SCIENTIFIC REPORTS

β-glucosidases from thermophilic cellulolytic fungus *Humicola insolens*

Wei Xia , *, Yingguo Bai *, Ying Cui , Xinxin Xu , Lichun Qian , Pengjun Shi , Wei Zhang , Huiying Luo , Xiuan Zhan & Bin Yao

The fungus ***Humicola insolens*** is one of the most powerful decomposers of crystalline cellulose. However, studies on the β

β-glucosidases from ***Humicola insolens***

Pichia pastoris

Hi

towards aryl β-glucosides while ***Hi*** ***Hi***

β ***Hi***

p

Hi ***Hi***

β

+

The depletion of fossil fuel at enhanced rate and accompanied adverse effects on the global economic and environment has accelerated the research on its alternatives. Cellulosic materials like agricultural wastes and crop by-products (corn stover, wheat straw, bagasse, etc) are the most abundant polysaccharides in nature and represent the most valuable source of renewable energy[1]. Thus efficient utilization of cellulose biomass has been attracting attentions worldwide for the sustainable development and eco-efficiency[2–4]. Cost-effective process of enzymatic hydrolysis requires low production cost and highly active enzymes with great inhibitor tolerance and synergistic actions. In nature, complete hydrolysis of cellulose needs the synergistic action of a whole cellulolytic enzyme system, which includes endo-β-glucanase (EC 3.2.1.4), cellobiohydrolase (EC 3.2.1.91) and β-glucosidase (EC 3.2.1.21)[1,5,6]. β-Glucosidase can accelerate the decomposition of cellulose and improve the glucose yield by catalyzing the rate-limiting step of cellobiose hydrolysis[7,8]. The hyperproducing mutant strains of *Trichoderma reesei* are commercial producers of highly active cellulase (i.e. Celluclast 1.5

www.cazy.org/)[9,10]. Besides biomass conversion, β-glucosidases are applied in other biological processes, such as biogenesis of various functional molecules (e.g., terpenols, flavonoids, phytohormones) from glycoside precursors[11–13]. Considering the great differences among homologous GH3 β-glucosidases[14], an insight into the substrate specificity of β-glucosidases is beneficial for better utilization of this multifunctional biocatalyst. With the development of crystal determination, there have been several resolved GH3 β-glucosidase structures[15–17]. And several conserved substrate bind sites were verified in single protein by crystallisation of inhibitor complex or experimental determination, such as Arg156 and Tyr511 of GH3 β-glucosidase AaBGL1 from *Aspergillus aculeatus* (PDB: 4IIB)[15], and Trp49 of β-glucosidase from *Aspergillus niger*[18]. But few researches were conducted to investigate the functional diversity and substrate specificity of multiple β-glucosidases from the same species.

Thermophilic filamentous fungi are excellent microbial sources of highly-active, thermostable β-glucosidases for industrial purposes[19]. To date, filamentous fungi including *Trichoderma*, *Aspergillus* and *Penicillium* are the main microbial sources of industrial cellulases[20,21]. Thermophilic *Humicola* spp. are also reported to have the [illegible] to produce various cellulolytic enzymes[22–28]. However, the studies on β-glucosidases from this fungus appeared rather scanty. Up to now, an intracellular glucose- and xylose-stimulated β-glucosidase of GH1 (BglHi1) and a purified extracellular β-glucosidase of GH3 (BglHi2) from *H. insolens* have been biochemically characterized[24,25]. In this study, three GH3 β-glucosidase encoding genes (*Hibgl3A–C*) were cloned from *H. insolens* strain Y1 and successfully expressed in *Pichia pastoris* strain GS115. The enzymes were all most active under neutral [illegible] conditions, but showed distinguished substrate specificity, catalytic efficiency and glucose tolerance. Further site-directed mutagenesis revealed the vital role of three residues in the substrate specificity of GH3 β-glucosidases.

Results

Gene cloning a[illegible]

cDNA)

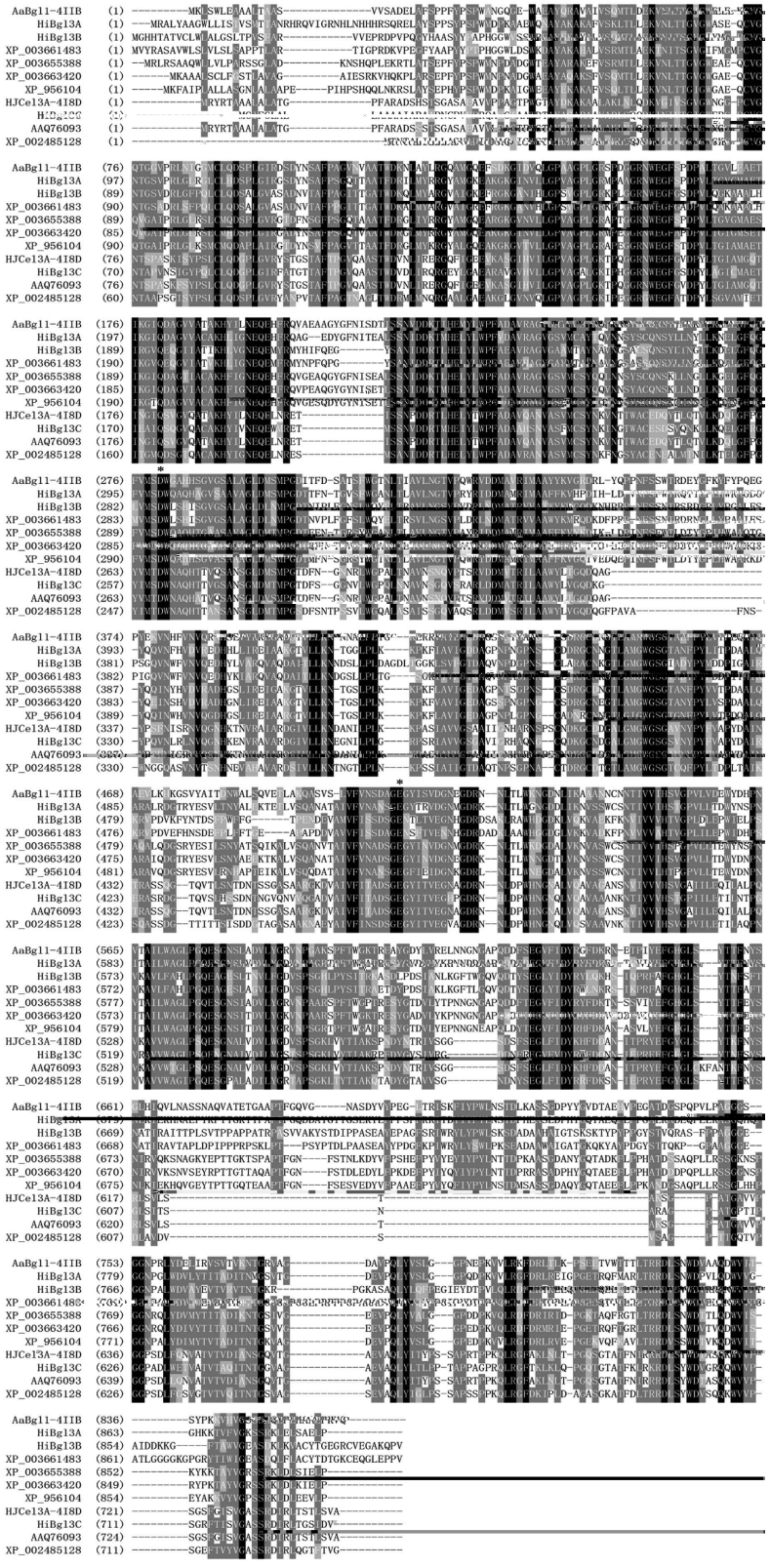

Figure 1. Sequence alignment of *Hi*Bgl3A, *Hi*Bgl3B and *Hi*Bgl3C with other GH3 β-glucosidases. The source and PDB codes or genebank accession numbers of these β-glucosidases are *Aspergillus aculeatus* (4IIB), *Hypocrea jecorina* (3ZYZ), *Myceliophthora thermophila* ATCC 42464 (XP_003663420), *Thielavia terrestris* NRRL 8126 (XP_003655388), *Neurospora crassa* OR74A (XP_956104), *Trichoderma viride* (AAQ76093) and *Talaromyces stipitatus* ATCC 10500 (XP_002485128). Identical and similar amino acids are indicated by black and gray shades, respectively. The putative catalytic residues were marked with asterisks. The three unique residues existed in HiBgl3B were framed by red rectangle.

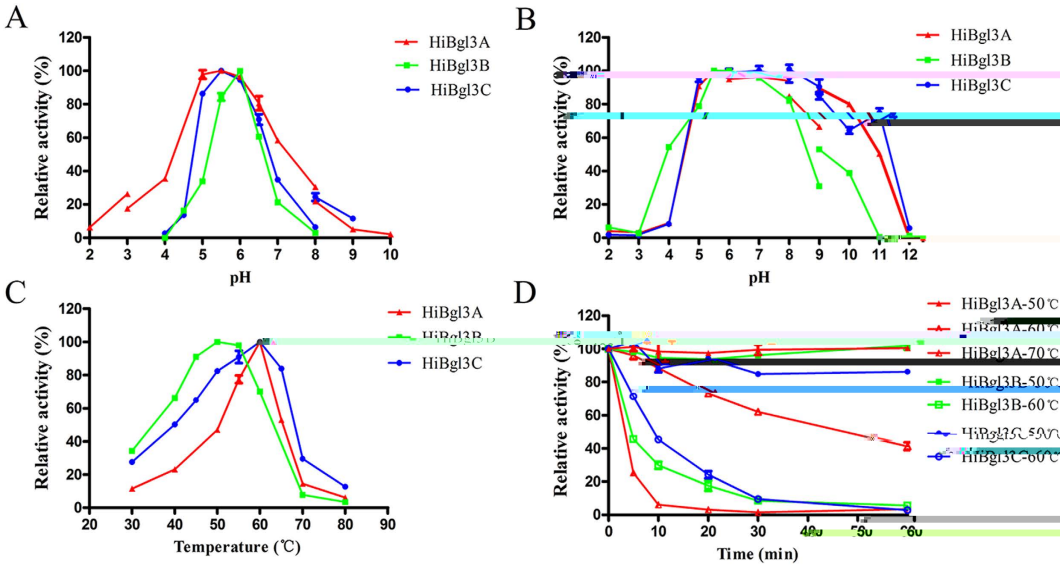

Figure 2. Enzymatic properties of the purified recombinant β-glucosidases using *p*NPG as the substrate. The relative activities of *Hi*Bgl3A (red triangle), *Hi*Bgl3B (green square) and *Hi*Bgl3C (blue circle) were plotted in the line chart. (**A**) Effect of pH on activities. (**B**) pH stability. (**C**) Effect of temperature on activities. (**D**) Thermostability at different temperatures. Each value in the panel represents the means ± SD (n = 3).

Substrate[a]	Specific activity (U/mg)		
	HiBgl3A	HiBgl3B	HiBgl3C
Disaccharide			
Cellobiose	36.3 ± 0.2	ND[b]	56.4 ± 1.2
Sophorose	94.6 ± 0.9	ND	113.9 ± 0.6
Gentiobiose	103.4 ± 1.6	ND	50.4 ± 0.5
Aryl β-glycoside			
*p*NPG	57.5 ± 0.3	31.6 ± 0.6	158.8 ± 2.0
*p*NPC	3.5 ± 0.2	1.7 ± 0.2	17.8 ± 0.2
*p*NPX	1.6 ± 0.1	0.8 ± 0.2	4.1 ± 0.2
*p*NPGal	1.3 ± 0.1	0.8 ± 0.3	1.9 ± 0.2
*p*NPAf	3.1 ± 0.2	ND	ND
Daidzin	20.4 ± 0.5	62.6 ± 2.3	80.0 ± 1.7
Genistin	25.5 ± 0.8	28.6 ± 0.4	22.1 ± 0.1
Glycitin	19.0 ± 1.1	15.9 ± 0.8	8.8 ± 0.5

Table 2. Substrate specificity of the three *H. insolens* β-glucosidases. [a]The final concentration of each substrate is 1 mM. [b]ND, not detected.

The substrate specificities of the three *H. insolens* β-glucosidases are shown in Table 2. When using disaccharides of different linkages as the substrate, the enzymes showed different preference, gentiobiose (β-1,6 linkage) > sophorose (β-1,2 linkage) > cellobiose (β-1,4 linkage) for *Hi*Bgl3A, sophorose > cellobiose > gentiobiose for *Hi*Bgl3C, respectively, and no *Hi*Bgl3B activity against all tested disaccharides. Aryl β-glycoside substrates (4-nitrophenyl compounds and soy isoflavones) that have a phenyl at subsite +1 were also tested. All enzymes showed much higher activities towards *p*NPG (over 8 fold) than against *p*NPC, *p*NPX, *p*NPGal and *p*NPAf, but varied in the hydrolysis of soy isoflavones. The activities of *Hi*Bgl3B and *Hi*Bgl3C against soy isoflavones followed the order of daidzin > genistin > glycitin, while *Hi*Bgl3A had relatively low but similar activities against the three tested soy isoflavone substrates. All enzymes had no observable activity on polysaccharides (barley β-glucan, sodium carboxymethylcellulose, Avicel, laminarin and lichenin).

The kinetics of *H. insolens* β-glucosidases on substrates *p*NPG and cellobiose are shown in Table 3. In comparison with the other two counterparts, *Hi*Bgl3C exhibited much higher substrate affinity (the lowest K_m) and catalytic efficiency (k_{cat}/K_m). The glucose inhibition was also evaluated using *p*NPG as the substrate. *Hi*Bgl3B and *Hi*Bgl3C exhibited relatively high tolerance to glucose than that of *Hi*Bgl3A.

Species	Enzymes	Optimum		K_m (mM)		k_{cat}/K_m (/s/mM)		K_i (mM)		References
		pH	T (°C)	*p*NPG	CB	*p*NPG	CB	*p*NPG	CB	
Humicola insolens Y1	*Hi*Bgl3A	5.5	60	0.90	8.44	81.6	11.1	25.0	–	This work
H. insolens Y1	*Hi*Bgl3B	6.0	50	1.51	–	28.6	–	55.2	–	This work
H. insolens Y1	*Hi*Bgl3C	5.5	55	0.20	6.63	1557	23.0	37.1	–	This work
Aspergillus niger	N188	4.8	50	0.57	0.88	41	36	2.7	–	30
Aspergillus foetidus		4.8	65	0.41	–	–	–	8.1	–	31
Aspergillus japonicus		5.0	40	0.60	0.95	432	368	2.73	–	32
Aspergillus oryzae		5.0	50	0.29	1.96	1270	510	2.9	5	33
Fomitopsis palustris		5.0	50	0.12	4.8	6160	21	0.35	–	34
Hypocrea jecorina		5.0	50	0.09	0.75	466	157	0.51	–	32
Myceliophthora thermophila		5.0	40	0.39	2.64	376	17	0.28	–	35
Trichoderma reesei		4.5	50	–	0.54	–	41	–	0.29	36
Thermoanaerobacterium thermosaccharolyticum DSM 571	rBGL	6.4	70	0.63	7.9	–	13.3	600	–	37
Uncultured bacterium		6.0	40	0.39	20.4	–	0.65	1000	–	38
Neocallimastix patriciarum	*Npa*BGS	6.0	40	–	–	–	–	–	–	39
Cow rumen metagenome	LAB25g2	5.2	50	0.45	4.88	0.92	0.2	–	–	40

Table 3. Property comparison of microbial β-glucosidases[a]. [a]*p*NPG was used for the determination of optimal conditions, and *p*NPG and cellobiose (CB) were used for the determination of kinetics and glucose inhibition.

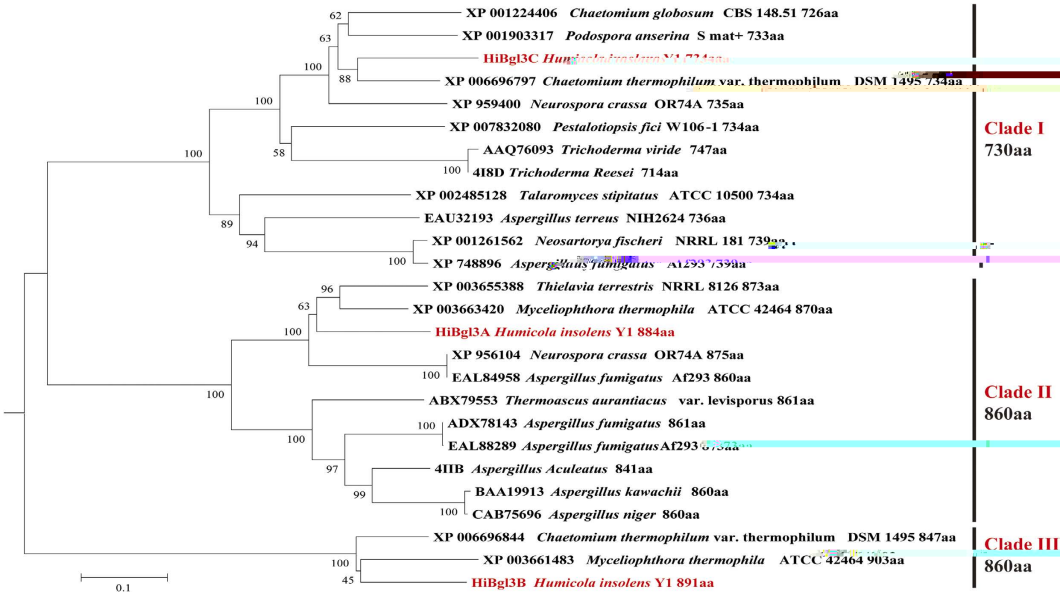

Figure 3. The phylogenetic tree generated from the analysis of *H. insolens* β-glucosidases and other closely related β-glucosidases amino acid sequences in the NCBI database using the Neighbor-Joining method. The numbers on nodes correspond to the percentage bootstrap values for 1,000 replicates. The accession number of each SOD in GenBank is labelled prior to the species name.

Design, construction and specific activities determination of mutants. Three isoenzymes exhibited distinct features in terms of substrate specificity. To investigate the evolutionary relationship of GH3 β-glucosidases, a phylogenetic analysis on the amino acid sequences of *H. insolens* β-glucosidases and counterparts obtained from the NCBI database using the Neighbor-Joining (NJ) method (shown in Fig. 3) indicated that *Hi*Bgl3A, *Hi*Bgl3B and *Hi*Bgl3C belonged to different evolutionarily related clades. Those belonging to clade II and III have similar length, about 860 amino acids, with different substrate specificities. Sequence alignment suggested that three distinct unique residues (framed by red rectangle in Fig. 1) existed in *Hi*Bgl3B and the homologous protein XP_003661483, i.e. Ile48, Ile278 and Thr484, which were generally Trp, Phe, and Tyr, respectively, in most GH3 β-glucosidases. These three residues are all located at the entrance of the enzyme's catalytic pocket, and may relate to substrate specificity. Thus we conducted site-directed mutagenesis on *Hi*Bgl3A and *Hi*Bgl3B of similar lengths to verify the impact of these three sites on recognizing +1 subsite of different substrates. The specific activities and kinetic values of mutants towards aryl β-glycoside *p*NPG and three disaccharides were measured

(shown in Table 4). Although the single and combined mutants I48W, I278F, T484Y and I48W/I278F/T484Y of *Hi*Bgl3B showed decreased activities towards *p*NPG, all the mutants conferred obvious hydrolysis activities on sophorose, which were much higher than the activities on *p*NPG of their own. However, no activity on the other two disaccharides was detected. In contrast, negative mutants of *Hi*Bgl3A all lost hydrolysis activities towards disaccharides, and the mutants W69I and F304I even became completely inactivated towards *p*NPG (data not shown). The turnover numbers (k_{cat}) of *Hi*Bgl3B mutants all decreased substantially while no significant change was found in the K_m values, leading to declines in catalytic efficiency. However, introduction of mutation Y509T caused the increase of the K_m value, but did not affect the k_{cat} value.

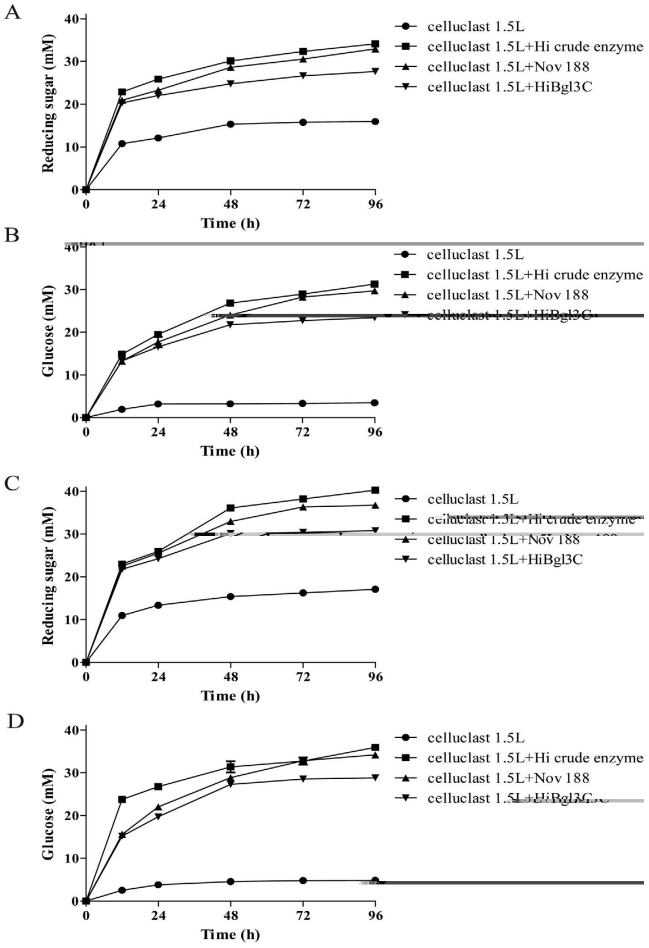

Figure 4. Enzymatic saccharifications of pretreated cellulose materials. The concentrations (mM) of reducing sugar and glucose in the hydrolyzates of corn stover and Avicel by different enzyme combinations (cellulase at 5 FPU/g and β-glucosidase at 11 BGU/g, resepctivley) were shown in (**A**–**D**), respectively. Circle: Celluclast 1.5 L only; square: Celluclast 1.5 L plus crude enzyme of *H. insolens* Y1; upper triangle: Celluclast 1.5 L plus Novozyme 188; lower triangle: Celluclast 1.5 L plus *Hi*Bgl3C.

a very low specific activity on *p*NPG (0.6 U/mg)[14]. This kind of β-glucosidase is defined as aryl β-glucosidase, and widely exists in microorganisms[44]. A previous study has reported that GH3 β-glucosidases have a strict stereochemical requirement to accommodate β-D-glucopyranose at subsite −1, while subsites + seem insignificant in both substrate binding and hydrolysis[10]. This relative plasticity at subsite +1 might account for the broad substrate specificity of GH3 β-glucosidases towards different aglycon structures. Besides, β-glucosidases of different sequence clades may vary in conformation at the subsite +1, consequently leading to variations in the activities on substrates with different aglycons at this site. Based on the multiple sequence alignment and structure analysis, three conserved substrate recognizing residues for the subsite +1 of cellobiose were identified in the β-glucosidases capable of hydrolyzing cellobiose[15,17], for instance, Trp68, Phe305 and Tyr511 of *Aa*BGL1 from *Aspergillus aculeatus* (PDB: 4IIB), were substituted by Ile48, Ile278 and Thr484 of *Hi*Bgl3B, respectively (see Fig. 1)[15]. The lack of ability to interact with the second sugar ring of disaccharides might be responsible for its unique substrate specificity. It could be speculated that this sort of β-glucosidases does not contribute to the deconstruction of cellulose, though the natural function is unclear yet. This difference indicates their functional diversity.

Activity changes between wild type and mutant enzymes suggested that Ile48, Ile278 and Thr484 have effects on the substrate recognition of enzymes. To investigate the interactions between I48, I278 and T484 and sophorose, a molecular docking was performed, and the results were shown in Fig. 5A. The interactions were analyzed by software Ligplot^{+} and shown in the planar graph Fig. 5B. W48, F278 and Y484 are supposed to interact with ligand sophorose at the subsite +1 by hydrophobic and polar interactions, which are absent in wild type *Hi*Bgl3B. And compared to other two disaccharides β-1-4-cellobiose and β-1-6-gentiobiose, the hydroxyl at the C6 site of +1 subunit of β-1-2-sophorose was outward, thus requiring smaller spatial position to locate into the binding pocket. It may explain why *Hi*Bgl3B mutants had no activity on cellobiose and gentiobiose. In contrast, wild type *Hi*Bgl3B is able to combine the highly hydrophobic aromatic ring at its +1 binding site. However, the binding affinity seems to vary significantly among different sequences and structures because the same case did not apply

to the negative mutants of HiBgl3A. For example, the single mutants of *Hi*Bgl3A had no activity towards disaccharides, and no accumulated effect was found in the combination mutant of *Hi*Bgl3B. We conjecture that these three residues in *Hi*Bgl3A must have an integral role in the binding with disaccharide substrates. Without Trp69 and Phe304, the enzyme lost hydrolysis activities on both aryl aglycones and disaccharides, indicating the indis-

were purchased from Promega (Madison, WI). The cDNA synthesis kit was purchased from TransGen. Barley β-glucan, Avicel, 4-nitrophenyl β-D-glucopyranoside (*p*NPG), 4-nitrophenyl β-D-xylopyranoside (*p*NPX), 4-nitrophenyl α-L-arabinofuranoside (*p*NPAf), 4-nitrophenyl α-D-galactopyranoside (*p*NPGal), 4-nitrophenyl α-L-arabinopyranoside (*p*NPAb), 4-nitrophenyl β-D-cellobioside (*p*NPC), disaccharides cellobiose, sophorose and gentibiose and soybean flavones daidzin, genistin and glycitin were all purchased from Sigma-Aldrich. Sodium carboxymethylcellulose (CMC-Na), laminarin and lichenin were obtained from Megazyme (Wicklow, Ireland). All other chemicals used were of analytical grade and commercially available.

Gene cloning and sequence analysis. The total RNA of *H. insolens* Y1 was extracted from the mycelia after 3 days' growth in the inducing wheat bran medium, and was reverse transcribed into cDNA by TransScript® One-Step gDNA Removal and cDNA Synthesis SuperMix kit (TransGen). The gene fragments without the signal peptide-coding sequence were amplified using *H. insolens* Y1 cDNA as the template, at an annealing temperature of 60 °C, with three specific primer sets (shown in Supplementary Table S2). The PCR products were purified and ligated into the pEASY-T3 vector for sequencing. Vector NTI Advance 10.0 software (Invitrogen) was used to analyze the DNA sequence and to predict the molecular weight and *p*I of proteins and perform multiple sequence alignments. The signal peptide and the potential *N*-glycosylation sites were predicted by the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) and the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/), respectively. The neighbor-joining phylogenetic tree based on the coding sequences of was performed by using MEGA (Version 6.0).

Expression and purification of the recombinant enzymes.

Kinetic parameters and glucose inhibition. The K_m, V_{max} and k_{cat} values of *H. insolens* β-glucosidases were determined under each optimal conditions for 5 min in 100 mM Na_2HPO_4-citric acid containing 1–10 mM *p*NPG

23. Karlsson, J. *et al.* Enzymatic degradation of carboxymethyl cellulose hydrolyzed by the endoglucanases Cel5A, Cel7B, and Cel45A from *Humicola insolens* and Cel7B, Cel12A and Cel45Acore from *Trichoderma reesei*. *Biopolymers.* **63,** 32–40 (2002).
24. Souzaa, F. H. M., Inocentesc, R. F., Warda, R. J., Jorgeb, J. A. & Furriel, R. P. M. Glucose and xylose stimulation of a β-glucosidase from the thermophilic fungus *Humicola insolens*: A kinetic and biophysical study. *J. Mol. Catal. B. Enzym*. **94,** 119– 128 (2013).
25. Meleiro, L. P. *et al.* A novel β-glucosidase from *Humicola insolens* with high potential for untreated waste paper conversion to sugars. *Appl. Biochem. Biotechnol.* **173,** 391–408 (2014).
26. Boisset, C., Pétrequin, C., Chanzy, H., Henrissat, B. & Schülein, M. Optimized mixtures of recombinant *Humicola insolens* cellulases for the biodegradation of crystalline cellulose. *Biotechnol. Bioeng.* **72,** 339–45 (2001).
27. Varrot, A. *et al.* Structural basis for ligand binding and processivity in cellobiohydrolase Cel6A from *Humicola insolens*. *Structure* **11,** 855–64 (2003).
28. Otzen, D. E., Christiansen, L. & Schülein, M. A comparative study of the unfolding of the endoglucanase Cel45 from *Humicola insolens* in denaturant and surfactant. *Protein. Sci.* **8,** 1878–87 (1999).
29. Lambertz, C. *et al.* Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnol. Biofuels.* **7,** 135 (2014).
30. Chauve, M. *et al.* Comparative kinetic analysis of two fungal β-glucosidases. *Biotechnol. Biofuels.* **3,** 3 (2010).
31. Decker, C. H., Visser, J. & Schreier, P. β-Glucosidases from five black *Aspergillus* species: study of their physico-chemical and biocatalytic properties. *J. Agric. Food. Chem.* **48,** 4929–36 (2000).
32. Korotkova, O. G. *et al.* Isolation and properties of fungal β-glucosidases. *Biochem. Mosc.* **74,** 569–577 (2009).
33. Langston, J., Sheehy, N. & Xu, F. Substrate specificity of *Aspergillus oryzae* family 3 β-glucosidase. *Biochim. Biophys. Acta*. **1764,** 972–978 (2006).
34. Yoon, J. J., Kim, K. Y. & Cha, C. J. Purification and characterization of thermostable β-glucosidase from the brown-rot basidiomycete *Fomitopsis palustris* grown on microcrystalline cellulose. *J. Microbiol.* **46,** 51–55 (2008).
35. Karnaouri, A., Topakas, E., Paschos, T., Taouki, I. & Christakopoulos, P. Cloning, expression and characterization of an ethanol tolerant GH3 β-glucosidase from *Myceliophthora thermophila*. *Peer. J.* **1,** e46 (2013).
36. Schmid, G. & Wandrey, C. Characterization of a cellodextrin glucohydrolase with soluble oligomeric substrates: experimental results and modeling of concentration-time-course data. *Biotechnol. Bioeng.* **33,** 1445–1460 (1989).
37. Pei, J., Pang, Q., Zhao, L., Fan, S. & Shi, H. *Thermoanaerobacterium thermosaccharolyticum* β-glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose. *Biotechnol. Biofuels.* **5,** 31 (2012).
38. Zemin, F. *et al.* Cloning and characterization of β-glucosidase from marine microbial metagenome with excellent glucose tolerance. *J. Microbiol. Biotechnol.* **20,** 1351–1358 (2010).
39. Chen, H. L. *et al.* A highly efficient β-glucosidase from the buffalo rumen fungus *Neocallimastix patriciarum* W5. *Biotechnol. Biofuels*. **5,** 24 (2012).
40. Del Pozo, M. V. *et al.* Microbial β-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail. *Biotechnol. Biofuels.* **5,** 73 (2012).
41. Christakopoulos, P. *et al.* Purification and characterisation of an extracellular β-glucosidase with transglycosylation and exo-glucosidase activities from *Fusarium oxysporum*. *Eur. J. Biochem.* **224,** 379–85 (1994).
42. Bronnenmeier, K. & Staudenbauer, W. L. Purification and properties of an extracellular β-glucosidase from the cellulolytic thermophile *Clostridium stercorarium*. *Appl. Microbiol. Biotechnol.* **28,** 380–386 (1988).
43. Hong, M. R. *et al*. Characterization of a recombinant β-glucosidase from the thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *J. Biosci. Bioeng.* **108,** 36–40 (2009).
44. Gao, J. & Wakarchuk, W. Characterization of five β-glycoside hydrolases from *Cellulomonas fimi* ATCC 484. *J. Bacteriol.* **196,** 4103–10 (2014).
45. Du, Y. *et al.* Characterisation of three novel thermophilic xylanases from *Humicola insolens* Y1 with application potentials in the brewing industry. *Bioresour. Technol.* **130,** 161–167 (2013).
46. Miller, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31, 426–428 (1959).

This research was supported by the National High Technology Research and Development Program of China (863 Program, 2012AA022105 and 2012AA022208), the China Modern Agriculture Research System (CARS-42) and the Key Scientific and Technological Innovation Group Program "Feed Research and Security" of Zhejiang Province (2011R50025-12).

W.X. and Y.B. performed the enzyme production, activity assay, acquisition of data and drafted the manuscript. Y.C. and X.X. performed the TLC analysis. P.S. and W.Z. performed the enzymatic saccharification. H.L. and X.Z. performed the data processing and interpretation, and participated in revising the manuscript. L.Q. and B.Y. were the corresponding authors; they designed the study and revised the manuscript critically for important intellectual content. All authors read and approved the final version of the manuscript.

Accession codes: The cDNA sequences have been submitted to GenBank, and accession numbers are KT203370, KT203372 and KT203372 for Hibgl3A, Hibgl3B, and Hibgl3C, respectively.

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xia, W. *et al.* Functional diversity of family 3 β-glucosidases from thermophilic cellulolytic fungus *Humicola insolens* Y1. *Sci. Rep.* **6**, 27062; doi: 10.1038/srep27062 (2016).