present within 3.5 Å of the ligand on a per-binding-site basis in VMD every 0.1 ns. The value for a given binding site was then averaged over the 250 ns simulations. A higher number of water molecules is generally indicative of unfavorable contributions to entropy due to solvation effects.

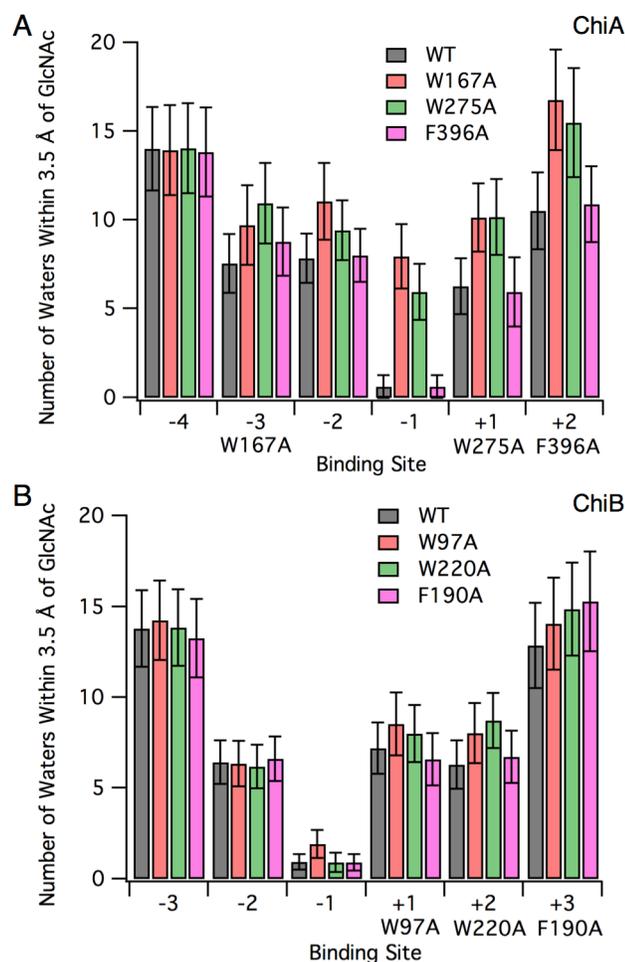


Figure S3. Active site solvation is represented by the average number of water molecules within 3.5 Å of the ligand in each of the six binding sites in wild-type and aromatic variants of (A) ChiA and (B) ChiB. Labels below the *x*-axis indicate the position of the aromatic residues relative to the binding site. Error bars were obtained through 2.5 ns block averaging.

RMSF of ChiB catalytic residues

The MD simulations were initialized such that the Glu-144 side chain formed a hydrogen bond with both Asp-142 and the glycosidic linkage between the +1 and -1 pyranose rings; however, in each case, the Glu-144 side chain eventually rotated such that only the Asp-142 hydrogen bond persisted. In wild-type ChiB, this conformational change occurred after 45 ns

(Figure S4A & B). In both ChiB-W220A (not shown) and ChiB-F190A (Figure S4C), the reorientation of the Glu-144 side chain occurred after ~100 ns. In W97A, the side chain of Asp-142 rotated almost immediately after the simulation started (within 10 ns) to interact with Asp-140 instead of Glu-144 and remained in that state for the remainder of the simulation (not shown). Though the catalytic side chain of Glu-144 and Asp-142 behave similarly in W220A and F190A after 100 ns, the W220A mutation still results in higher fluctuation in the catalytic residues, as the mutation also affects the flexibility of Tyr-214 and Asp-215.

The ChiB-F190A conformational change proceeds via a slightly different route than wild-type, which also adds to the early-stage increase in RMSF of the catalytic residues. The ChiB-F190A Glu-144 side chain hydrogen bonds with the Asp-142 side chain (Figure S4D) throughout the simulation and the glycosidic linkage intermittently (Figure S4F), as does wild-type. However, the reduced interaction of Glu-144 in ChiB-F190A with the glycosidic linkage encourages hydrogen bond formation with Tyr-145 (Figure S4E), which does not occur in wild-type. This behavior subsides after the simulation reaches equilibrium.

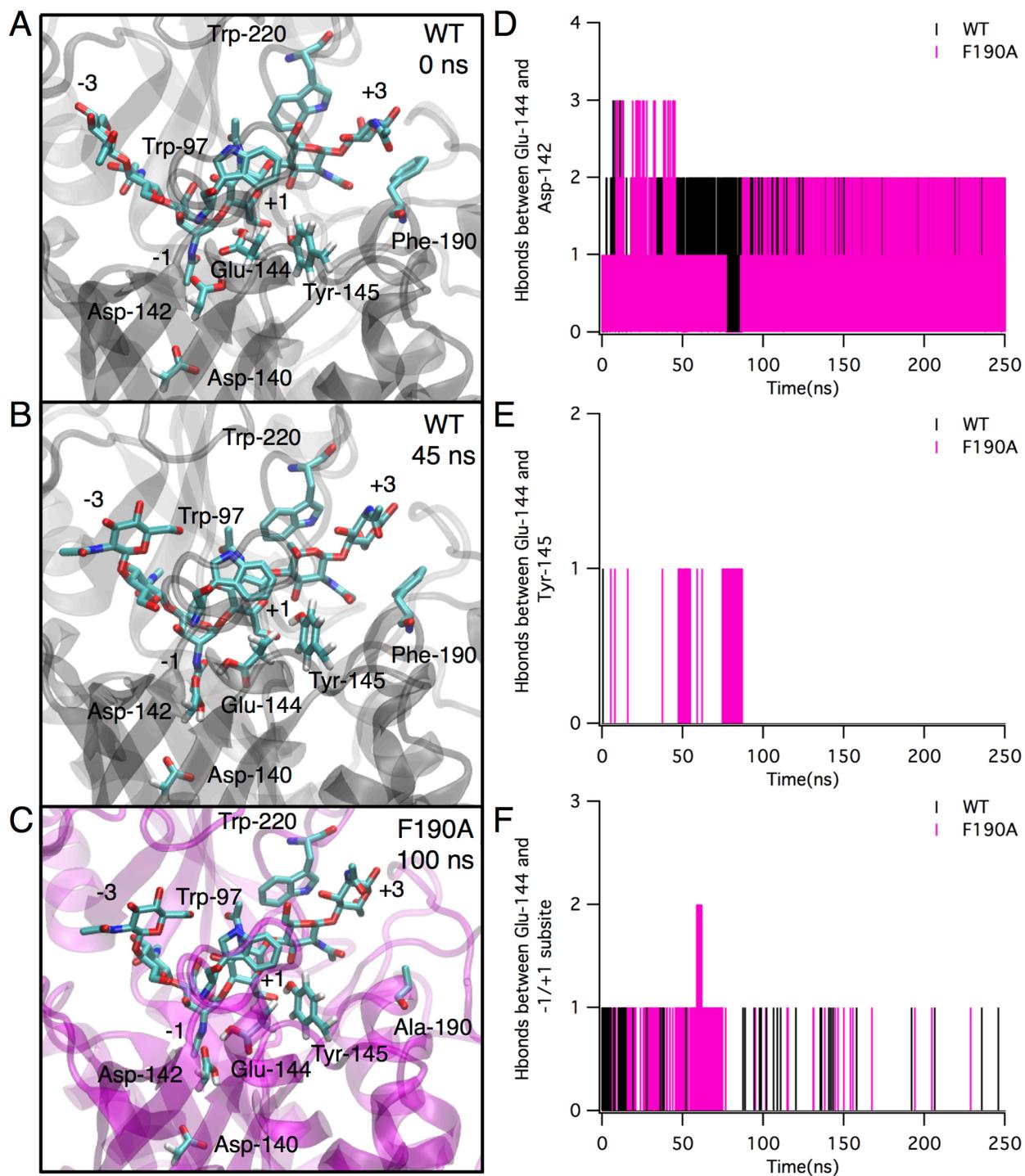


Figure S4. (A) Snapshot from the ChiB wild-type (WT) MD simulation at 0 ns, where the Glu-144 forms a hydrogen bond with Asp-142 and the glycosidic oxygen between -1 and +1 sites. (B) Snapshot from the ChiB WT simulation at 45 ns, where the Glu-144 side chain reorients to

interact with only Asp-142. This state is maintained through the remainder of the simulation. (C) Snapshot of ChiB-F190A at 100 ns, where the Glu-144 side chain reorients in the same fashion as ChiB-WT Glu-144, interacting with only Asp-142, which persisted through the remainder of the simulation. Panels (D) through (E) illustrate hydrogen bond formation (Hbonds) over the 250 ns simulation between: (D) Glu-144 and Asp-142 in ChiB-WT and ChiB-F190A; (E) Glu-144 and Tyr-145 in ChiB-WT and ChiB-F190A; and (F) Glu-144 and the glycosidic oxygen between the -1/+1 subsite in ChiB-WT and ChiB-F190A. A distance cut off of 3.4 Å and an angle cutoff of 60° from linear was used to determine the number of hydrogen bonds formed between the selections.³⁴

TI data

As described in the Computational Methods section of the manuscript (Figure 2), the relative change in binding free energy due to mutation ($\Delta\Delta G_{TI}$) is determined by calculating individual electrostatic and van der Waals components in the presence of a ligand (bound) and in the absence of a ligand (free) (Table S1). The sum of the electrostatic and van der Waals components describe the free energy of a given state (ΔG), where $\Delta G_{Bound(M-WT)}$ and $\Delta G_{Free(M-WT)}$ describe the bound and free ligand states, respectively. The difference of $\Delta G_{Bound(M-WT)}$ and $\Delta G_{Free(M-WT)}$, as described in equation 3 of the manuscript, is the relative change in binding free energy, $\Delta\Delta G_{TI}$. As an example, in the ligand bound state of ChiA, the mutation of residue Trp-167 to alanine causes a free energy change, $\Delta G_{Bound(M-WT)}$, of 12.8 ± 0.4 kcal/mol, which is the sum of the 9.3 ± 0.1 kcal/mol from electrostatic interactions and the 3.5 ± 0.4 kcal/mol from van der Waals (Table S1). Similarly in the ligand free state, the mutation changes free energy,

$\Delta G_{Free(M-WT)}$, by 7.8 ± 0.1 kcal/mol (electrostatic: 5.7 ± 0.0 kcal/mol and van der Waals: 2.1 ± 0.1 kcal/mol). The difference of $\Delta G_{Bound(M-WT)}$ and $\Delta G_{Free(M-WT)}$ gives a $\Delta\Delta G_{TI}$ of 5.0 ± 0.4 kcal/mol, which is the overall effect of mutating Trp-167 to alanine upon ligand binding.

Table S1. Detailed ligand binding free energies calculated from TI

ChiA	Energy	Bound		Free	
		Energy (kcal/mol)	Error (kcal/mol)	Energy (kcal/mol)	Error (kcal/mol)
W167A	Electrostatics	9.3	± 0.1	5.7	± 0.0
	VDW	3.5	± 0.4	2.1	± 0.1
	ΔG	12.8	± 0.4	7.8	± 0.1
	$\Delta\Delta G_{TI}$ (kcal/mol)	5.0 ± 0.4			
W275A	Electrostatics	8.5	± 0.1	5.2	± 0.0
	VDW	2.5	± 0.3	2.2	± 0.1
	ΔG	11.0	± 0.3	7.4	± 0.1
	$\Delta\Delta G_{TI}$ (kcal/mol)	3.6 ± 0.4			
F396A	Electrostatics	4.3	± 0.0	3.5	± 0.0
	VDW	-2.8	± 0.1	-1.8	± 0.1
	ΔG	1.5	± 0.1	1.7	± 0.1
	$\Delta\Delta G_{TI}$ (kcal/mol)	-0.2 ± 0.2			
ChiB	Energy	Bound		Free	
		Energy (kcal/mol)	Error (kcal/mol)	Energy (kcal/mol)	Error (kcal/mol)
W97A	Electrostatics	8.4	± 0.1	6.8	± 0.1
	VDW	2.4	± 0.4	1.2	± 0.2
	ΔG	10.8	± 0.4	8.0	± 0.2
	$\Delta\Delta G_{TI}$ (kcal/mol)	2.8 ± 0.4			
W220A	Electrostatics	9.6	± 0.1	7.8	± 0.0
	VDW	2.9	± 0.3	1.2	± 0.1
	ΔG	12.5	± 0.3	9.0	± 0.1
	$\Delta\Delta G_{TI}$ (kcal/mol)	3.5 ± 0.3			
F190A	Electrostatics	-8.4	± 0.0	-7.7	± 0.0
	VDW	6.3	± 0.1	5.9	± 0.1
	ΔG	-2.1	± 0.1	-1.8	± 0.1
	$\Delta\Delta G_{TI}$ (kcal/mol)	-0.3 ± 0.2			

5. Multiple sequence alignment

```

>gi|48425130|SmarChiB| 58 .....LECAWDPATNDA.....K
>gi|497397203|DspChiB| 269 .....ANGEVVVGDVYADTDKAFPGD..CWEFG...C
>gi|573461615|LputChiB| 225 .....TNGEIVVGDYADTDKFFPGD..CWDPG...C
>gi|550993861|LsacChiB| 221 .....ANGEIVVGDYADTDKAFPGD..CWDPG...C
>gi|557622580|BtoyChiB| 1  MQNKEVPTLVLGEPWADVTKSYPGSGTTWEDCDKYA
>gi|163938376|BweiChiB| 105  LQNKEVPTLVLGEPWADVTKSYPGSGTTWEDCDKYA
>gi|376264400|BcerChiB| 105  LQNKEVPTLVLGEPWADVTKSYPGSGTTWEDCDKYA
>gi|345548065|GsteChiB| 66   LQNKEVPTLVLGEPWADVTKSYPGSGTTWEDCDKYA
>gi|259558661|BthuChiB| 73   LQNKEVPTLVLGEPWADVTKSYPGSGTTWEDCDKYA
>gi|56962420|BclaChiB| 101  G.VIDVPTNGSIVMGDPWIDAQKSNPGD..TWE...P
>gi|15613479|BhalChiB| 102  G.VIDVPTNGSIVMGDPWIDAQKSNPGD..TWE...P
>gi|498185052|PelgChiB| 332  G.NINVPTNGSIVQGDWADTGMSTYPGD..TWDQ...P
>gi|517582257|PsanChiB| 105  GQTIHVPTNGTIVLGDWIDTGMSTYPGD..VWDQ...P
>gi|927653|KzopChiB| 106  SQTINVPTNGTIVLGDWIDTGMSTYPGD..TWDQ...P
>gi|20150306|BcirChiB| 74   SQTINVPTNGTIVLGDWIDTGMSTYPGD..TWDQ...P
>gi|90023510|SdegChiB| 262  .....LECFIDVEKAD.....A
>gi|321173846|PspChiB| 58   .....LECAWDPATNDA.....K
>gi|518650018|SplyChiB| 58   .....LECAWDPATNDA.....K
>gi|22652066|SliqChiB| 58   .....LECAWDPATNDA.....K
>gi|157371713|SproChiB| 58   .....LECAWDPATNDA.....K
consensus>70 .....ng..vlgd.w.d..k...g....wd.....

```

```

>gi|48425130|SmarChiB| 71  ARDVVNRLEALKAHNP.SLRIMFSIGGWYYSN.DLGVSHA
>gi|497397203|DspChiB| 296  KRGNFNQLNKLKQKHPHLKTLISVGGWTWSNNDLGVSHA
>gi|573461615|LputChiB| 252  KRGNFNQLQKLKAAYPHLKTLISVGGWTWSGGDLGVSHA
>gi|550993861|LsacChiB| 248  KRGNFNQLTKLKQKYPHLKTLISVGGWTWSGGDLGVSHA
>gi|557622580|BtoyChiB| 39   RCGNFGBLKR.LKAKYPHLKTIISVGGWTWSNRDLGVSHA
>gi|163938376|BweiChiB| 143  RCGNFGBLKR.LKAKYPHLKTIISVGGWTWSNRDLGVSHA
>gi|376264400|BcerChiB| 143  RCGNFGBLKR.LKAKYPHLKTIISVGGWTWSNRDLGVSHA
>gi|345548065|GsteChiB| 104  RCGNFGBLKR.LKAKYPHLKTIISVGGWTWSNRDLGVSHA
>gi|259558661|BthuChiB| 111  RCGNFGBLKR.LKAKYPHLKTIISVGGWTWSNRDLGVSHA
>gi|56962420|BclaChiB| 132  IRGNFKQLQKLKEEHPHLKTLISVGGWWSNRDLGVSHA
>gi|15613479|BhalChiB| 133  LRGNFKQLNKLKEEHPHLKTLISVGGWTWSNRDLGVSHA
>gi|498185052|PelgChiB| 363  LKGSFNQLIKLKKANPNLKTLISVGGWWSNRDLGVSHA
>gi|517582257|PsanChiB| 137  YAGNIHQLNKLKQANPHLKTIISVGGWTWSNRDLGVSHA
>gi|927653|KzopChiB| 138  IAGNINQLNKLKQINPNLKTIISVGGWTWSNRDLGVSHA
>gi|20150306|BcirChiB| 106  IAGNINQLNKLKQINPNLKTIISVGGWTWSNRDLGVSHA
>gi|90023510|SdegChiB| 275  ETQIIAELQALKNWNADLKILFSVGGWAESNDAAETVS
>gi|321173846|PspChiB| 71  ARDVVNRLEALKAHNP.SLRIMFSIGGWYYSN.DLGVSHA
>gi|518650018|SplyChiB| 71  ARDVVGRLEALKAHNP.SLRIMFSIGGWYYSN.DLGVSHA
>gi|22652066|SliqChiB| 71  ARDVVGRLEALKAHNP.SLRIMFSIGGWYYSN.DLGVSHA
>gi|157371713|SproChiB| 71  ARDVVSRLEALKAHNP.SLRIMFSIGGWYYSN.DLGVSHA
consensus>70 ..g...qL..LK...p.Lkt.is!GGW.wSn.....

```

```

>gi|48425130|SmarChiB| 175 QALPYQLTTIAGAGGAFFLSRYYSKLAQIVAPLDYINIM
>gi|497397203|DspChiB| 401 E...YLLTTIASGANPNYV..KNTELQVQAQLDWINIM
>gi|573461615|LputChiB| 357 K...YLLTTIASGASPNYV..QNTELQVKDYVDWINIM
>gi|550993861|LsacChiB| 353 H...YLLTTIAGASPSYY.NDNTELDKVKNYVDWINIM
>gi|557622580|BtoyChiB| 146 Q...YLLTTIASGASQRYA..DHTELKKISQILDWINIM
>gi|163938376|BweiChiB| 250 Q...YLLTTIASGASQRYA..DHTELKKISQILDWINIM
>gi|376264400|BcerChiB| 250 Q...YLLTTIASGASQRYA..DHTELKKISQILDWINIM
>gi|345548065|GsteChiB| 211 Q...YLLTTIASGASQRYA..DHTELKKISQILDWINIM
>gi|259558661|BthuChiB| 218 Q...YLLTTIASGASQRYA..DHTELKKISQILDWINIM
>gi|56962420|BclaChiB| 239 E...YLVTIASGASSEYV..ENNKLAEIAEVVDWINIM
>gi|15613479|BhalChiB| 240 D...YLLTTIASGASPGYV..ENNKLNEIAEIVDWINIM
>gi|498185052|PelgChiB| 470 Q...YLLTTIASGAGPNYV..NNTELSKIAQTLDWINIM
>gi|517582257|PsanChiB| 244 H...YLLTTIASGASPAFV..NNTELGNIASIVDWINIM
>gi|927653|KzopChiB| 245 K...YLLTTIASGASTTYA..ANTELANIASIVDWINIM
>gi|20150306|BcirChiB| 213 K...YLLTTIASGASATYA..ANTELAKIAAIVDWINIM
>gi|90023510|SdegChiB| 373 ..NGELVTIAGAGGAFFLSRYYSKLAAIVEQLDFINIM
>gi|321173846|PspChiB| 175 QALPYQLTTIAGAGGAFFLSRYYSKLAQIVAPLDYINIM
>gi|518650018|SplyChiB| 175 QALPYQLTTIAGAGGAFFLSRYYSKLPQIVASLDYINIM
>gi|22652066|SliqChiB| 175 QALPYQLTTIAGAGGAFFLSRYYSKLPQIVASLDYINIM
>gi|157371713|SproChiB| 175 QALPYQLTTIAGAGGAFFLSRYYSKLPQIVASLDYINIM
consensus>70 q...y11TIA.ga...%.....1...!....DwINiM

```

```

>gi|48425130|SmarChiB| 213 TYDLACPWEKVT..NHQAALFGDAAGPTFYNALREANL
>gi|497397203|DspChiB| 434 TYDFHGGWENVS..GHNAPLYFDPKDPS.....
>gi|573461615|LputChiB| 390 TYDFHGGWENVS..GHNAPLYVDPSDPS.....
>gi|550993861|LsacChiB| 387 TYDFRGWENAN..GHNAPLYADPSDPH.....
>gi|557622580|BtoyChiB| 179 TYDFHGGWEATS..NHNAALYKDPNDPA.....
>gi|163938376|BweiChiB| 283 TYDFHGGWEATS..NHNAALYKDPNDPA.....
>gi|376264400|BcerChiB| 283 TYDFHGGWEATS..NHNAALYKDPNDPT.....
>gi|345548065|GsteChiB| 244 TYDFHGGWEATS..NHNAALYKDPNDPA.....
>gi|259558661|BthuChiB| 251 TYDFHGGWEATS..NHNAALYKDPNDPA.....
>gi|56962420|BclaChiB| 272 TYDFNGGWONTS..GHNAPLYFDPATAD.....
>gi|15613479|BhalChiB| 273 TYDFNGGWONIS..GHNAPLYYDPATAN.....
>gi|498185052|PelgChiB| 503 TYDFHGGWESKS..GHNAPLYFDPADNS.....
>gi|517582257|PsanChiB| 277 TYDFNGGWQMT..AHNAPLYFDPAAEA.....
>gi|927653|KzopChiB| 278 TYDFNGAWQKVS..AHNALLNYDPAASA.....
>gi|20150306|BcirChiB| 246 TYDFNGAWQKIS..AHNAPLNYDPAASA.....
>gi|90023510|SdegChiB| 409 TYDLNGPWNGVTKTNFHAHLYGNNQEPRFYNALREADL
>gi|321173846|PspChiB| 213 TYDLACPWEKVT..NHQAALFGDAAGPTFYNALREANL
>gi|518650018|SplyChiB| 213 TYDLACPWEKIT..NHQAALFGDSAGPTFYNALREANL
>gi|22652066|SliqChiB| 213 TYDLACPWEKIT..NHQAGLFGDSAGPTFYNALREANL
>gi|157371713|SproChiB| 213 TYDLACPWEKIT..NHQAGLFGDSAGPTFYNALREANL
consensus>70 TYDf.G.W#.....hNA.LY.#p..p.....

```

Figure S5. Multiple sequence alignment of species phylogenetically related to *S. marcescens* ChiB. The multiple sequence alignment was generated with Clustal Omega³⁵⁻³⁷ and prepared for publication using the ESPrift 3.0 server.³⁸ The cutoff consensus used was 70%. Fully conserved residues are shown in bold white characters on a red background. Chemically similar residues are shown in bold black characters boxed in yellow. In the consensus line, uppercase letters represent full conservation of the residues at a given sequence location (100%), and lowercase

letters represent similarity greater than 70%. The symbol “.” represents no conservation at a given location. Additional symbols include: “!” representing Ile or Val, “\$” representing Leu or Met, “%” representing Phe or Tyr, and “#” representing Asn, Asp, Gln, or Glu. ChiB residues relevant for the present study are indicated by black arrows. The alignment shows that Trp-97 and Trp-220 of *S. marcescens* ChiB are well-conserved aromatic residues, while Phe-190 is quite variable. The short abbreviations of the species used for sequence comparisons are: Smar = *Serratia marcescens*, Dsp = *Desmospora sp.*, Lput = *Laceyella putida*, Lsac = *Laceyella sacchari*, Btoy = *Bacillus toyonensis*, Bwei = *Bacillus weihenstephanensis*, Bcer = *Bacillus cereus*, Gste = *Geobacillus stearothermophilus*, Bthu = *Bacillus thuringiensis serovar kurstaki*, Bcla = *Bacillus clausii*, Bhal = *Bacillus halodurans*, Pelg = *Paenibacillus elgii*, Psan = *Paenibacillus sanguinis*, Kzop = *Kurthia zopfii*, Bcir = *Bacillus circulans*, Sdeg = *Saccharophagus degradans*, Psp = *Pseudomonas sp.*, Sply = *Serratia plymuthica*, Sliq = *Serratia liquefaciens*, Spro = *Serratia proteamaculans*.

References

1. Papanikolau, Y.; Prag, G.; Tavlas, G.; Vorgias, C. E.; Oppenheim, A. B.; Petratos, K., High Resolution Structural Analyses of Mutant Chitinase A Complexes with Substrates Provide New Insight into the Mechanism of Catalysis. *Biochemistry* **2001**, *40* (38), 11338-11343.
2. van Aalten, D. M. F.; Komander, D.; Synstad, B.; Gaseidnes, S.; Peter, M. G.; Eijsink, V. G. H., Structural Insights into the Catalytic Mechanism of a Family 18 Exo-Chitinase. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (16), 8979-8984.
3. Vaaje-Kolstad, G.; Houston, D. R.; Rao, F. V.; Peter, M. G.; Synstad, B.; van Aalten, D. M. F.; Eijsink, V. G. H., Structure of the D142N Mutant of the Family 18 Chitinase ChiB from *Serratia marcescens* and its Complex with Allosamidin. *Biochim. Biophys. Acta, Proteins Proteomics* **2004**, *1696* (1), 103-111.
4. Schrödinger, L. *The PyMOL Molecular Graphics System*, 1.5.0.4; 2010.
5. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., et al., CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545-1614.
6. 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.* **2012**, *40* (W1), W537-W541.
7. Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A., A Simple Clustering Algorithm can be Accurate Enough for Use in Calculations of pKs in Macromolecules. *Proteins: Struct., Funct., Genet.* **2006**, *63* (4), 928-938.

8. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., H⁺⁺: A Server for Estimating pK(a)s and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368-W371.
9. Nosé, S., and Klein, M. L., Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50* (5), 1055-1076.
10. Hoover, W. G., Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695-1697.
11. 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., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586-3616.
12. Mackerell, A. D.; Feig, M.; Brooks, C. L., 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. Comput. Chem.* **2004**, *25* (11), 1400-1415.
13. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D., Additive Empirical Force Field for Hexopyranose Monosaccharides. *J. Comput. Chem.* **2008**, *29* (15), 2543-2564.
14. Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5* (9), 2353-2370.
15. Durell, S. R.; Brooks, B. R.; Bennaim, A., Solvent-Induced Forces between 2 Hydrophilic Groups. *J. Phys. Chem.* **1994**, *98* (8), 2198-2202.

16. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926-935.
17. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781-1802.
18. Grest, G. S.; Kremer, K., Molecular Dynamics Simulation for Polymers in the Presence of a Heat Bath. *Phys. Rev. A* **1986**, *33* (5), 3628-3631.
19. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of Cartesian Equations of Motion of a System with Constraints - Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327-341.
20. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577-8593.
21. Kirkwood, J. G., Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3* (5), 300-313.
22. Straatsma, T. P.; Mccammon, J. A., Multiconfiguration Thermodynamic Integration. *J. Chem. Phys.* **1991**, *95* (2), 1175-1188.
23. Frenkel, D., and Smit, B., *Understanding Molecular Simulations: From Algorithms to Applications*. 2nd ed.; Academic Press: 2002.
24. Pohorille, A.; Jarzynski, C.; Chipot, C., Good Practices in Free-Energy Calculations. *J. Phys. Chem. B* **2010**, *114* (32), 10235-10253.

25. Steinbrecher, T.; Mobley, D. L.; Case, D. A., Nonlinear Scaling Schemes for Lennard-Jones Interactions in Free Energy Calculations. *J. Chem. Phys.* **2007**, *127* (21), 214108/1-214108/13.
26. Hamre, A. G.; Jana, S.; Holen, M. M.; Mathiesen, G.; Väljamäe, P.; Payne, C. M.; Sørlie, M., Thermodynamic Relationships with Processivity in *Serratia marcescens* Family 18 Chitinases. *J. Phys. Chem. B* **2015**, *119* (30), 9601-9613.
27. Norberg, A. L.; Karlsen, V.; Hoell, I. A.; Bakke, I.; Eijsink, V. G. H.; Sørlie, M., Determination of Substrate Binding Energies in Individual Subsites of a Family 18 Chitinase. *FEBS Lett.* **2010**, *584* (22), 4581-4585.
28. Horn, S. J.; Sørlie, M.; Vårum, K. M.; Väljamäe, P.; Eijsink, V. G. H., Measuring Processivity. *Cellulases* **2012**, *510*, 69-95.
29. Hamre, A. G.; Lorentzen, S. B.; Väljamäe, P.; Sørlie, M., Enzyme Processivity Changes with the Extent of Recalcitrant Polysaccharide Degradation. *FEBS Lett.* **2014**, *588* (24), 4620-4624.
30. Norberg, A. L.; Dybvik, A. I.; Zakariassen, H.; Mormann, M.; Peter-Katalinic, J.; Eijsink, V. G. H.; Sørlie, M., Substrate Positioning in Chitinase A, A Processive Chito-Biohydrolase from *Serratia marcescens*. *FEBS Lett.* **2011**, *585* (14), 2339-2344.
31. Kurasin, M.; Väljamäe, P., Processivity of Cellobiohydrolases Is Limited by the Substrate. *J. Biol. Chem.* **2011**, *286* (1), 169-177.
32. Igarashi, K.; Uchihashi, T.; Koivula, A.; Wada, M.; Kimura, S.; Okamoto, T.; Penttilä, M.; Ando, T.; Samejima, M., Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science* **2011**, *333* (6047), 1279-1282.

33. Kurasin, M.; Kuusk, S.; Kuusk, P.; Sørlie, M.; Våljamäe, P., Slow Off-rates and Strong Product Binding Are Required for Processivity and Efficient Degradation of Recalcitrant Chitin by Family 18 Chitinases. *J. Biol. Chem.* **2015**, *290* (48), 29074-85.
34. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual Molecular Dynamics. *J. Mol. Graph. Model.* **1996**, *14* (1), 33-38.
35. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W. Z.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J., et al., Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
36. Goujon, M.; McWilliam, H.; Li, W. Z.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R., A New Bioinformatics Analysis Tools Framework at EMBL-EBI. *Nucleic Acids Res.* **2010**, *38*, W695-W699.
37. McWilliam, H.; Li, W. Z.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R., Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **2013**, *41* (W1), W597-W600.
38. Robert, X.; Gouet, P., Deciphering Key Features in Protein Structures with the New ENDscript Server. *Nucleic Acids Res.* **2014**, *42* (W1), W320-W324.