Supporting Information

Catalytic Cycle of Glycoside Hydrolase BglX from *Pseudomonas aeruginosa* and Its Implication in Biofilm Formation

Kiran V. Mahasenan[†§], María T. Batuecas[‡§], Stefania De Benedetti[†], Choon Kim[†], Neha Rana[†], Mijoon Lee[†], Dusan Hesek[†], Jed F. Fisher[†], Julia Sanz-Aparicio[‡], Juan A. Hermoso[‡*], and Shahriar Mobashery[†*]

[†]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556, USA
[‡]Department of Crystallography and Structural Biology, Institute of Physical Chemistry "Rocasolano", CSIC, 28006 Madrid, Spain.
[§]The first two authors contributed equally to this work
*To whom correspondence should be addressed. Email: SM, mobashery@nd.edu; JAH, xjuan@iqfr.csic.es

Table of contents

Supplementary materials and methods

Cloning of the *bglX* (PA1726) gene and purification of BglX wild-type and D286N variant

The gene for the wild-type BglX (amino acids 21–774; without the periplasmic signal peptide) was cloned from the *P. aeruginosa* PAO1 genomic DNA into pET28a(+) using restriction enzymes NdeI and HindIII. The plasmid pET28a(+) containing the *bglX* wild-type gene was used as the template for site-directed mutagenesis to produce the variant D286N. The Q5 Hot Start High Fidelity DNA polymerase (New England Biolabs) was used. The list of primers used for cloning and mutagenesis is shown in Table S1. The purity of the PCR reactions was determined using 1% agarose gel for 25 min at 100 V. The sequence of the wild type and the mutant gene were confirmed by DNA sequencing on both strands.

Escherichia coli BL21 star (DE3) cells (Invitrogen) were transformed with pET28a_*bglX* plasmid (wild type or mutant). A single colony was cultured overnight in LB media containing 30 μg/mL of kanamycin. This culture was transferred in 1 L of fresh LB media and was allowed to grow at 37 °C until an OD_{600} of 0.6. Protein expression was induced with 0.5 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside, IBI Scientific) at 16 C overnight. Cell pellet was resuspended in 40 mL of lysis buffer (50 mM Tris buffer, pH 8, 300 mM NaCl, 20 mM imidazole) and the proteins were released from the cells by sonification on ice (10 × 2 min cycles with 2 min rest using a Branson 450 Sonifier). After 45 min of centrifugation at 18,000 *g*, the supernatant was loaded onto 5 mL of Ni-NTA resin (Macherey-Nagel). The resin was washed with 50 mL of lysis buffer and the protein was eluted with a gradient of 20 to 500 mM imidazole in lysis buffer (200 mL). The fractions containing the recombinant protein, as verified by SDS-PAGE gel, were collected and concentrated with Amicon Ultra Centrifugal Filter, 10-kDa cut off. For both proteins, the purification yielded to 60 mg of protein from 1 L culture. The proteins were stable both at 4 °C and at –80 °C and survived freezing-thaw cycles.

Generation of a *bglX*::Tn strain harboring a BglX-producing plasmid (*bglX*::Tn/pBglX)

In order to produce BglX in the *bglX*::Tn background, the DNA fragment including the promoter, the full-length *bglX* gene and the transcription terminator was amplified from *P. aeruginosa* PAO1 genomic DNA with the primers listed in Table S1. The PCR product was inserted into an *E. coli – P. aeruginosa* shuttle vector, pCN61 at BamHI and XbaI sites.[1] The resulting plasmid, pBglX, was first introduced into *E. coli* DH5α by heat-shock transformation, followed by selection of transformants on an LB plate containing 100 μg/mL of ampicillin. The sequence in pBglX was confirmed by DNA sequencing. Then, pBglX was introduced into the *bglX*::Tn background strain by electroporation.[2] The transformants (*bglX*::Tn/pBglX) were selected on an LB plate containing 50 μg/mL of streptomycin as *P. aeruginosa* PAO1 and *bglX*::Tn are susceptible to streptomycin.

BglX activity assays

Activity assays for BglX was determined using the chromogenic substrate 4-nitrophenyl *β*-D-glucopyranoside (Glc-β-PNP, Sigma). The release of 4-nitrophenolate ion was spectrophotometrically monitored at 400 nm. The reaction (100 μL) was prepared in HEPES buffer (pH 7) with 5 μM enzyme and 1–20 mM Glc-β-PNP. The assays were performed in triplicate at 25 °C for 5 min. For inhibition assays, the reactions were performed under the same conditions as the activity assay, with fixed concentrations of Glc-β-PNP (10 mM) and compounds (40–100 mM). The kinetic parameters are given in Table S6.

BglX phosphorylase assay

Some GH3 enzymes exhibit phosphorylase activity.[3] . We explored if BglX could form glucose-*O*-1-phosphate in the presence of phosphate in the buffer. Wild-type BglX (5 μM) was incubated with sophorose (400 μM) in HEPES pH 7 buffer with or without 100 mM sodium phosphate at room temperature for 2 h and

18 h. The reaction mixtures were flash-frozen. We analyzed the results by LC/MS. Under the condition given in Table S2 the standard glucose-6-phosphate eluted at 4.2 min with an observed *m*/*z* value of 259.0203 (*m*/*z* $[M–H]^-$, calcd 259.0224 for $C_6H_{12}O_9P$). The presence of phosphate in the buffer did not enhance the BglX activity and glucose-1-phosphate was not detected as a product. Hence, BglX is unlikely to have phosphorylase activity.

Bioinformatics analysis of BglX

The *bgIX* gene (locus tag PA1726) is located at 1866 kb on the *P. aeruginosa* PAO1 chromosome adjacent to *mucR* and *pscL* genes. The *mucR* gene encodes MucR, a membrane-associated regulator of alginate biosynthesis, which is important in the biofilm development.[4-6] The *pscL* gene encodes the cytoplasmic type III export protein PscL.[7] The *pscL* gene contains an AmrZ-binding site motif signature. AmrZ is a transcription factor that positively regulates alginate synthesis.[8] Although both genes adjacent to *bglX* are involved in alginate biosynthesis, bioinformatics analysis from the DOOR database[9] indicates that they are not part of the same operon.

Crystallization

The complex of the wt BglX with 1-deoxynojirimycine was obtained by fast soaking the native protein crystals in a 40 mM compound solution during a few seconds. These crystals were cryoprotected as described for the native crystals.

The BglX variant D286N was concentrated to 12 mg/mL in 50 mM Tris, 300 mM NaCl pH 8 buffer. Crystals grew in a condition with 100 mM Bis Tris, 17% PEG 10000, 100 mM ammonium acetate pH 5.5 buffer using the sitting drop vapor-diffusion method at 291K. The crystals were soaked into a cryoprotectant solution (30% PEG 400) before flash cooling at 100 K.

The complex of BglX-D286N with laminaritriose (BglX-D286N:LMT) was obtained by soaking mutant crystals in a 30 mM compound solution during 10 minutes. These crystals were cryoprotected as described for the D286N variant.

The complex of BglX-D286N with two molecules of glucose (a ternary complex, BglX-D286N:BGC:BGC) was obtained by sequential soakings. First, the D286N crystals were soaked in a 15 mM laminaribiose solution for a few seconds and then in a 100% (w/v) glucose solution for a few seconds. These crystals were cryoprotected in a 70% paratone/30% parafine solution.

The complex of BglX-D286N with glucose (BglX-D286N:BGC) was obtained by soaking mutant crystals in a 25% glucose solution for a few seconds. These crystals were cryoprotected as described for the D286N variant.

The complexes of BglX-D286N with the lactose (BglX-D286N:Lactose) and cellobiose (BglX-D286N:Cellobiose) were obtained by soaking the D286N crystals in a 35 mM of each compound solutions separately for 45 minutes. These crystals were cryoprotected as described for the D286N variant.

The complex of BglX-D286N with xylotriose (BglX-D286N:Xylotriose) was obtained by soaking the D286N crystals in a 30 mM compound solution for 10 minutes. These crystals were cryoprotected as described for the D286N variant.

Non-denaturing mass spectrometry

Non-denaturing mass spectrometry of wild-type BglX was performed as previously described with minor modifications. [10] The BglX (20 μM) was buffer exchanged into 1 M ammonium acetate using a Zeba Desalting Column, followed by an Amicon Ultra 0.5 mL Centrifugal Filter (10-kDa molecular weight cut-off). ESI MS

spectra of BglX were acquired in the mass range 500–10000 Da using a Bruker MicrOTOF-QII operating in the positive-ion mode with the following parameters: end plate offset –500 V, capillary voltage 3.4 kV, nebulizer gas pressure 0.8 bar, dry gas flow rate 3 L/min, dry gas temperature 180 °C, funnel 1 RF 400 V, funnel 2 RF 500 V, hexapole RF 600 V, quadrupole ion energy 5 eV, collision energy 8 eV, collision cell RF 3.4 kV, ion transfer time 100 μs, and pre-pulse ion storage 25 μs. Solutions were infused at a flow rate of 10 μL/min. Mass spectra were collected for 10 minutes. The mass spectrum of BglX under this non-denaturing mass spectrometry condition is given in Figure S7.

Analysis of the solved X-ray structures in the GH3 family

All solved X-ray structures of GH3 family members were obtained from PDB databank (111 X-ray structures). Structures with more than 95% sequence identity were not considered to remove the redundancy, which resulted in 38 diverse enzymes of the family (Table S6). The structures were superimposed to the BglX:LMT X-ray structure with MAESTRO program (v. 2018, Schrödinger, LLC; New York, NY). Multiple Sequence Viewer module of MAESTRO aligned amino-acid sequences based on the structural superimposition and extracted the structure-based sequence alignment of the 39 structures. The multiple-sequence alignment was imported to MEGA-X software (Build# 10180925) for the phylogenetic-tree construction.[11] Neighbor-Joining method deduced the evolutionary history.[12] The Poisson-correction method computed the evolutionary distances.[13] The phylogenetic tree (Figure S8) was plotted with ggtree package (v. 1.2.17) in R program.

Modeling of protein-ligand complex

The X-ray complex of the laminaritriose (LMT) with BglX-D286N mutant was computationally mutated to D286, and prepared using the Protein Preparation Wizard module of the MAESTRO interface (v. 2017). In this stage appropriate bond orders were assigned and hydrogen atoms were added. The protonation state of the histidine residues was assigned based on calculation with the PROPKA program. The orientations of the crystal water molecules were optimized based on hydrogen-bond network analysis. The prepared complex was energy-minimized (OPLS2005 forcefield) and used as the enzyme-substrate Michaelis complex for simulation. For modeling the covalent glycosyl species (Figure 4C) we structurally aligned the previously reported covalent glycosyl enzyme model of NagZ [14] to the BglX-D286N:LMT structure. This alignment brought the glycosylated D244 of NagZ into the same space as D286N of BglX (Figure S5 and S12). Based on the alignment, the BglX complex with C1 atom of glucose covalently bound with D286 at the –1 site was manually modelled and energy-minimized with OPLS2005 forcefield.

QM/MM (Quantum Mechanics/Molecular Mechanics) simulation

For QM/MM analysis, the snapshots were extracted from MD trajectories and were subjected to steered MD (sMD) based on pre-defined reaction coordinate (RC). AMBER16 package with QM/MM support was used to perform MD with semi-empirical method, PM6-D, for treating QM region and Amber forcefield ff14SB for treating MM region. [15] Side chains of E517, D286, and substrate/product atoms were included in the QM region. The rest of the solvated system was treated with MM. Following the protocol for modeling net retaining GH mechanisms, the reaction was split into two halves, glycosylation and deglycosylation [16]. For the deglycosylation step, the hydrolytic water molecule was also included in the QM region. Both halves were modeled individually using Linear Combination of Distances (LCOD) approach in AMBER16. For glycosylation, the reaction coordinates were chosen as:

$$\boldsymbol{RC_1 = d_{O3-H_{E2(E517)}} - d_{H_{E2(E517)}-O_{E2(E517)}}} \text{ and } \boldsymbol{RC_2 = d_{O_{D2(D286)}-C1} - d_{C1-O3}}$$

For deglycosylation, the reaction coordinates were chosen as:

$$\boldsymbol{RC_1 = d_{O_{\varepsilon 2(E517)}-HW} - d_{HW-OW}} \text{ and } \boldsymbol{RC_2 = d_{OW-C1} - d_{C1-O_{D2(D286)}}}$$

The sampling was performed along 21/19 (d_1/d_2) points and 20/20 (d_1/d_2) points in steering paths for glycosylation and deglycosylation steps, respectively, with 100 kcal/mol·Å^2 harmonic restraints. The SHAKE was turned off inside the QM region and atoms beyond 8 Å of QM region were heavily restrained. The MD with semi-empirical QM treatment was run for 500 ps. The Potential of Mean Force (PMF) profile was computed for each step.

Analysis of the solved homodimeric hydrolase enzymes

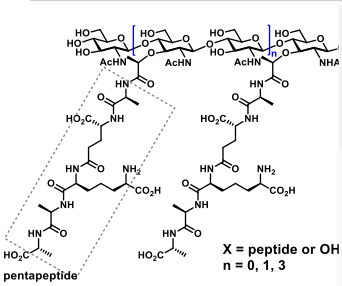
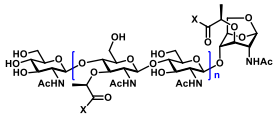
All the solved structures in the hydrolase enzyme family (28,973 PDB entries; Table S8) deposited at the RCSB PDB databank (www.rcsb.org) was searched for homodimeric arrangement (EnzymeClassificationTree Search for Hydrolases and Stoichiometry in biological assembly of A2) which resulted in 5,189 PDB entries. Presence of a ligand at the interface of the dimer was searched with Protein interfaces, surfaces and assemblies program (PISA) module of CCP4 package [17]. From the PISA calculation, PDB entries with at least one ligand in contact with both monomer chains were enumerated. After removing redundant entries of the same enzymes (95% sequence identity cut-off), 110 enzymes were identified to have a binding site at the interface among the total 7,956 representative hydrolase enzyme structures.

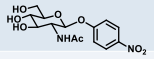
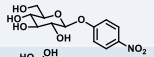
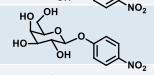
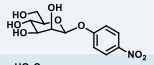
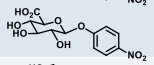
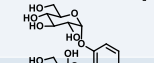
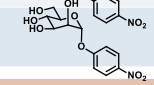
Analysis of the taxomony of BglX homologs

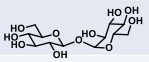
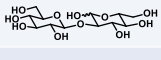
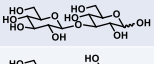
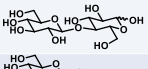
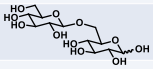
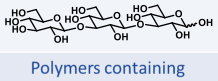
Sequences of BglX and the other four known structural homologs (Table S6) were used to perform a BLAST search to the NCBI RefSeq database. A total of 20,000 hits were retrieved for each with at least 30% identity to the query sequence. The results were combined to generate a collection of 100,000 sequences. Identical sequences and those with more than 95% identity were removed from the pool to obtain representative enzymes using USEARCH sequence clustering module, 'cluster_fast', and resulted in a final collection of 8,815 homologs. A similar BLAST search was conducted with sequences of the L9 loop of each of the five enzymes, resulting in 10,233 homologus enzymes. Between the two sets, 6,678 enzymes were common, which represented BglX homologs with an L9 loop (the loop from the second monomer that participates in the active-site assembly and catalytic reaction). Taxonomy of the homologs were available from the NCBI taxonomy database. They were retrieved using the taxonomizR package (https://github.com/sherrillmix/taxonomizR) and plotted with the sunburstR package (https://github.com/timelyportfolio/sunburstR) of R program (Figure S13).

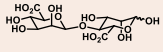
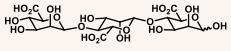
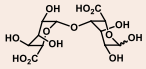
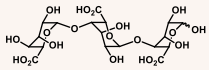
Figures S1–S13

A

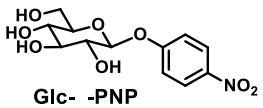
Synthetic peptidoglycans		Peptide	OH	BglX substrate
	MurNAc	n = 0	n = 0	X
		n = 1	n = 1	X
		n = 3	n = 3	X
	anhMurNAc	n = 0	n = 0	X
		n = 1	n = 1	X
		–	n = 3	X
		–	n = 2-9	X
	GlcNAc	–	chitin	X

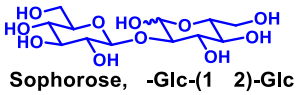
Monosaccharide *O-p*-nitrophenyl ether		BglX substrate
GlcNAc-β-PNP		X
Glc-β-PNP		O
Gal-β-PNP		O
Man-β-PNP		X
GlcA-β-PNP		X
Glc-α-PNP		X
Man-α-PNP		X

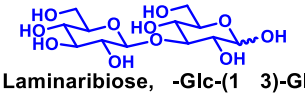
Disaccharide Glc-Glc, β-(1→x)		BglX substrate
α,β-(1→1) α,β-Trehalose		X
β-(1→2) Sophorose		O
β-(1→3) Laminaribiose		O
β-(1→4) Cellobiose		X
β-(1→6) Gentiobiose		X
β-(1→3)-Laminaritriose		O
Laminarin	Polymers containing β-(1→3)-glycosidic link with β-(1→6) branches	O

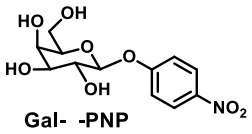
Alginate		BglX substrate
D-mannuronic acid β-(1→4) disaccharide		X
D-mannuronic acid β-(1→4) trisaccharide		X
L-guluronic acid α-(1→4) disaccharide		X
L-guluronic acid α-(1→4) trisaccharide		X
Alginate	Polymer of β-(1→4) D-mannuronate and α-(1→4) L-guluronate	X

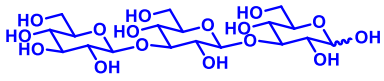
B



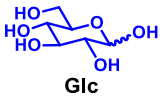


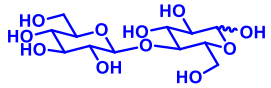


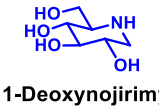


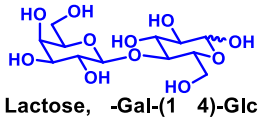


C









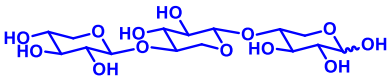


Figure S1. Chemical structures of (A) compounds screened for BglX enzymatic activity and of (B) compounds turned over by BglX. The symbol 'X' indicates lack of turnover and 'O' the presence of turnover by BglX. (C) Compounds used to obtain X-ray structures of BglX complexes. Compounds colored in blue in panel B are substrates that were also used to obtain X-ray structures of BglX complexes.

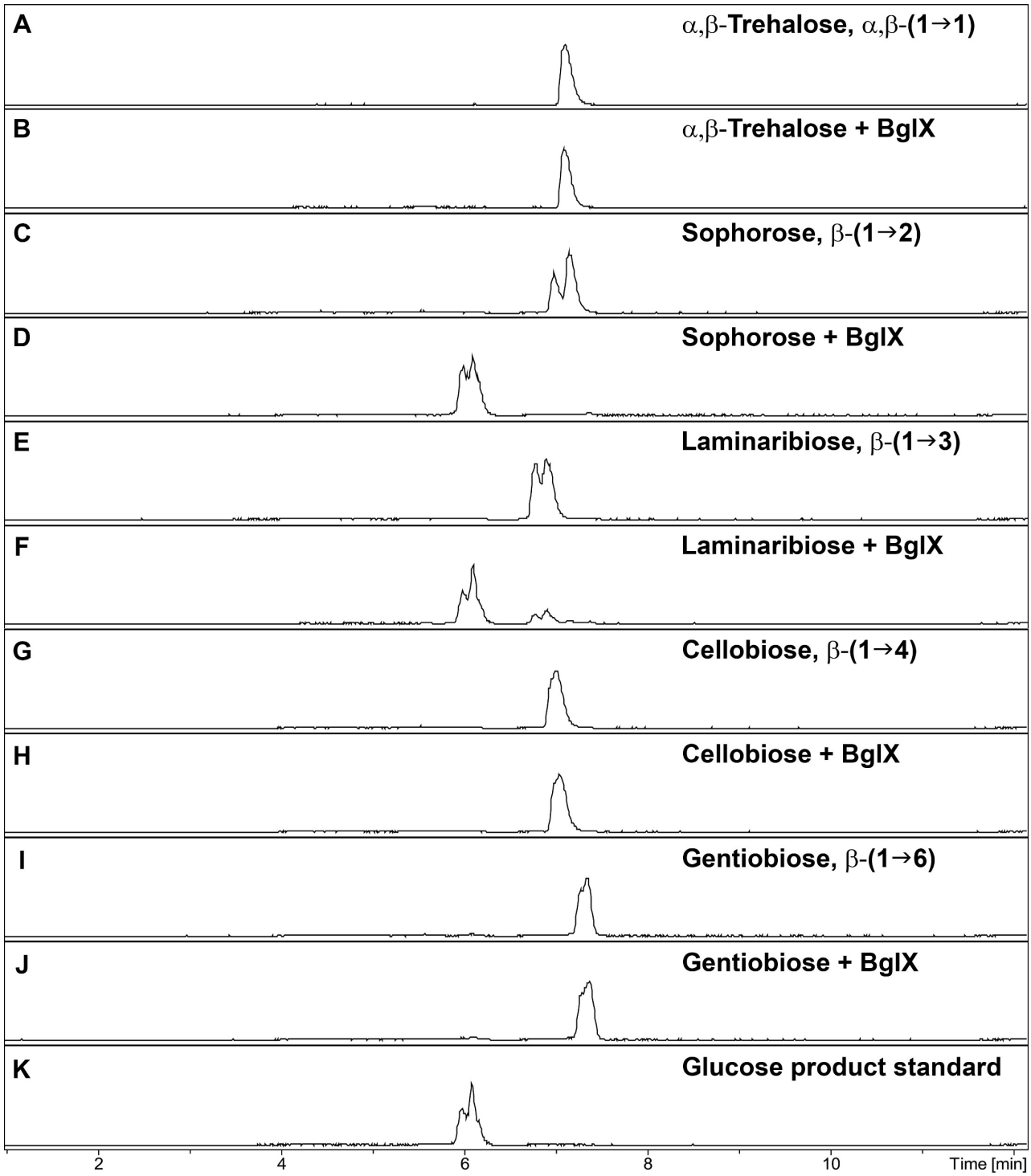


Figure S2. LC/MS traces of BglX reactions with five Glc-Glc isomers with various β-(1→x)-glycosidic linkage. Chromatograms before (A, C, E, G, and I) and after the addition of BglX (B, D, F, H, and J) and standard of reaction product, glucose (K) are given.

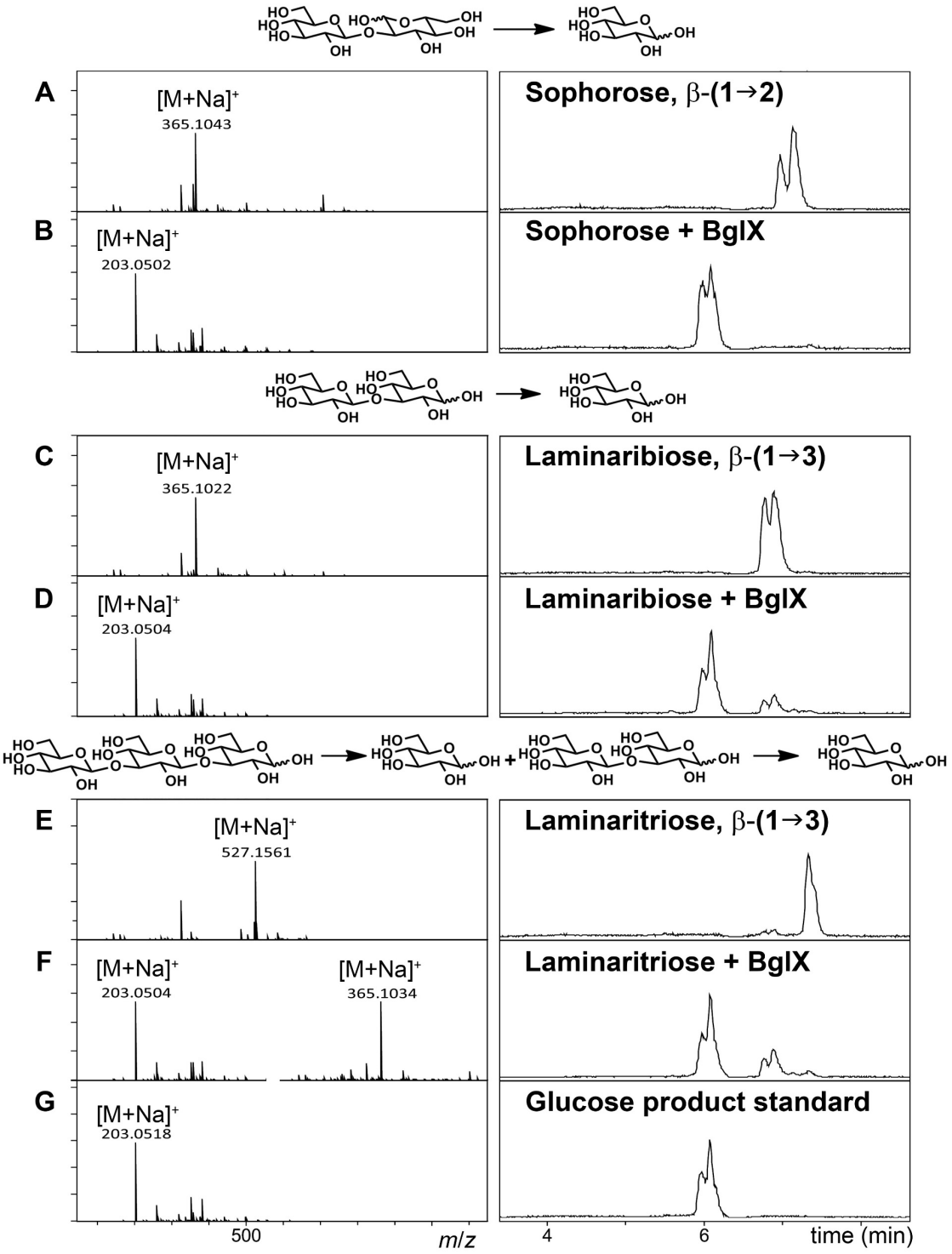


Figure S3. Turnover of sophorose, laminaribiose, and laminaritriose by BglX. MS spectra of substrate and reaction products are given in the left panels.

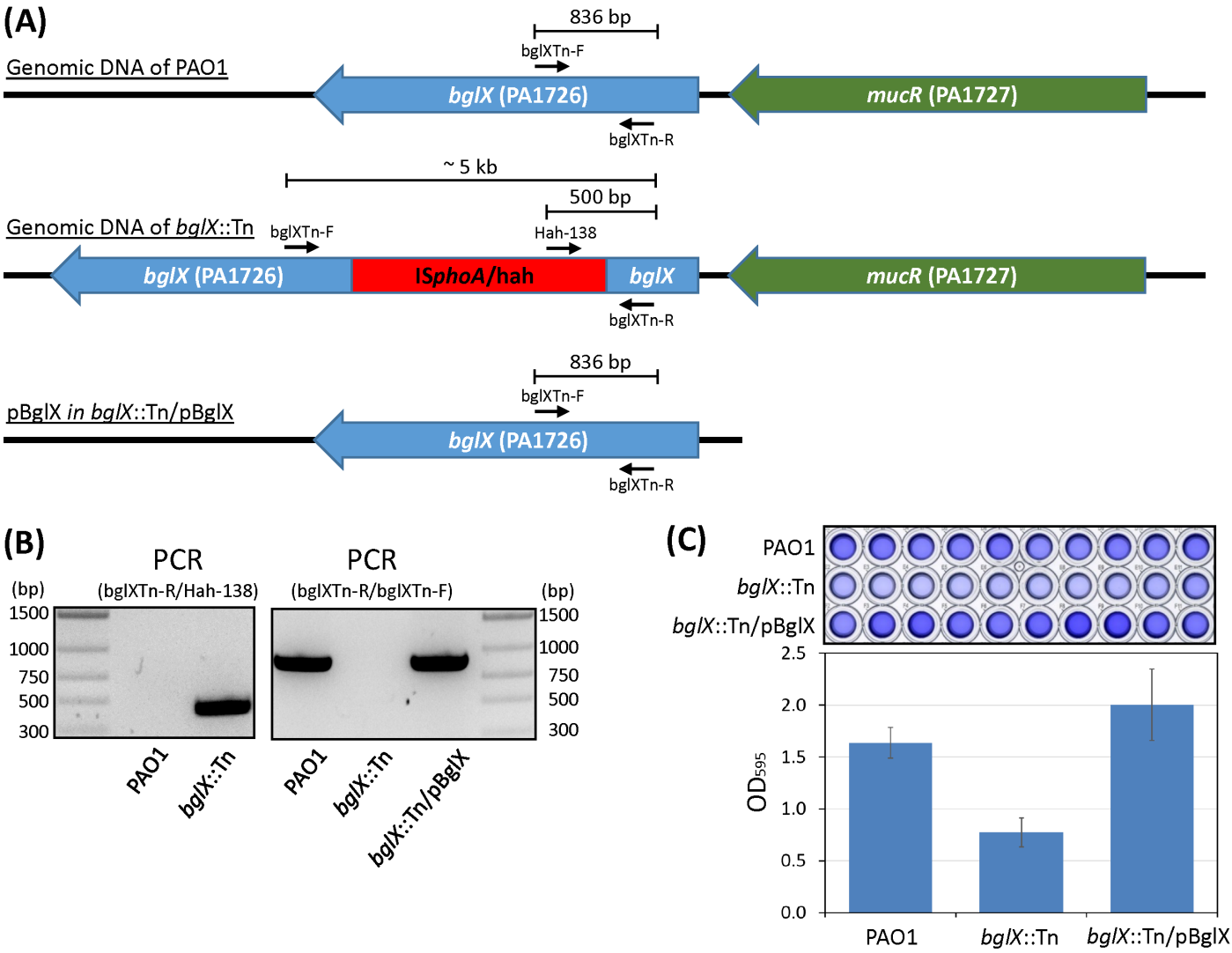


Figure S4. Confirmation of transposon insertion in *bglX* and biofilm formation of *P. aeruginosa* PAO1 WT, *bglX*::Tn and *bglX*::Tn/pBglX. (**A**) Structural diagram of the *bglX* region in genomic DNA in PAO1 and *bglX*::Tn and the *bglX* gene of pBglX plamid in *bglX*::Tn/pBglX. (**B**) PCR of genomic DNAs from PAO1 and *bglX*::Tn and the plasmid from *bglX*::Tn/pBglX with two different sets of primers. PCR with primers bglX-R and Hah-138 produced the 500-bp product only in *bglX*::Tn since Hah-138 is a Tn-specific primer. PCR with primers bglX-R and bglX-F amplified the 836-bp products in wild-type PAO1 and *bglX*::Tn/pBglX, confirming that *bglX*::Tn/pBglX carries the plasmid. (**C**) Biofilm formation in *P. aeruginosa* PAO1, *bglX*::Tn and *bglX*::Tn/pBglX was measured with 0.1% crystal violet. Biofilm formation was decreased in the transposon mutant *bglX*::Tn, while it was restored as much as the wild type when the *bglX* gene was inserted (*bglX*::Tn/pBglX).

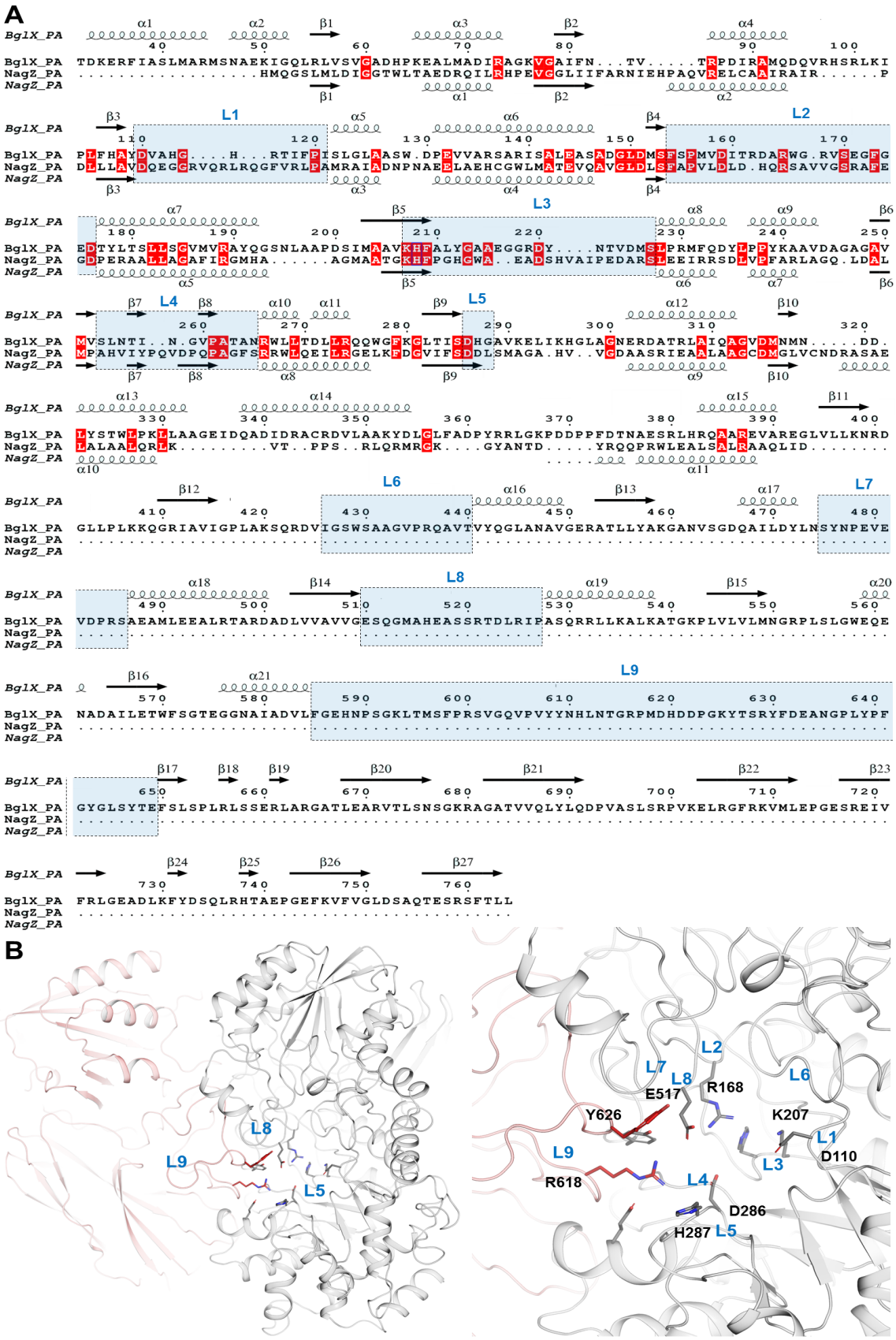


Figure S5. Constitution of the BglX active site. (A) The sequence and secondary-structural elements of BglX in comparison to NagZ. The loops that form the active site are highlighted in boxes with blue shade. The top and bottom rows depict the type of secondary structures. The residues are numbered in the second and third rows. (B) Residues in the active site that participate in catalysis and the corresponding loops (based on the apo X-ray structure of BglX), as depicted in panel A and Table S5, are shown. Loop L9 reaches into the active site of the second monomer.

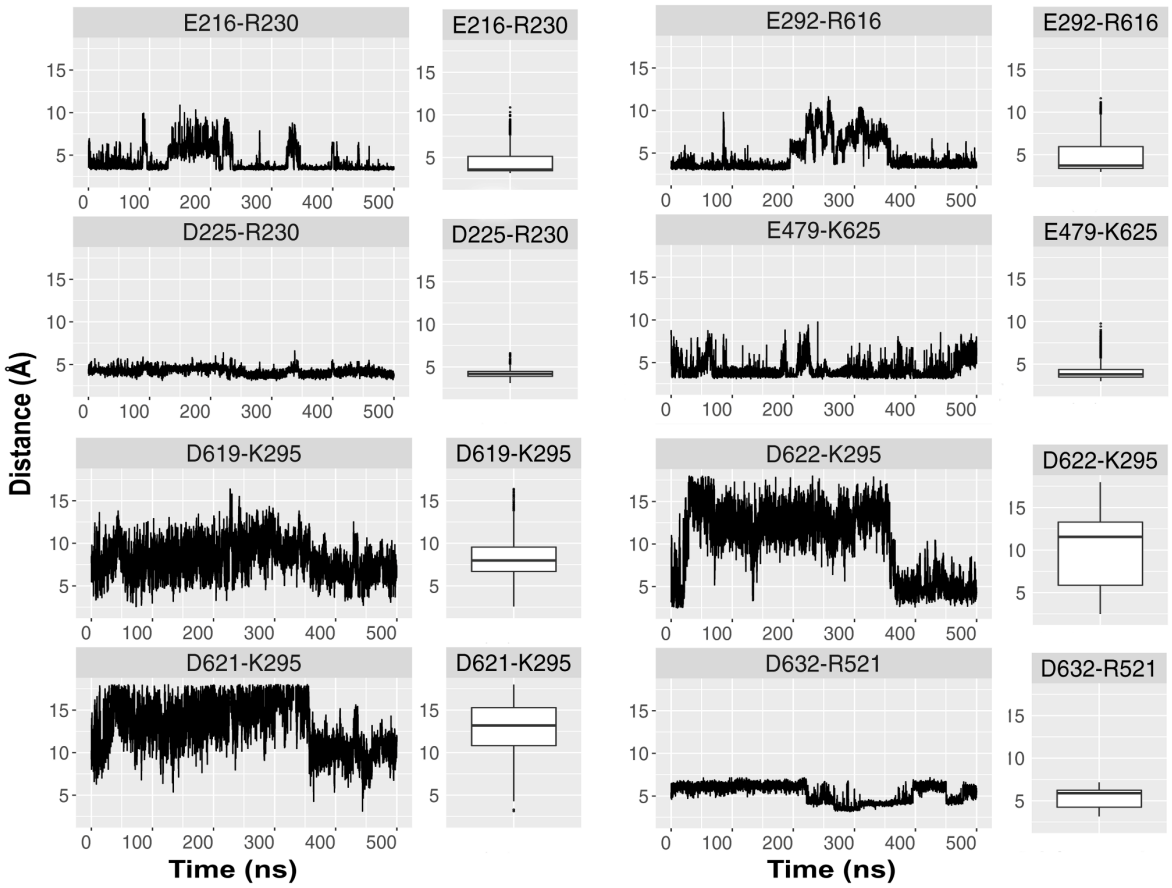


Figure S6. Distance fluctuation of salt-bridges at the dimer interface of the apo BglX complex during 500-ns MD simulation, displayed by line-plot (distance against time; on the left) and box-plot (to the right). Data points were calculated from 10,000 snapshots of the 500-ns trajectory. The distance scale was cut-off at 18 Å. Five of the eight salt-bridges were maintained over the simulation, indicating contribution to the dimer stability, of which E292·R616 is of particular interest as it is closely involved with the catalytic cycle (Figure 4).

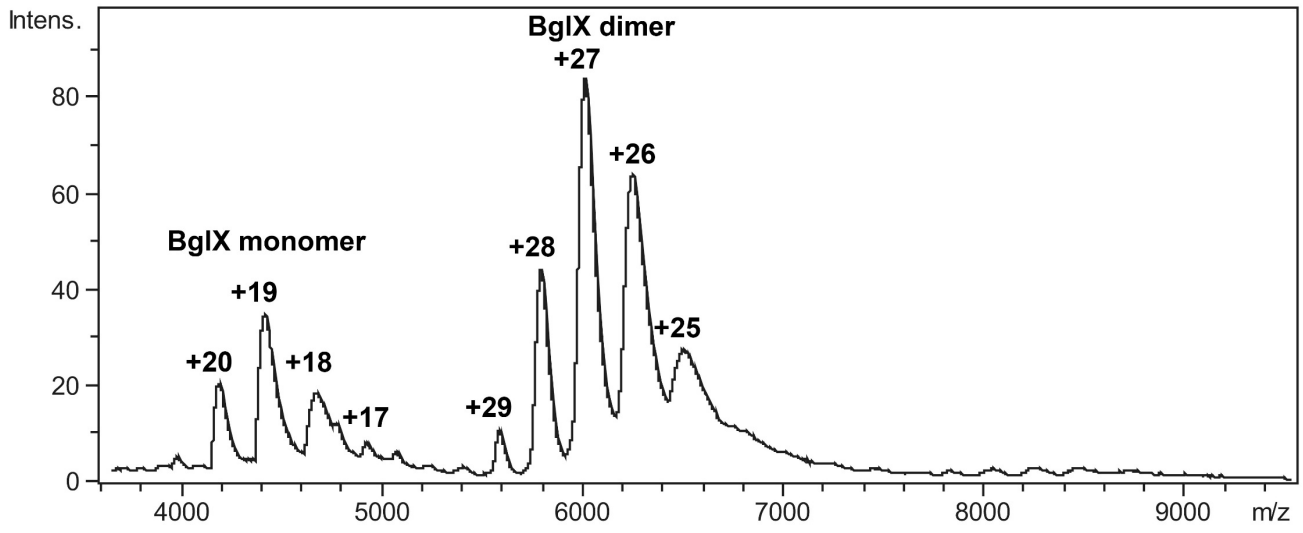


Figure S7. The Mass spectrum of BglX under non-denaturing mass spectrometry condition (average mass of dimer is 162,473 ± 4,000 Da).

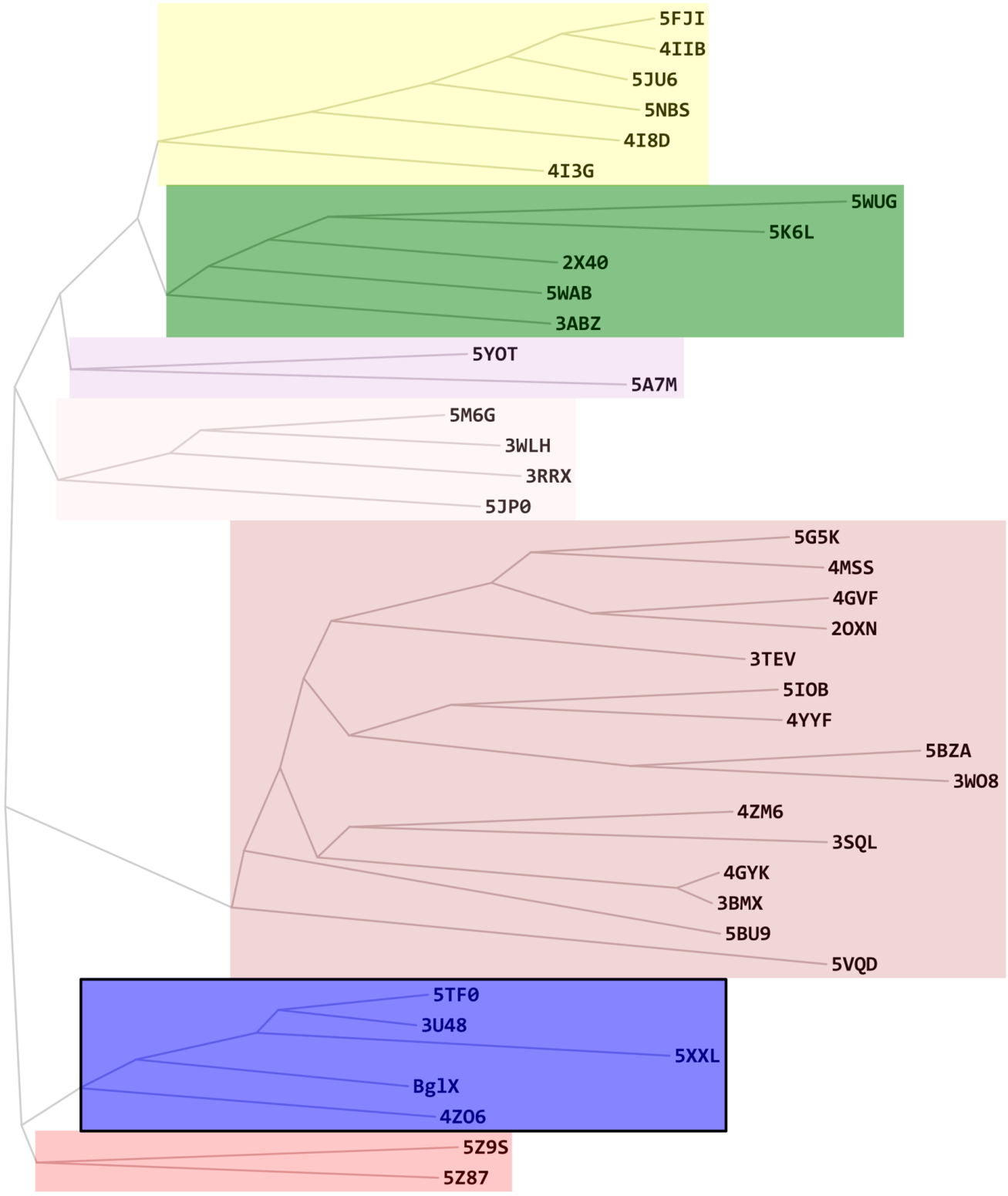
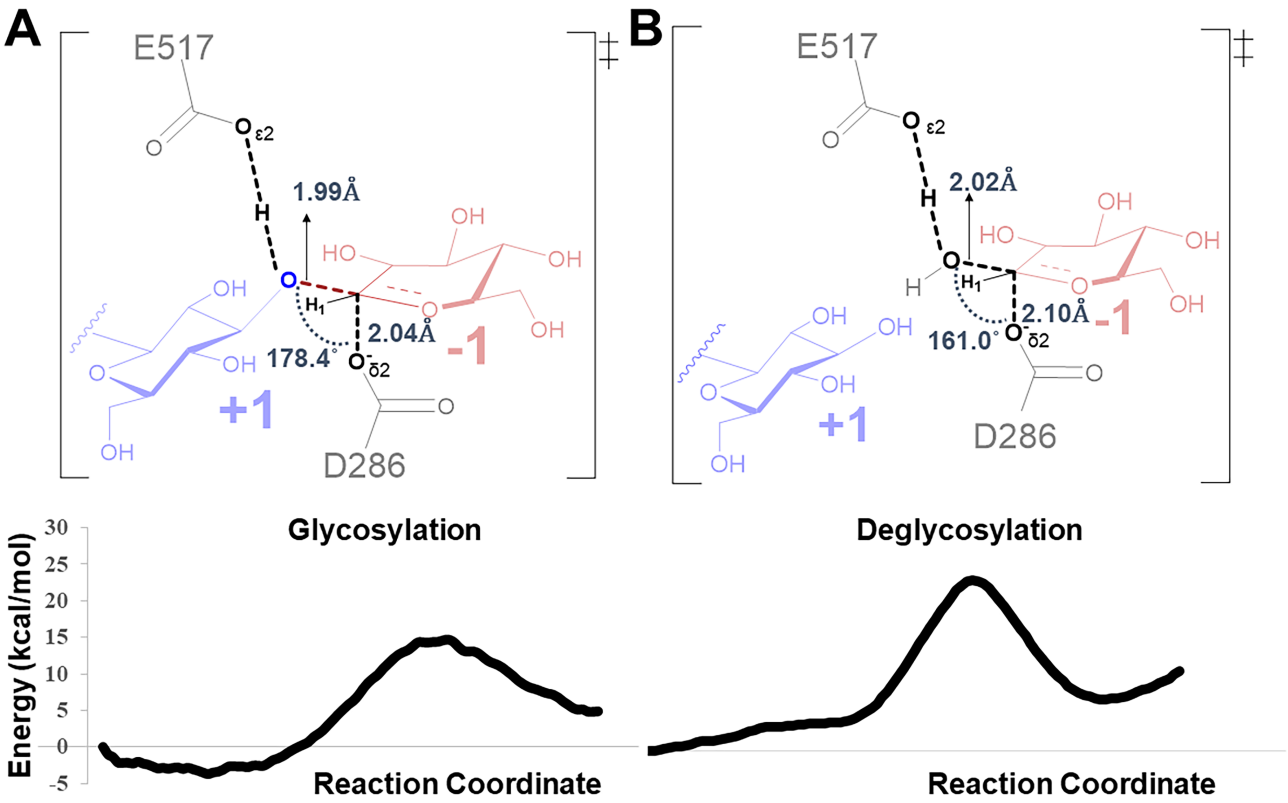


Figure S8. Phylogenetic tree showing the distribution of GH3 family enzymes with solved structures in the Protein Data Bank labeled with PDB codes. BglX homologs with homodimeric active site is clustered in the box in violet (second from bottom, with black border) while NagZ homologs (β-hexosaminidases) are in the pink box (third from bottom). The code for *P. aeruginosa* NagZ is 5G5K. The tree was generated with MEGA software using structure-based sequence alignment performed with MAESTRO program. The figure was generated with ggtree R package. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Enzymes and organism corresponding to the PDB codes are listed in the Table S5.



C

Characteristics*	Glycosylation			Deglycosylation		
	Michaelis Complex	First TS	Covalent-enzyme species	Covalent-enzyme species**	Second TS	Product
C1···O3 (or C1···OW)	1.49 Å	1.99 Å	2.44 Å	-	2.02 Å	1.40 Å
C1···$O_{δ2(D286)}$	3.09 Å	2.04 Å	1.56 Å	1.54 Å	2.10 Å	3.43 Å
O3···$H_{ε2(E517)}$ (or OW···HW)	1.85 Å	1.10 Å	1.01 Å	0.95 Å	1.11 Å	1.98 Å
$O_{ε2(E517)}$···$H_{ε2(E517)}$ (or $O_{ε2(E517)}$···HW)	0.96 Å	1.40 Å	1.75 Å	-	1.40 Å	0.97 Å
C1···O5	1.46 Å	1.29 Å	1.40 Å	1.41 Å	1.25 Å	1.41 Å
O3···C1···$O_{δ2(D286)}$ (or OW···C1···$O_{δ2(D286)}$)	125.9°	178.4°	147.3°	-	161.0°	168.3°
O3···C1···H1 (or OW···C1···H1)	112.6°	80.8°	53.5°	-	91.1°	103.6°
$O_{δ2(D286)}$···C1···H1	30.3°	80.6°	95.6°	113.3°	80.7°	69.1°
Sugar Puckering	$^4C_1$	$^4H_3$	^4E	^4E	$^4H_3$	E_3

* Distinct characteristics of deglycosylation in parenthesis.

** Covalent enzyme species after water molecule has penetrated the active site.

Figure S9. The transition states of the glycosylation (**A**) and deglycosylation (**B**) steps of BglX (top panels) and their corresponding reaction profiles (bottom panels), as calculated by QM/MM simulation (PM6-D). Energy barriers calculated for first TS and second TS are 14.6 kcal/mol and 23.0 kcal/mol, respectively. The characteristics of enzyme species are given in panel **C**.

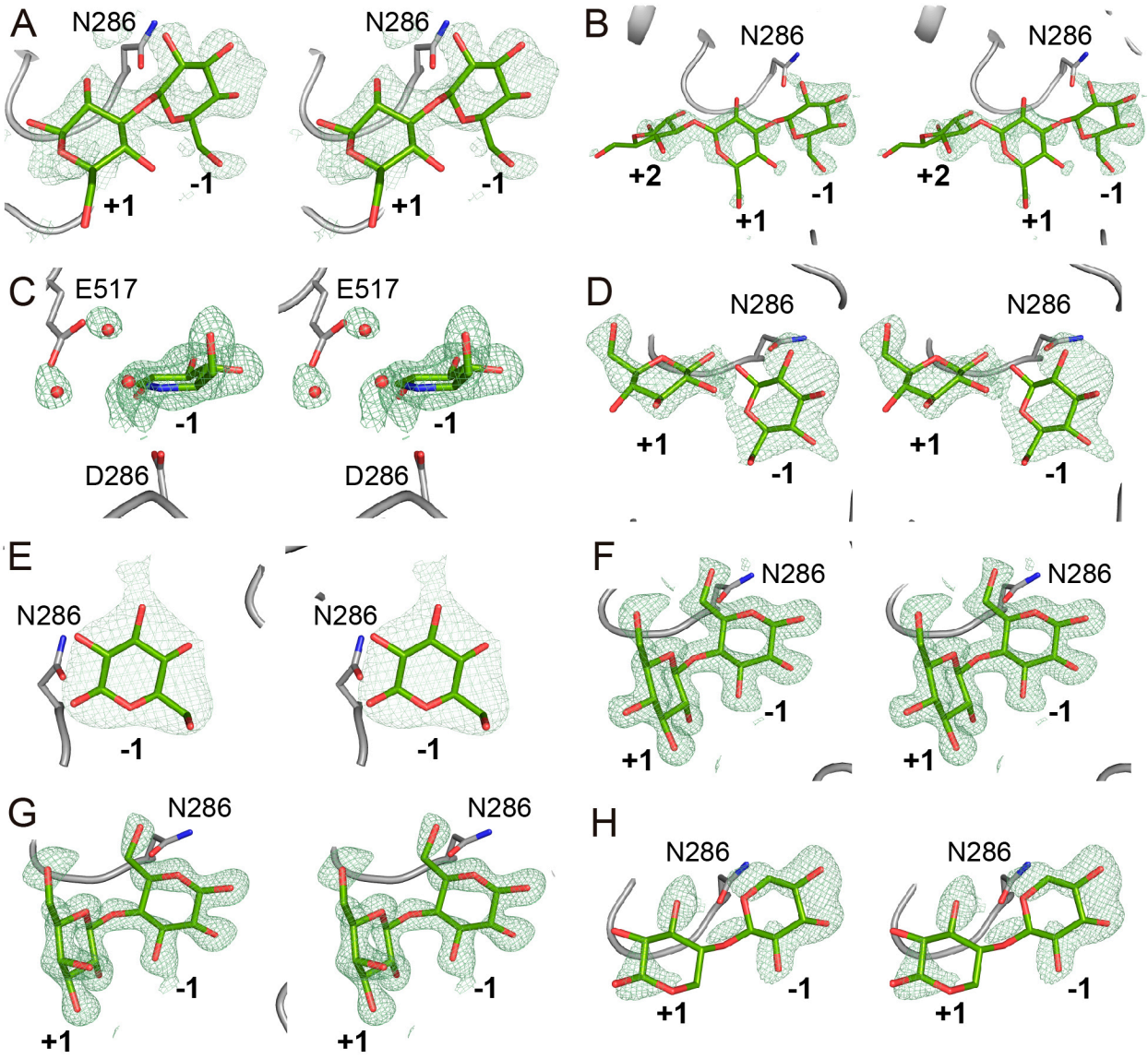


Figure S10. Stereo view showing the Fo-Fc electron-density maps for the BglX complexes. Initial electron-density maps Fo-Fc map contoured at 2σ for (**A**) the Chain B of the complex BglX-D286N:Laminaritriose (only two saccharide units are visible in subsites -1 and +1), (**B**) the Chain A of the complex BglX-D286N:Laminaritriose (laminaritriose occupies subsites -1, +1 and +2), (**C**) the complex BglX:1-deoxynojirimycin (compound is located in subsite -1, E517 and key water molecules are also showed), (**D**) the ternary complex BglX:Glucose:Glucose. (glucose molecules are in subsites -1 and +1), (**E**) the BglX-D286N:Glucose complex (glucose is in the -1 position), (**F**) the complex BglX:Cellobiose, (**G**) the complex BglX-D286N:Lactose and (**H**) the complex BglX-D286N:Xylotriose, where only two rings of the xylotriose are visible. Ligands in panels F, G and H occupy subsite -1, but they are not substrates of BglX.

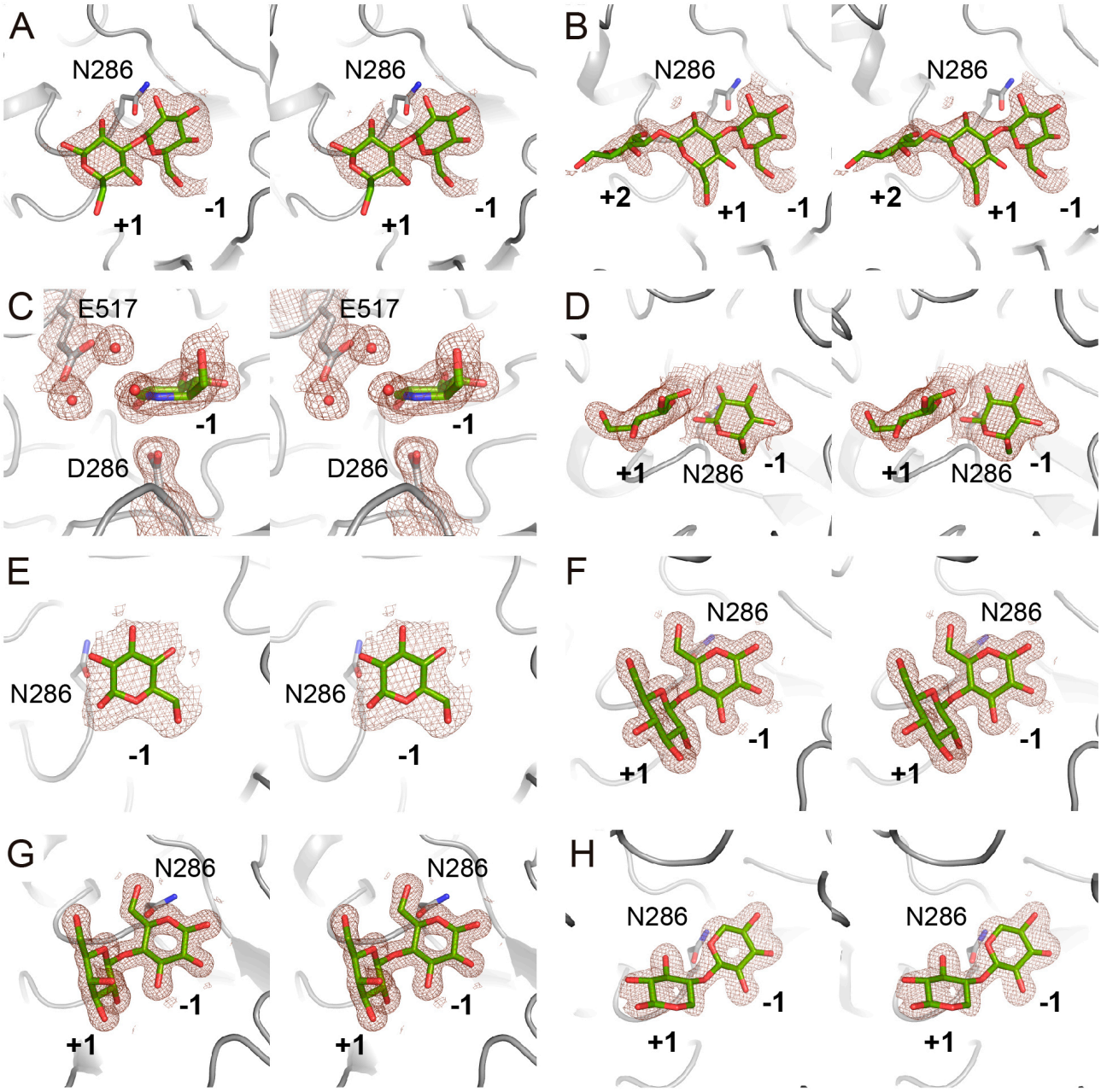


Figure S11. Stereo view showing the 2Fo-Fc electron-density maps for the BglX complexes. (**A**) Electron-density map contoured at 0.8σ and carve at 2.0 for the Chain B of the complex BglX-D286N:Laminaritriose. Only two saccharide units are visible in subsites -1 and +1. (**B**) Electron-density map contoured at 0.8σ and carve at 2.0 for the Chain A of the complex BglX-D286N:Laminaritriose. The laminaritriose occupies subsites -1, +1 and +2. (**C**) Electron-density map contoured at 1σ and carve at 1.8 for the complex BglX:1-deoxynojirimycin. The compound is located in subsite -1. E517 and key water molecules are also showed. (**D**) Electron-density map contoured at 1σ and carve at 2.0 for the ternary complex BglX:Glucose:Glucose. The glucose molecules are in subsites -1 and +1. (**E**) Electron-density map contoured at 0.8σ and carve at 2.0 for the BglX-D286N:Glucose complex. Glucose is in the -1 position. (**F**) Electron-density map contoured at 1σ and carve at 1.8 for the complex BglX:Cellobiose. (**G**) Electron-density map contoured at 1σ and carve at 1.8 for the complex BglX-D286N:Lactose. (**H**) Electron-density map contoured at 0.8σ and carve at 1.8 for the complex BglX-D286N:Xylotriose, where only two rings of the xylotriose are visible. Ligands in panels F, G and H occupy subsite -1, but they are not substrates of BglX.

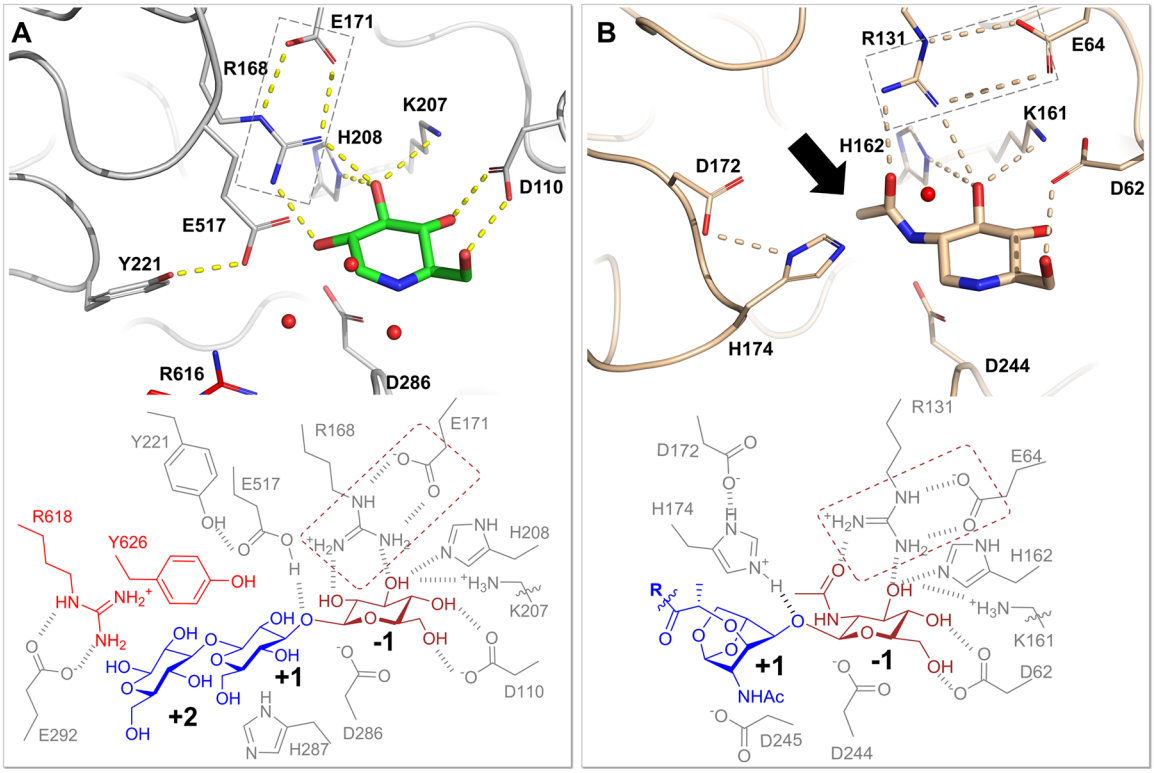


Figure S12. Comparison of (A) BglX active site (X-ray structure in complex with 1-deoxynojirimycin; present work) with that of (B) NagZ (in complex with 2-acetamido-1,2-dideoxy nojirimycin; PDB code: 5G5K). The nucleophilic amino-acid residue (D286 and D244, respectively) and the acid/base residue (E517 and H174, respectively) occupy the same three-dimensional space. Crystallographic water molecules (represented as red spheres) close to the C1 carbon and the acid/base residue appropriately occupy for donating the proton, which suggest their role in the catalytic mechanism of GH3 enzyme family. The saccharides at the –1 site for both enzymes show similar hydrogen-bond patterns. The alternate position of R131·E64 salt-bridge in the NagZ active site (highlighted in the box in broken lines) allows for occupancy of *N*-acetamido group (indicated by black arrow), while the R168·E171 salt-bridge hinders binding in BglX. A simplified 2D summary of the substrate-bound complex is given at the bottom of each panel.

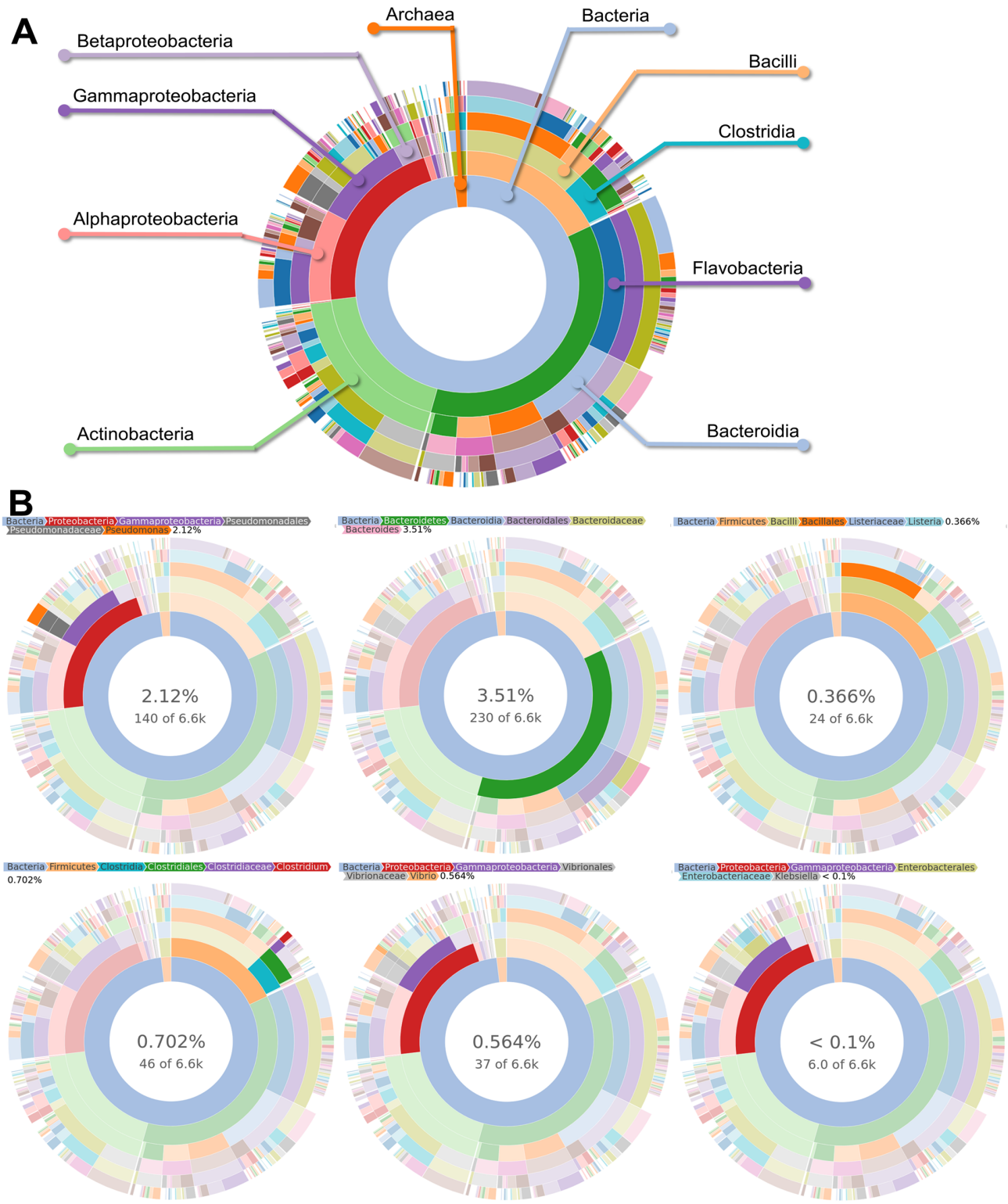


Figure S13. Taxonomy of organisms having BglX homologs from NCBI RefSeq protein sequence database. Taxonomy was retrieved from NCBI Taxonomy database and is represented as a sunburst plot. The innermost ring represents the superkingdom in taxonomy, in which more than 98% are bacteria and the rest are archea. The ring immediately outer to the superkingdom represents phylum, and is followed by the ring representing class. The top panel (A) is labeled for superkingdom and class. Rings to the outer are class, order, family, and genera, in that order. Bottom panel B shows selected six genera, of which X-ray structures are solved for the protein from *Pseudomonas* (this work), *Bacterioids*, and *Listeria* (top row). Other notable genera shown are *Clostridium*, *Vibrio,* and *Klebsiella* (bottom row). Taxomony was generated with taxonomizR package. The figure was plotted with sunburstR package, both using R program.

Tables S1–S7

Table S1. List of primers used for cloning and mutagenesis.

	Construct	Primer	Sequence*
1	WT	fwd rev	5′-GCGCGC**CATATG**GCTGGCCAGGGCGAATCC-3′ 5′-GCGCGC**AAGCTT**TCAGAGCAGGGTGAAGCTG-3′
2	D286N	fwd rev	5′-ACCATCAGTAACCACGGCGCG-3′ 5′-CGCGCCGTGGTTACTGATGGT-3′
3	pBglX	fwd rev	5′-AATTA**TCTAGA**ACGGGACCTGTCAGAAGCGGC-3′ 5′-AATTA**GGATCC**GAAGGGGACGCCGGTTGAGGGAACC-3′
4	Confirmation of transposon insersion	Hah-138 bglX-F bglX-R	5′-CGGGTGCAGTAATATCGCCCT-3′ 5′-CGACCAACTGCCGATGAC-3′ 5′-TGAGCTTCTCGCCTATGGTC-3′

*Restriction enzyme sites are in bold. Mutated codons are underlined.

Table S2. LC-MS conditions used for BglX substrate screening.

Compound	Column	LC gradient	Ionization
Synthetic peptidoglycans	Dionex Acclaim™ PolarAdvantage II C18	Variations of 45-min gradient[a] 98% A/2% B for 5 min 83% A/17% B for 40 min	Positive
Monosaccharide PNPs/disaccharide /trisaccharide /chitin	Acquity UPLC® BEH Amide	14-min gradient[a] 1% A/99% B at 0 min 70% A/30% B for 10 min 1% A/99% B for 0.1 min and keep for 3.9 min	Positive
Laminarin	Acquity UPLC® HSS T3 column	45-min gradient[a] 100% A/0% B for 5 min 87% A/13% B for 40 min	Positive
Alginate/Disaccharide/trisaccharide of Alginate component	Acquity UPLC® BEH Amide	20-min gradient[a] 20% A/80% B for 3 min 80% A/20% B for 10 min 1% A/99% B for 2 min and keep for 5 min	Negative
Phosphate	Acquity UPLC® BEH Amide	10-min gradient[b] 1% A/99% B at 0 min 70% A/30% B for 7 min 1% A/99% B for 0.1 min and keep for 2.9 min	Negative

[a]Mobile phase with acidic modifier: A = 0.1% formic acid in water, B = 0.1% formic acid in CAN. [b]Mobile phase with basic modifier: A = 1 mM ammonium bicarbonate in 95:5 = water:ACN, B = 1 mM ammonium bicarbonate in 5:95 = water:ACN

Table S3: Crystallographic data*

	BglX	BglX-D286N:LMT**	BglX:NOJ**	BglX-D286N:BGC:BGC**	BglX-D286N:BGC**	BglX-D286N:Cellobiose	BglX-D286N:Lactose	BglX-D286N:Xylotriose
Diffraction Data Statistics								
Wavelength (Å)	0.979270	0.979185	0.979310	0.979150	0.979160	0.979257	0.979257	0.979257
Space group	P 1	P 1	P 2_1 2_1 2_1	P 1	P 1	P 1	P 1	P 1
a, b, c (Å)	71.08,73.70,81.53	72.05,74.43,82.12	77.94,86.88,244.18	73.50,74.91,82.10	70.71,73.62,81.51	71.00,73.58,81.36	70.98,73.59,81.12	70.96,73.67,81.44
α, β, γ	65.58,73.91,69.70	65.63,73.94,69.46	90,90,90	65.64,74.18,68.90	65.38,73.41,69.89	65.60,73.67,69.70	65.55,73.42,69.68	65.55,73.46,69.66
Resolution range (Å)	45.57-1.80 (1.83-1.80)	46.08-2.15 (2.19-2.15)	86.88-2.00 (2.03-2.00)	46.72-2.40 (2.48-2.40)	45.16-2.85 (3.00-2.85)	45.37-1.65 (1.68-1.65)	45.17-1.60 (1.63-1.60)	45.27-1.80 (1.83-1.80)
Unique reflections	124434 (6074)	75583 (4474)	109537 (5282)	48392 (4979)	31258 (4491)	160758 (7738)	175334 (8534)	123955 (6063)
Completeness (%)	96.2 (94.7)	96.6 (95.6)	97.2 (96.2)	84.1 (93.7)	96.8 (95.4)	96.3 (94.2)	96.1 (94.5)	96.3 (94.8)
Multiplicity	2.4 (2.4)	3.5 (3.4)	8.9 (7.6)	2.6 (2.6)	2.9 (3.0)	3.6 (3.6)	3.3 (3.4)	3.0 (3.1)
CC1/2	0.97 (0.84)	0.98 (0.83)	0.99 (0.58)	0.99 (0.89)	0.98 (0.68)	0.99 (0.84)	0.99 (0.79)	0.99 (0.79)
Rpim	0.10 (0.24)	0.09 (0.44)	0.11 (0.58)	0.07 (0.33)	0.13 (0.46)	0.06 (0.39)	0.05 (0.51)	0.08 (0.42)
Avg. I/σ (I)	6.1(2.9)	8.1 (2.0)	6.8 (1.9)	5.9 (1.8)	5.9 (2.2)	10.8 (2.6)	10.8 (2.3)	8.1 (2.2)
Refinement Statistics								
Resolution range (Å)	45.57-1.80	46.08-2.15	86.88-2.00	46.72-2.40	45.16-2.85	45.37-1.65	45.17-1.60	45.27-1.80
R_{work}/R_{free}	0.17/0.21	0.24/0.28	0.16/0.19	0.23/0.27	0.17/0.22	0.13/0.16	0.14/0.17	0.16/0.20
No. atoms								
Protein	11296	11296	11296	11296	11266	11301	11296	11296
Water	1212	867	1249	480	193	1545	1440	1041
Ligand	17	59	44	50	26	124	86	64
Root-Mean-Square Deviations								
Bond length (Å)	0.009	0.008	0.009	0.005	0.003	0.013	0.013	0.011
Bond angles (deg)	1.45	1.53	1.49	1.45	1.29	1.83	1.81	1.67
Ramachandran favored/outliers (%)	96.9/0.4	96.1/0.2	96.9/0.2	94.8/0.6	95.8/0.4	96.8/0.3	96.6/0.4	96.4/0.4
Residues in the AU	1464	1464	1464	1460	1460	1461	1464	1464
PDB code	6R5I	6R5U	6R5N	6R5O	6R5P	6R5R	6R5T	6R5V

*Values in parentheses are referred to the highest resolution shell indicated.**LMT: laminaritriose; NOJ: 1-deoxynojirimycin; BGC: β-D-glucose,

Table S4. Loops and amino-acid residues that form the active site

Loop	*Residues of the loop*	*Residues at the dimer interface (residue identifier numbers)*
L1	D110-I121	None
L2	F154-D176	162,165
L3	K207-S227	216-223, 225-230
L4	S253-N265	254, 256, 258-261
L5	S285-G288	286-287 (292, 294, 295, 296, 297, 298, 299, 300)
L6	I426-T440	None
L7	S475-S486	472, 473, 476, 477
L8	E510-P527	514, 516-523, 525
L9	F585-E649	601-608, 611-617, 622, 624-633, 635

Table S5. GH3 family enzymes reported in the RCSB PDB database and comparisons to BglX

PDB ID	Dimer/Monomer	Macromolecule Name	Source organism	Identity to BglX	Similarity to BglX	Structure. Rel. Date
BglX	D	BglX	*Pseudomonas aeruginosa*	100.00	100.00	NA
4ZO6	D	Lin1840 protein	*Listeria innocua*	38.94	54.04	2016-05-18
3U48	D	JMB19063	*metagenome*	49.60	67.47	2013-04-10
5TF0	D	GH3 N-terminal domain protein	*Bacteroides intestinalis*	47.58	64.92	2016-10-05
5XXL	D	Periplasmic beta-glucosidase	*Bacteroides thetaiotaomicron*	33.56	47.65	2017-12-13
5Z87	M	EmGH1	*Erythrobacter marinus*	33.12	50.39	2019-02-06
5Z9S	M	Beta-glucosidase	*Bifidobacterium longum*	31.49	45.84	2018-03-14
5A7M	M	Beta-xylosidase	*Trichoderma reesei*	19.23	32.75	2016-08-10
5YOT	M	Isoprimeverose-producing enzyme	*Aspergillus oryzae*	28.63	46.91	2018-11-07
2X40	M	Beta-glucosidase	*Thermotoga neapolitana*	21.93	36.56	2010-02-09
3ABZ	M	Beta-glucosidase	*Kluyveromyces marxianus*	19.96	32.98	2010-08-11
5WAB	M	beta-glucosidase	*Bifidobacterium adolescentis*	23.57	36.87	2018-04-04
5WUG	M	Beta-glucosidase	*Paenibacillus barengoltzii*	6.15	11.42	2018-01-31
5K6L	M	Beta-glucosidase	*metagenome*	5.26	10.27	2016-10-05
3RRX	M	beta-glucanase	*Pseudoalteromonas sp. BB1*	21.56	33.42	2011-12-21
3WLH	M	ExoI	*Hordeum vulgare*	21.15	34.01	2015-03-25
5M6G	M	Beta-glucosidase	*Saccharopolyspora erythraea*	24.32	37.71	2017-11-29
5JP0	M	Beta-glucosidase BoGH3B	*Bacteroides ovatus*	27.74	45.15	2016-08-10
5JU6	M	Beta-glucosidase	*Rasamsonia emersonii*	16.77	26.05	2016-07-13
4I3G	M	Beta-glucosidase	*Streptomyces venezuelae*	21.26	33.10	2013-02-20
4I8D	M	Beta-glucosidase	*Trichoderma reesei*	19.39	32.73	2012-12-19
4IIB	M	Beta-glucosidase	*Aspergillus aculeatus*	17.06	27.44	2013-04-10
5FJI	M	Beta-glucosidase	*Aspergillus fumigatus*	16.95	26.84	2016-02-10
5NBS	M	Beta-glucosidase	*Neurospora crassa*	16.88	26.42	2018-03-21
5VQD	M	Beta-glucoside phosphorylase	*metagenome*	12.04	23.83	2018-01-17
3TEV	M	Beta-hexosaminidase	*Deinococcus radiodurans*	7.16	15.25	2011-08-31
3SQL	M	Beta-hexosaminidase	*Synechococcus sp. PCC 7002*	11.49	20.82	2011-07-20
5G5K	M	Beta-hexosaminidase	*Pseudomonas aeruginosa*	8.40	14.30	2017-04-12
2OXN	M	Beta-hexosaminidase	*Vibrio cholerae*	6.96	13.93	2007-06-12
3WO8	M	Beta-hexosaminidase	*Thermotoga maritima*	9.01	19.60	2014-12-24
4GVF	M	Beta-hexosaminidase	*Salmonella enterica*	6.04	12.34	2012-12-12
4GYK	M	Beta-hexosaminidase	*Bacillus subtilis*	14.93	27.67	2012-12-12
5BU9	M	Beta-hexosaminidase	*Beutenbergia cavernae*	9.57	16.43	2015-06-17
5BZA	M	Beta-hexosaminidase	*Thermotoga neapolitana*	9.61	19.47	2015-09-16
4MSS	M	Beta-hexosaminidase	*Burkholderia cenocepacia*	7.92	13.38	2013-10-30
4YYF	M	Beta-hexosaminidase	*Mycolicibacterium smegmatis*	9.41	15.69	2015-04-08
4ZM6	M	Beta-hexosaminidase	*Rhizomucor miehei*	10.43	20.47	2015-12-09
5IOB	M	Beta-hexosaminidase	*Corynebacterium glutamicum*	7.76	14.36	2016-03-23

.

Table S6. **Kinetics parameters for the wild-type and variant of BglX**

	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
wild type	14 ± 2	0.057 ± 0.004	4.1 ± 0.6
D286N	–	–	–

Table S7. Statistics of the hydrolase enzymes in the PDB databank (as of 26 June 2019)

Advanced search parameter at the RCSB PDB database	Number of hits (structures)
EnzymeClassificationTree Search for 3: Hydrolases	28,973
EnzymeClassificationTree Search for 3: Hydrolases and Representative Structures at 95% Sequence Identity	7,956
EnzymeClassificationTree Search for 3: Hydrolases and Stoichiometry in biological assembly: Stoichiometry is A2	5,189
EnzymeClassificationTree Search for 3: Hydrolases and Stoichiometry in biological assembly: Stoichiometry is A2 and Representative Structures at 95% Sequence Identity	1,803

Legends for Movies S1–S5

Movie S1. (See the supplementary movie file). Molecular-dynamic simulation of the wild-type apo dimer structure of BglX. The L9 loop of the monomer-II (M-II; red color) inserts into the active site of monomer-I (M-I; colored in gray). The L7 loop of the M-I supports the binding of L9 to the site. Relevant active-site residues are shown in sphere representation. All the supplementary movie files were rendered using VMD program.[18]

Movie S2. (See the supplementary movie file). Molecular-dynamic simulation of the native wild-type enzyme-substrate complex (Michaelis complex). The important residues are labeled. Catalytic residues (general acid/base E517 and nucleophile D286) are pointed with arrows.

Movie S3. (See the supplementary movie file). QM/MM simulation of the glycosylation step.

Movie S4. (See the supplementary movie file). Molecular-dynamic simulation of the model of glycosyl-enzyme intermediate with the first product at the +1/+2 site showing entry of water molecule in between the C1 carbon of the -1 site glycosyl species and the general acid/base residue E517 prior to the deglycosylation step. The first product of the reaction (laminaribiose) adjusts position for the entry for water molecule.

Movie S5. (See the supplementary movie file). QM/MM simulation of the deglycosylation step.

References

(1) Nieto, C., Fernandez-Tresguerres, E., Sanchez, N., Vicente, M., and Diaz, R. (1990) Cloning vectors, derived from a naturally occurring plasmid of *Pseudomonas savastanoi*, specifically tailored for genetic manipulations in Pseudomonas, *Gene 87*, 145-149.

(2) Choi, K. H., Kumar, A., and Schweizer, H. P. (2006) A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation, *J. Microbiol. Methods 64*, 391-397.

(3) Macdonald, S. S., Blaukopf, M., and Withers, S. G. (2015) *N*-acetylglucosaminidases from CAZy family GH3 are really glycoside phosphorylases, thereby explaining their use of histidine as an acid/base catalyst in place of glutamic acid, *J. Biol. Chem. 290*, 4887-4895.

(4) Hay, I. D., Remminghorst, U., and Rehm, B. H. (2009) MucR, a novel membrane-associated regulator of alginate biosynthesis in *Pseudomonas aeruginosa*, *Appl. Environ. Microbiol. 75*, 1110-1120.

(5) Rossi, E., Paroni, M., and Landini, P. (2018) Biofilm and motility in response to environmental and host-related signals in Gram negative opportunistic pathogens, *J. Appl. Microbiol.*

(6) Valentini, M., and Filloux, A. (2016) Biofilms and Cyclic di-GMP (c-di-GMP) signaling: Lessons from *Pseudomonas aeruginosa* and other bacteria, *J. Biol. Chem. 291*, 12547-12555.

(7) Hauser, A. R. (2009) The type III secretion system of Pseudomonas aeruginosa: infection by injection, *Nat. Rev. Microbiol. 7*, 654-665.

(8) Jones, C. J., Ryder, C. R., Mann, E. E., and Wozniak, D. J. (2013) AmrZ modulates *Pseudomonas aeruginosa* biofilm architecture by directly repressing transcription of the psl operon, *J. Bacteriol. 195*, 1637-1644.

(9) Mao, F., Dam, P., Chou, J., Olman, V., and Xu, Y. (2009) DOOR: a database for prokaryotic operons, *Nucleic Acids Res. 37*, D459-463.

(10) Dik, D. A., Dominguez-Gil, T., Lee, M., Hesek, D., Byun, B., Fishovitz, J., Boggess, B., Hellman, L. M., Fisher, J. F., Hermoso, J. A., and Mobashery, S. (2017) Muropeptide binding and the X-ray structure of the effector domain of the transcriptional regulator AmpR of *Pseudomonas aeruginosa*, *J. Am. Chem. Soc. 139*, 1448-1451.

(11) Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, *Mol. Biol. Evol. 35*, 1547-1549.

(12) Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol. 4*, 406-425.

(13) Zuckerkandl, E., and Pauling, L. (1965) Evolutionary Divergence and Convergence in Proteins, In *Evolving Genes and Proteins* (Bryson, V., and Vogel, H. J., Eds.), pp 97-166, Academic Press.

(14) Acebron, I., Mahasenan, K. V., De Benedetti, S., Lee, M., Artola-Recolons, C., Hesek, D., Wang, H., Hermoso, J. A., and Mobashery, S. (2017) Catalytic cycle of the *N*-acetylglucosaminidase NagZ from *Pseudomonas aeruginosa*, *J. Am. Chem. Soc. 139*, 6795-6798.

(15) de M Seabra, G., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007) Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the amber molecular dynamics package, *J. Phys. Chem. A 111*, 5655-5664.

(16) Liu, J., Wang, X., and Xu, D. (2010) QM/MM study on the catalytic mechanism of cellulose hydrolysis catalyzed by cellulase Cel5A from Acidothermus cellulolyticus, *J. Phys. Chem. B 114*, 1462-1470.

(17) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol. 372*, 774-797.

(18) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, *J. Mol. Graph. Mod. 14*, 33-38, 27-38.