J Ind Microbiol Biotechnol (2014) 41:489–498
DOI 10.1007/s10295-013-1377-0

BIOENERGY/BIOFUELS/BIOCHEMICALS

SIMB
Society for Industrial Microbiology and Biotechnology

Directed evolution of GH43 β-xylosidase XylBH43 thermal stability and L186 saturation mutagenesis

Sanjay K. Singh · Chamroeun Heng · Jay D. Braker · Victor J. Chan · Charles C. Lee · Douglas B. Jordan · Ling Yuan · Kurt Wagschal

Received: 12 March 2013 / Accepted: 24 October 2013 / Published online: 29 November 2013

Abstract Directed evolution of β-xylosidase XylBH43 using a single round of gene shuffling identified three mutations, R45K, M69P, and L186Y, that affect thermal stability parameter $K_t^{0.5}$ by -1.8 ± 0.1, 1.7 ± 0.3, and 3.2 ± 0.4 °C, respectively. In addition, a cluster of four mutations near hairpin loop-D83 improved $K_t^{0.5}$ by ~3 °C; none of the individual amino acid changes measurably affect $K_t^{0.5}$. Saturation mutagenesis of L186 identified the variant L186K as having the most improved $K_t^{0.5}$ value, by 8.1 ± 0.3 °C. The L186Y mutation was found to be additive, resulting in $K_t^{0.5}$ increasing by up to 8.8 ± 0.3 °C when several beneficial mutations were combined. While k_{cat} of xylobiose and 4-nitrophenyl-β-D-xylopyranoside were found to be depressed from 8 to 83 % in the thermally improved mutants, K_m, K_{ss} (substrate inhibition), and K_i (product inhibition) values generally increased, resulting in lessened substrate and xylose inhibition.

S. K. Singh and C. Heng contributed equally to the manuscript.

Electronic supplementary material The online version of this article (doi:10.1007/s10295-013-1377-0) contains supplementary material, which is available to authorized users.

S. K. Singh · L. Yuan (✉)
Department of Plant and Soil Sciences, Kentucky Tobacco Research and Development Center, University of Kentucky, Lexington, KY 40546, USA
e-mail: lyuan3@uky.edu

C. Heng · V. J. Chan · C. C. Lee · K. Wagschal (✉)
USDA Agricultural Research Service, Western Regional Research Center, Albany, CA 94710, USA
e-mail: kurt.wagschal@ars.usda.gov

J. D. Braker · D. B. Jordan
USDA Agricultural Research Service, National Center for Agricultural Utilization Research, Peoria, IL 61604, USA

Keywords Glycosyl hydrolase · Directed evolution · Gene shuffling · Thermal stability · Protein engineering

Introduction

The half-life of a glycosidic bond in nature is on the order of 5×10^6 years [32], underlying the ubiquity of cellulose and hemicelluloses as plant cell wall structural components and key repositories of biologically-fixated carbon. After cellulose, hemicelluloses are the second most abundant plant cell wall polymer type; they are heterogeneous in nature as far as monosaccharide constituents, main-chain and branched-chain monosaccharide linkage types, linkages between hemicelluloses and lignin, and further modifications such as acetylation and feruylation. One of the most common hemicelluloses is arabinoxylan, composed mainly of a xylose backbone with arabinose branch points, and abundantly present in agricultural crop residues including maize and all cereal grasses [5].

Glycosyl hydrolases (GH's), both ubiquitous and unending in variety like their substrates, catalyze glycosyl polymer deconstruction with such efficiency that they comprise some of the best enzyme catalysts known. A combination of enzyme types, primarily GH's, are necessary for complete arabinoxylan biomass enzymic deconstruction; these include β-D-xylosidases (EC 3.2.1.37), responsible for hydrolysis of xylooligosaccharides, predominately xylobiose (X2) to xylose [8, 9]. The β-xylosidases XylBH43 [27] and SXA [13], used for molecular breeding in this study, are members of glycosyl hydrolase family 43 (GH43), whose members include *endo*- and *exo*-α-L-arabinases (EC 3.2.1.99), α-L-arabinofuranosidases (EC 3.2.1.55), xylanases (EC 3.2.1.8), and galactan 1,3-β-galactosidases (EC 3.2.1.145) (http://cazy.org) [3]. SXA and XylBH43

are among the top five most efficient GH43 xylosidases reported for the hydrolysis of natural substrate X2 [16]. However, the GH43 xylosidases studied to date lose activity precipitously above 50 °C [16], while in biomass saccharification reactors, a higher thermal stability than that of the natural enzymes thus far encountered may be desired for faster reaction times, reduced matrix viscosity, and lower microbial contamination [6]. In nature, thermal stability is selected for directly by organisms pioneering thermally extreme environments. In established populations, the evolution of protein function likely occurs mostly under selection pressure for function within a temperature range, while also evolving indirectly for thermal stability since this confers increased protein structural stability. The latter in turn enables and promotes evolution by allowing proper folding, despite significant mutation [2]. Thus, evolution-based protein engineering of properties other than thermal stability may benefit by engineering thermally improved variants as starting materials.

Strategies to obtain thermally improved enzymes, which are often most effective when combined [6], include directed evolution [24], rational engineering, and screening thermophilic organisms. We show here that screening a small, ~3,840-member library created by a single round of gene shuffling resulted in identifying three amino acid residues and a cluster of mutations near loop-D83 that significantly affect thermal stability. Furthermore, we combined some of the mutations, resulting in an 8.8 ± 0.5 °C increase in $K_t^{0.5}$. Also identified were mutants with altered pH and inhibitor profiles, potentially desirous characteristics for xylosidase use in mechanistic studies and commercial biomass saccharification.

Materials and methods

DNA shuffled library

The DNA sequence of XylBH43 (PDB ID 1YRZ) was synthesized (Genscript USA Inc., Piscataway, NJ) with internal *Nde*I and *Xho*I restriction sites, removed as described previously [7], and subcloned into the pET29b(+) expression plasmid (EMD Chemicals, Inc., Gibbstown, NJ). For use in DNA shuffling, the DNA sequence of SXA (PDB ID 3C2U) was optimized by inspection for maximal identity with that of XylBH43, with the constraint of maintaining the native amino acid sequence (Supplementary Fig. S1), and was then synthesized and subcloned into pET29b(+). DNA shuffling was performed as described [24, 36], with some modifications. Parental DNA was amplified using primers 5′-CTCAGCTTCCTTTCGGGCTTTGTT-3′ and 5′-TGTGAGCGGATAACAATTCCCCTCTAG-3′, which correspond to regions of the pET29b(+) vector that are upstream and downstream of the *Nde*I and *Xho*I restriction sites. We mixed 5 μg of DNA from each parent, and subjected them to DNAse I digestion at 37 °C using supplied buffer in the presence of 1 mM Mg^{2+} and 0.16 U DNAse I in a final volume of 100 μl. The reaction was allowed to proceed for 14 min, and then a 20-μl aliquot was removed and quenched by mixing with 5 μl 500 mM EDTA on ice every 2 min thereafter. The polymerase chain reaction (PCR) reassembly reaction mixture contained 30 μl of fragments (100–200 bp), 5 μl of 10× Pfu DNA polymerase buffer, 2 μl of 10 mM dNTP mix, and 2.5 U of Pfu DNA polymerase, in a final volume of 50 μl. Reactions were performed as follows: 96 °C for 1.5 min, 35 cycles of denaturation at 94 °C for 30 s, nine hybridization steps separated by 3 °C from 65 to 41 °C for 1.5 min each, elongation step 72 °C for 4 min, and the final extension at 72 °C for 10 min. Fragments were amplified using the primers 5′-AGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG-3′ and 5′-GTGGTGGTGGTGGTGCTCGAG-3′, which contain the *Nde*I and *Xho*I restriction sites, respectively. In this step, 5 μl of reassembled product was amplified using 5 U of Pfu DNA polymerase, 2 μl of 10 mM dNTP mix, and the supplied buffer in a 100 μl reaction. The PCR protocol was: 95 °C for 2 min; 25 cycles of 1 min denaturation at 95 °C, hybridization step at 55 °C for 1 min, elongation at 72 °C for 4 min, and a final extension step at 72 °C for 10 min. DNA was purified using a 0.8 % agarose gel and Zymo Gel DNA purification kit (Zymo Research Corp., Irvine CA), and subcloned into the *Nde*I and *Xho*I restriction sites of the pET29b(+) expression plasmid. Electroporation was performed using electrocompetent *E. coli* BL21(DE3) cells according to the manufacturer's protocol (Lucigen Corp., Middleton, WI, USA).

L186 saturation mutagenesis

We constructed a gene library encoding variants containing all the possible amino acids at position 186, following the methods described by Pattanaik et al [20]. Two primer pairs (Supplementary Table 1) were designed to randomize position 186 in the nucleotide sequence. PfuUltra High-Fidelity DNA polymerase (Agilent Technologies, USA) was used in sPCR to minimize random point mutations. PCR using pET29b(+) XylBH43-L186 as a template was performed with a Peltier Thermal Cycler (MJ Research, USA) using the following program: preheat 94 °C for 2 min; 16 cycles of 94 °C for 30 s, 55 °C for 1 min and 68 °C for 8 min; followed by incubation at 68 °C for 1 h. The resulting randomized PCR products were purified using Wizard SV Gel and a PCR clean-up system (Promega, USA), and then treated with *Dpn*I (New England Biolabs, USA) for 1 h to eliminate the parental plasmid. The *Dpn*I enzyme was then heat inactivated at 75 °C for 15 min. The resulting plasmid

library was transformed into chemically competent TB1 cells. Plasmids were isolated from the overnight cultures of randomly selected colonies using Wizard Plus SV miniprep kit (Promega, USA) and sequenced to verify the mutation. Plasmids encoding variants with all 20 amino acids at position 186 were identified from 90 randomly selected bacterial colonies.

Thermal stability library screening

The XylBH43/SXA shuffled library was plated onto LB agar containing 30 μg/ml kanamycin (LB_{kan}) using 22 × 22 cm Q-trays (Molecular Devices LLC, Sunnyvale, CA), and incubated overnight at 30 °C. A Q-pix colony picker was used to transfer ~3,800 colonies to Greiner Fluorotrac 200 384-well microplates containing LB_{kan} and the Novagen Overnight Express™ autoinduction system. A whole-cell, in vivo 1° screen was performed using the substrate 4-methylumbelliferyl-β-D-xylopyranoside (MUX), as described previously [28], and ~1,300 clones with active enzyme expression were rearrayed into 96-well hit plates. Two rounds of a 2° screen were performed using crude enzyme lysates prepared using the hit plates as described [7], with a thermal challenge at temperatures of 55 °C for 1 h in round 1, and 57 °C for 1 h with an additional 45 min at 65 °C in round 2, to obtain 10 unique clones for further testing. For hit verification, we performed protein expression and purification from the 2° screen hits using Ni–NTA affinity chromatography followed by size-exclusion chromatography, as previously described [7]. Finally, we determined the temperature at which activity is halved over a 1 h incubation period ($K_t^{0.5}$), as described below.

$K_t^{0.5}$ determination

Hydrolytic activity rates before and after thermal treatment were assessed by measuring the rate of *p*-nitrophenyl leaving the group appearance, using a wavelength of 400 nm and 4 mM 4NPX substrate. Enzyme concentration was adjusted prior to the thermal challenge, such that initial rates were normalized to ~50 μM/min at 25 °C. A thermocycler (MJ Research, Watertown MA) was used to establish a temperature gradient of 20–25 °C, depending on the enzyme being tested, and the selected enzyme was incubated at various temperatures for 1 h in an assay buffer consisting of 100 mM phosphate, pH 6.5, and 1 mg/ml bovine serum albumin (BSA). After the thermal challenge, the temperature was lowered to 0–4 °C for 10 min, and then allowed to refold at 30 °C for at least 10 min. $K_t^{0.5}$ was calculated by fitting the data to Eq. 1 (see "Equations"), where *a* is an exponential term, *t* is the temperature, and $K_t^{0.5}$ is the temperature where *p* is half of *P*.

Activity versus pH profile

The effect of pH on V_{max} for 4NPX hydrolysis was determined using buffers containing 1 mg/ml BSA, with 100 mM succinate being used for pH 3–7, 100 mM phosphate for pH 6.5–8, and 100 mM pyrophosphate for pH 7.5–9. Endpoint assays were performed in quadruplicate using 32.5 nM enzyme, 6.4 mM 4NPX; the reactions were induced at 25 °C for 15 min, and then quenched by adding an equal volume of Na_2CO_3 before spectrophotometric analysis at 400 nm.

Xylobiose hydrolysis analysis and kinetics

Xylobiose (X2) was obtained from Wako Chemicals (Richmond, VA). All other reagents were of high quality. Water was purified by a Milli-Q Academic A10 unit (Millipore; Billerica, MA). A DX500 Dionex HPLC system equipped with an ED40 electrochemical detector (pulsed amperometry) was used for oligosaccharide separation and detection (Dionex; Sunnyvale, CA). Reactions used to calculate k_{cat} and K_m values contained varied concentrations of X2 (1.20–48.4 mM) in 100 mM sodium succinate, pH 6.0 (ionic strength = 0.3 M), in a final volume of 0.3 mL, and were at 25 °C for all enzymes tested, except for XylBH43 and L186K. These two were also tested at elevated temperatures corresponding to their respective $K_t^{0.5}$ values of 51 and 59 °C. Reactions were initiated with 7 μl of appropriately diluted enzyme. Progress was monitored discontinuously by quenching 150 μl of the reaction with 150 μl 200 mM sodium phosphate buffer, with the pH adjusted so that the final quenched reaction was pH ~10.7, followed by HPLC analysis. The time zero reaction mixtures contained no enzymes. Product analysis was performed on a Dionex HPLC. Separation of the products and reactant was achieved using a CarboPac PA-1 column with a separation step (15 mM NaOH), followed by a column washing step (250 mM sodium acetate in 100 mM NaOH) and a column reequilibration step (15 mM NaOH). Data were fitted to Eq. 2. Reactions used to calculate K_{ss}^{X2} values were performed as described previously, where an enzyme-coupled spectrophotometric assay is used for xylose quantitation [26]. The reaction conditions were 25 °C in 100 mM phosphate buffer, with a pH of 6.5, 1 mg/ml BSA, and 1.25–1.75 nM enzyme, and contained eight different X2 concentrations ranging from 1.88–120 mM in a final volume of 30 μL. Data were fitted to Eq. 3.

Enzyme kinetics, and inhibition constants using 4NPX

Michaelis–Menten kinetic parameters were obtained by continuous monitoring of hydrolyzed nitrophenyl chromophores, as previously described [25], using a reaction buffer

containing 100 mM sodium phosphate, 0.1 % BSA, pH 6.5 and 25 °C, and 4NPX concentrations of 80–9,600 μM. Monosaccharide inhibition constants were determined, as previously described [7], using 4NPX and xylose concentrations of 0, 75, 150, and 300 mM, and the data were fitted to either Eq. 4 or Eq. 5.

Equations

$$p = \frac{P}{1 + \left(\frac{t}{K_t^{0.5}}\right)^a} \quad (1)$$

$$\frac{v}{E_T} = \frac{k_{cat} \times S}{K_m + S} \quad (2)$$

$$\frac{v}{E_T} = \frac{k_{cat} \times S}{K_m + S\left(1 + \frac{S}{K_{ss}}\right)} \quad (3)$$

$$\frac{v}{E_T} = \frac{k_{cat} \times S}{K_m\left(1 + \frac{I}{K_i}\right) + S\left(1 + \frac{S}{K_{ss}}\right)} \quad (4)$$

$$\frac{v}{E_T} = \frac{k_{cat} \times S}{K_m\left(1 + \frac{I}{K_i}\right) + S} \quad (5)$$

Data were fitted to equations using either the computer program Grafit (Erithacus Software Ltd., Surrey, UK) or GraphPad Prism 5 (GraphPad Software, LaJolla, CA, USA). Symbol definitions are the following: v is the observed initial (steady-state) rate of catalysis; E_T is the total enzyme concentration; k_{cat} is the turnover number of catalysis; S is the substrate concentration; K_m is the Michaelis constant; I is the inhibitor concentration; K_i is the dissociation constant for I from EI; K_{ss} is the dissociation constant for S from ESS (substrate inhibition/activation); p is the determined parameter at a single temperature; P is the temperature-independent value of the parameter; t is the temperature; $K_t^{0.5}$ is the temperature where p is half of P; a is an exponential term.

Results

Shuffled library construction and screening

The native DNA sequences of *SXA* and *XylBH43* exhibit 56 % identity, while family shuffling [4] is typically performed using genes with >60 % identity to increase the cross-over frequency during recombination [35]. Thus, to maximize overall DNA homology and thereby increase recombination events [24], *SXA* was optimized for shuffling with *XylBH43* by aligning the DNA sequences and changing the individual *SXA* codons to match *XylBH43* codons, with the constraint that the amino acid sequence must be retained. Using this strategy, the overall DNA sequence identity with *XylBH43* increased by 17 %, from 56 to 73 % for optimized *SXA* (Supplementary Fig. 1) used to generate a DNA family shuffled library. About 3,840 library members were initially winnowed using an in vivo, active/inactive library purification 1° screen, wherein genotype and phenotype remain conveniently linked [28]. This screen selects for a combination of sufficient protein expression coupled with enzyme activity acting on MUX; in our study, it resulted in the identification of ~1,300 active clones (~33 % active) that were then rearrayed into 96-well hit plates. After the thermal challenge, 10 clones were selected for hit verification entailing protein expression and purification, followed by determination of $K_t^{0.5}$. Hit verification confirmed five XylBH43 variants (4B9, 2B10, 5F1, 6F11, and 10C4) with $\Delta K_t^{0.5}$ values that were 3.0 ± 0.3 to 5.6 ± 0.4 °C higher than XylBH43, and containing from 5–11 amino acid changes compared to XylBH43 (Table 1). These variants comprise 15 different amino acid substitutions in total from SXA into XylBH43, resulting from a single round of gene shuffling.

Thermal stability $K_t^{0.5}$ determinations

Site-directed mutagenesis was used to create the corresponding single mutation species for all 15 mutations in Table 1, except S54A and S55H, which were tested as the

Table 1 Thermal stability hit amino acid changes and L186Y additivity

Shuffled hit/$\Delta K_t^{0.5}$ (°C)[a]	XylBH43 amino acid changes	Addition of L186Y mutant/$\Delta K_t^{0.5}$ (°C)[a]
4B9/3.0 ± 0.3	5: V22A; C78D; H82A; T85K; Y87W	4B9:L186Y/6.8 ± 0.1
2B10/4.6 ± 0.6	7: V22A; M69P; N70D; C78D; H82A; T85K; Y87W	2B10:L186Y/8.8 ± 0.5
5F1/4.7 ± 0.1	11: V22A; S54A; S55H; T58S; M69P; N70D; C78D; H82A; T85K; Y87W; H419N	5F1:L186Y/7.7 ± 0.01
6F11/2.8 ± 0.2	6: V22A; R45K; K48V; C78D; Q185A; L186Y	Not applicable
10C4/5.6 ± 0.4	6: M69P; N70D; H82A; T85K; Q185A; L186Y	Not applicable

[a] XylBH43 $K_t^{0.5}$ = 51.0 ± 0.3 °C

Table 2 Site-directed mutant thermal stability

Site-directed mutant	$\Delta K_t^{0.5}$ (°C)[a]	Site-directed mutant	$\Delta K_t^{0.5}$ (°C)[a]
V22A	0.6 ± 0.3	C78D	0.0 ± 0.6
R45K	−1.8 ± 0.1	H82A	−0.5 ± 0.2
K48V	−0.8 ± 0.1	T85K	0.6 ± 0.1
S54A/S55H	−0.3 ± 0.2	Y87W	−0.3 ± 0.5
T58S	0.2 ± 0.1	Q185A	−0.1 ± 0.5
M69P	1.7 ± 0.3	L186Y	3.2 ± 0.4
N70D	−0.4 ± 0.3	H419N	0.5 ± 0.2
M69P/N70D	1.9 ± 0.4		

[a] XylBH43 $K_t^{0.5}$ = 51.0 ± 0.3 °C

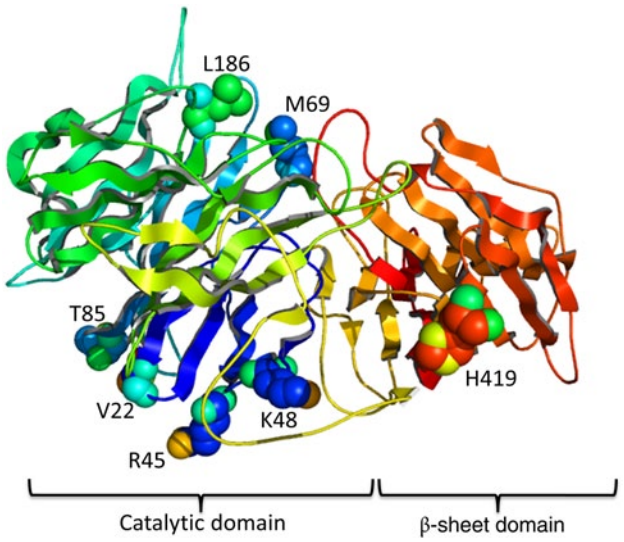

Fig. 1 Location of mutations affecting $\Delta K_t^{0.5}$, with structure representation created using PyMol (Schrödinger LLC)

double mutant (Table 2). When tested singly, three mutations significantly affected thermal stability compared to wild-type XylBH43 ($K_t^{0.5}$ = 51.0 ± 0.3 °C): M69P $\Delta K_t^{0.5}$ = 1.7 ± 0.3 °C, L186Y $\Delta K_t^{0.5}$ = 3.2 ± 0.4 °C, and R45K $\Delta K_t^{0.5}$ = −1.8 ± 0.1. The L186Y mutation, having the largest individual positive $\Delta K_t^{0.5}$ value, was present in the shuffled hits 6F11 and 10C4. Shuffled hit 4B9 contained the fewest number of mutations (five), comprised largely of a cluster of four mutations (C78D, H82A, T85K, and Y87W) near a two-residue (DG) hairpin loop, termed loop-D83 (Fig. 1), that in combination appeared to increase $K_t^{0.5}$ by ~3 °C, but when tested singly did not significantly affect $K_t^{0.5}$. Shuffled hits 2B10 and 5F1 had in common with 4B9 the loop-D83 mutations, and additionally contained the M69P mutation. Site-directed mutagenesis creation of 4B9:L186Y, 2B10:L186Y, and 5F1:L186Y (Table 1) for additivity analysis resulted in $K_t^{0.5}$ increases

Table 3 L186 saturation mutagenesis thermal stability

L186 mutant	$\Delta K_t^{0.5}$ (°C)[a]	L186 mutant	$\Delta K_t^{0.5}$ (°C)[a]
Lys (K)	8.1 ± 0.3	Tyr (Y)	3.2 ± 0.4
Gln (Q)	6.7 ± 0.2	Gly (G)	3.2 ± 0.2
Asn (N)	5.9 ± 0.2	Ala (A)	2.2 ± 0.2
His (H)	5.7 ± 0.3	Thr (T)	1.7 ± 0.2
Cys (C)	5.2 ± 0.2	Trp (W)	1.2 ± 0.2
Ser (S)	5.0 ± 0.2	Leu (L)	≡0
Glu (E)	4.4 ± 0.1	Val (V)	−0.4 ± 0.2
Phe (F)	4.4 ± 0.2	Asp (D)	−0.9 ± 0.2
Arg (R)	4.4 ± 0.3	Pro (P)	−1.5 ± 0.2
Met (M)	4.2 ± 0.1	Ile (I)	−1.9 ± 0.1

[a] XylBH43 $K_t^{0.5}$ = 51.0 ± 0.3 °C

of 3.8 ± 0.3, 4.2 ± 0.8, and 3.0 ± 0.1 °C, respectively, due to the addition of the L186Y mutation compared to a $K_t^{0.5}$ increase of 3.2 ± 0.4 °C for L186Y in isolation.

Saturation mutagenesis of L186 resulted in all variants being active (Table 3). Most L186 changes increased $K_t^{0.5}$, with L186K showing the largest increase (8.1 ± 0.3 °C). Four mutations lowered $K_t^{0.5}$, with reductions ranging from −0.42 ± 0.2 °C (L186V) to −1.9 ± 0.1 °C (L186I). The Lys mutation was introduced into the corresponding position of the parental enzyme SXA (SXA:Y184K), resulting in reduced thermal stability; $\Delta K_t^{0.5}$ = −2.4 ± 0.1 °C (data not shown).

pH curves

Concerning the pH optima of 4NPX hydrolysis, the acidic limbs of the pH curves of the XylBH43 variants are shifted ~0.5–1.0 unit closer to the acidic side relative to XylBH43 (Fig. 2). In particular, the mutants 10C4 (curve not shown), L186E (curve not shown), and L186R maintain >80 % maximal activity at pH 5, whereas XylBH43 activity is <60 % of the maximal activity at this pH. L186R also has a particularly broad pH optimum (~2 pH units) from ~pH 5.5–7.5 (Fig. 2). The basic limbs of five of the mutants (L186V, D, P, I, and T) are significantly shifted toward the acidic side (Fig. 2, circled). They have <40 % relative activities at pH 8, and all of these except L186T are also the only variants with $K_t^{0.5}$ values lower than wild-type L186 (Table 3).

Kinetic parameters acting on X2

The thermal stability hits (Table 1), the L186Y mutant (Table 2), the L186Y additivity constructs (Table 1), and the top six most thermally stable L186 site-directed mutants (Table 3) were evaluated for X2 catalysis at pH 6 and 25 °C (Table 4). The determined k_{cat}^{X2} values were

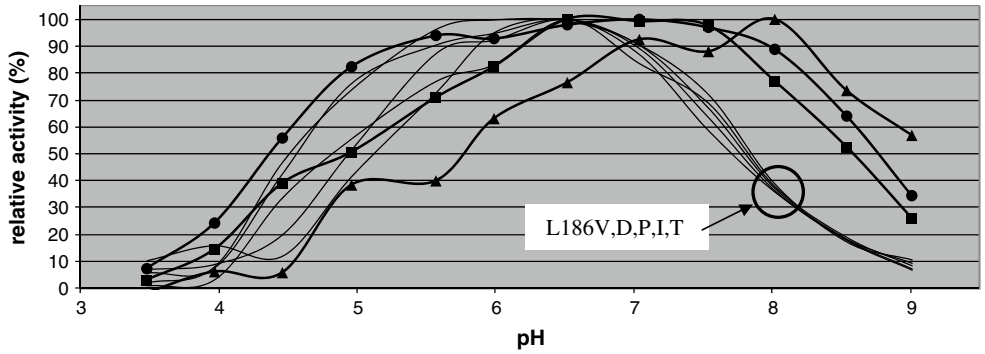

Fig. 2 pH curves at 25 °C for 4NPX hydrolysis by XylBH43 and select variants. *Filled square* xylBH43, *filled triangle* 2B10, *filled circle* L186R

Table 4 Kinetic parameters for catalyzed reactions acting on X2

Enzyme	k_{cat}^{X2} (s^{-1})[a]	k_{cat}^{X2}/K_m^{X2} (s^{-1} mM^{-1})[a]	K_m^{X2} (mM)[a]	K_{ss}^{X2} (mM)[b]
XylBH43 (25 °C)	91.9 ± 1.4	17.3 ± 0.3	5.3 ± 0.2	130 ± 10
XylBH43[c] (51 °C)	267 ± 4.1	30.7 ± 1.1	8.7 ± 0.4	ND
L186K (25 °C)	54.1 ± 0.9	6.5 ± 0.2	8.3 ± 0.3	700 ± 160
L186K[d] (59 °C)	262 ± 4.6	11.4 ± 0.3	23 ± 0.9	ND
L186Q	15.8 ± 0.2	6.0 ± 0.3	2.6 ± 0.1	ND
L186N	73.4 ± 1.3	3.1 ± 0.1	23.5 ± 0.9	ND
L186H	49.7 ± 0.6	1.2 ± 0.01	40.3 ± 0.8	360 ± 470
L186C	73.4 ± 1.4	9.5 ± 0.3	7.8 ± 0.4	ND
L186S	60.1 ± 0.6	6.7 ± 0.2	9.0 ± 0.3	ND
L186Y	68.8 ± 0.7	8.9 ± 0.2	7.7 ± 0.2	330 ± 30
5F1	73.6 ± 05	10.8 ± 0.2	6.8 ± 0.1	260 ± 60
5F1:L186Y	64.2 ± 1.7	5.5 ± 0.3	11.7 ± 0.8	ND
2B10	73.5 ± 0.7	10.3 ± 0.2	7.2 ± 0.2	140 ± 30
2B10:L186Y	56.4 ± 0.2	5.6 ± 0.04	10.2 ± 0.1	ND
4B9	84.2 ± 0.5	14.2 ± 0.2	5.9 ± 0.1	200 ± 20
4B9:L186Y	72.6 ± 0.8	7.6 ± 0.2	9.6 ± 0.3	ND
6F11	60.4 ± 0.7	5.9 ± 0.1	10.2 ± 0.3	470 ± 80
10C4	54.2 ± 0.3	5.4 ± 0.1	10.0 ± 0.1	320 ± 170

ND not determined

[a] Reactions were in 100 mM sodium succinate, pH 6.0, $I = 0.3$ M, 25 °C; the analysis was performed using HPAEC/PAD, and the data were fitted using Eq. 2

[b] Reactions were in 100 mM sodium phosphate, pH 6.5, 25 °C; the analysis was performed using a spectrophotometric enzyme-coupled assay, and the data were fitted using Eq. 3

[c] Reactions were in 100 mM sodium succinate, pH 6.0, $I = 0.3$ M, 51 °C; the analysis was performed using HPAEC/PAD, and the data were fitted using Eq. 2

[d] Reactions were in 100 mM sodium succinate, pH 6.0, $I = 0.3$ M, 59 °C; the analysis was performed using HPAEC/PAD, and the data were fitted using Eq. 2

seen to decrease by a factor of 0.17–0.92, while the K_m^{X2} values increased for all variants tested except L186Q, for which K_m^{X2} decreased by ~½; the L186Q variant also had the largest observed k_{cat}^{X2} decrease of 83 %. The determined K_{ss}^{X2} parameters (substrate inhibition) were increased for all variants except 2B10. For variant L186H, the lack of appreciable substrate inhibition at the highest X2 concentrations tested (Fig. 4b) led to poor fitting of this parameter using Eq. 3 ($K_{ss}^{X2} = 361 \pm 467$ mM).

For XylBH43, the k_{cat}^{X2} increased 2.9-fold to 267 ± 4 s^{-1} when determined at its $K_t^{0.5}$ (51 °C), compared to at 25 °C, and K_m^{X2} increased 1.6-fold to 8.7 ± 0.4 mM, resulting in k_{cat}^{X2}/K_m^{X2} nearly doubling to 30.7 ± 1.1 s^{-1} mM^{-1} (Table 4). For mutant L186K, the k_{cat}^{X2} determined at its $K_t^{0.5}$ temp (59 °C), instead of at 25 °C, increased 4.9-fold to 262 ± 5 s^{-1}, and K_m^{X2} increased 2.8-fold to 23 ± 0.9 mM, such that k_{cat}^{X2}/K_m^{X2} also nearly doubled to 11.4 ± 0.3 s^{-1} mM^{-1} (Table 4).

Kinetic parameters acting on 4NPX

Kinetic parameters of 4NPX hydrolysis and inhibition by xylose were determined for L186H, L186K and L186Y (Table 5; Fig. 4a). The k_{cat}^{4NPX}values ranged from 25 to 79 % of those seen for XylBH43. The K_m^{4NPX}, K_i^{xylose}, and K_{ss}^{4NPX} values are larger for all variants, up to 2-fold for K_m^{4NPX} and K_i^{xylose}. The K_{ss}^{4NPX}values are larger by ~3-fold for L186Y and L186K, and no substrate inhibition was observed for L186H at the highest 4NPX concentration employed.

Discussion

In gene shuffling of closely-related enzymes, multiple mutations can be tolerated, leading to relatively high library quality [4, 35]. Indeed, one of the selectants (5F1)

Table 5 Kinetic parameters of XylBH43 wild-type and selected mutant-catalyzed reactions acting on 4NPX

Enzyme	k_{cat}^{4NPX} (s^{-1})	$k_{cat}^{4NPX}/K_m^{4NPX}$ (s^{-1} mM^{-1})	K_m^{4NPX} (mM)	K_i^{xylose} (mM)	K_{ss}^{4NPX} (mM)
XylBH43	17.9 ± 1.3	61.7 ± 13.5	0.3 ± 0.1	30.2 ± 5.0	11.5 ± 2.4
L186Y	14.1 ± 0.9	34.4 ± 6.3	0.4 ± 0.1	38.5 ± 5.1	31.4 ± 9.6
L186K	6.4 ± 0.5	9.8 ± 2.0	0.7 ± 0.1	59.3 ± 9.0	31.5 ± 12.3
L186H	4.4 ± 0.1	9.3 ± 0.9	0.5 ± 0.1	58.0 ± 7.2	N/A[a]

Reactions were in 100 mM sodium phosphate, pH 6.5, 0.1 % BSA, 25 °C; data were fitted using Eq. 4

[a] N/A = not applicable to the 4NPX concentrations tested; data were fitted using Eq. 5

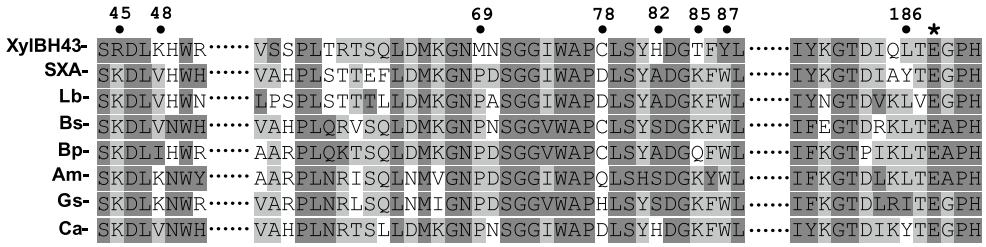

Fig. 3 Sequence alignment of XylBH43 and high-activity GH43 xylosidase homologs. *Filled circle* location of amino acid change in shuffled hits; *asterisk* active site catalytic Glu residue; *XylBH43* PDB ID 1YRZ; *SXA* PDB ID 3C2U; *Lb Lactobacillus brevis* ATCC 367, GenBank ABJ65333.1; *Bs Bacillus subtilis* 168, GenBank CAB13642.2; *Bp Bacillus pumilus*, GenBank CAA29235.1; *Gs Geobacillus stearothermophilus*, PDB ID 2EXH; *Am Alkaliphilus metalliredigens*, GenBank ABR49445.1; *Ca Clostridium acetobutylicum*, GenBank AAK81382

contained numerous (11) amino acid mutations. Screening a small, ~3,840 member library in a single round of family shuffling identified three individual mutations (R45K, M69P, L186Y) (Table 2) and a cluster of four mutations near loop-D83 that significantly affected thermal stability. Other recent examples of GH43-enzyme engineering are broadening the substrate acceptance of an arabinofuranosidase to include *endo*-xylanase activity [19], and lowering inhibitor affinities of β-xylosidase SXA [7, 15].

The domain architecture of GH 43 enzymes consist mainly of two types: either a single, α/β barrel catalytic domain, or an N-terminal catalytic domain combined to a similar-sized C-terminal β-sandwich domain with carbohydrate binding properties [34]. XylBH43 and SXA are of the latter type, and all of the mutations in the hits occurred in the catalytic domain, except H419N on the β-sandwich non-catalytic domain (Fig. 1). Figure 3 shows mutation locations and the amino acid alignment for XylBH43, SXA, and six other closely-related GH43 enzymes (45–55 % identity), consisting of the most efficient GH43 xylosidases reported to date for xylooligosaccharide hydrolysis, with k_{cat}^{X2} values ranging from 40 s^{-1} for the *Bacillus pumilus* enzyme (Bp), to the highest reported k_{cat}^{X2} of 420 s^{-1} for the *Lactobacillus brevis* enzyme (Lb) at 25 °C [16]. Interestingly, the mutations occurred mainly in loop regions or in β-sheets immediately before and after loop regions, suggesting that increasing loop rigidity may be a common underlying mechanism for increased thermal stability. Structural rigidity can confer thermal stability [1, 11], and amino acid *B*-values (organized by the XylBH43 crystal structure data from PDB 1YRZ for each amino acid using the program B-fitter [21]; see Supplementary Table 2) are a measure of electron density smearing that have been used successfully to identify regions for thermal stability improvement [12, 21, 22]. Thermally destabilizing mutations R45K ($\Delta K_t^{0.5} = -1.8 \pm 0.1$ °C) and K48V ($\Delta K_t^{0.5} = -0.8 \pm 0.1$ °C) are on a six-residue loop (R45-W50; Fig. 1), separating two β-strands. R45 ($B = 33.67$, 4th highest Arg *B*-value of 23 Arg residues having an average *B*-value = 24.52) and K48 ($B = 32.74$, 6th highest Lys *B*-value of 27 Lys residues having an average *B*-value = 27.52) have the largest B-factors of the shuffled mutations, ranking 41 and 37 out of 522 amino acids, respectively. M69 is located 15 Å from the active site carboxylates on the longest (24 residue) extended loop (S54-P77; Fig. 1), and XylBH43 is the only highly efficient GH43 xylosidase reported that contains Met at this position, otherwise occupied by Pro (Fig. 3). Stabilization has previously been shown by Pro substitution into loop regions, by making the loop more rigid and lowering the number of conformations and thus entropy of the unfolded state [1, 18, 29].

The thermal stability increases observed for hits 4B9, 2B10, and 5F1, containing the cluster of mutations C78D, H82A, T85K, and Y87W (Table 1), were not ascribable to single amino acid changes, as tested by variants in Table 2

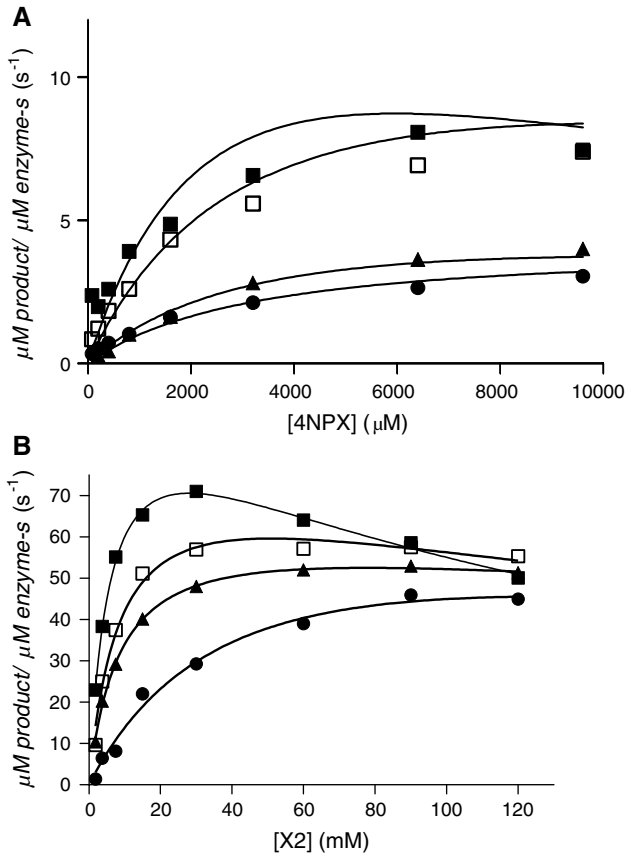

Fig. 4 Xylose and xylobiose inhibition of XylBH43 and select variants. *Filled square* wild-type; *open square* L186Y; *filled triangle* L186K; *filled circle* L186H. **a** 4NPX hydrolysis in the presence of 300 mM xylose; **b** X2 hydrolysis; reaction conditions were 100 mM phosphate, pH 6.5, 0.1 % BSA, 25 °C

where only T85K modestly improved $K_t^{0.5}$. However, together these four mutations appear to increase $K_t^{0.5}$ by ~3 °C. H82A and T85K comprise the last and first amino acids, respectively, of β-strands having an intervening two-residue (D83G84) hairpin loop, termed loop-D83 (Figs. 1, 3), and the increased thermal stability may have arisen from stabilizing this loop. Recently, the $K_t^{0.5}$ of a polyol dehydrogenase was increased by 7 °C in a chimera created by exchanging a loop region from a less thermostable homologous enzyme [33]. Residues H82A and T85K have moderately-high B-factors (ranked 71/522 and 73/522, respectively), while loop residue D83 has a high B-factor, ranked 17 out of 522 amino acids. Hit 4B9 is of particular interest for further thermal improvement, since the kinetic parameters determined for X2 were little affected for this variant (Table 4). Variants 2B10 and 5F1 have in common the additional mutations M69P and N70D, as compared to 4B9, and the observed $K_t^{0.5}$ increase of ~1.7 °C for these two hits compared to 4B9 is consistent with M69P simple additivity.

Mutation L186Y in the selected hits (Table 1) occurs on a 12-residue, active-site loop near catalytic general acid Glu188 (Fig. 1). Most 10C4 stabilization is likely due to the L186Y and M69P mutations, and assuming the stability conferred by all the mutations in 10C4 are simply additive, their sum of $\Delta K_t^{0.5} = 4.5 \pm 0.8$ °C can account for the observed $\Delta K_t^{0.5}$of 5.6 ± 0.4 °C. Most 6F11 stabilization appears due to the L186Y mutation, though the stability is somewhat more than expected based on L186Y alone, since 6F11 also contains the destabilizing mutations R45K and K48V. Saturation mutagenesis [31] of L186 resulted in all but four mutants with $K_t^{0.5}$ values greater than that of XylBH43, with the largest effect seen in L186K, having an $\Delta K_t^{0.5} = 8.1 \pm 0.3$ °C (Table 3). The effect of amino acid mutation is, in general, context-dependent [1], and when Lys was introduced into the corresponding position of the parental enzyme SXA, a decrease in $K_t^{0.5}$ of 2.4 °C was observed. The breadth of the effect on the $K_t^{0.5}$ of the L186 saturation mutagenesis, ranging from −1.9 to 8.1 °C, supports the saturation mutagenesis of other residues that appear to slightly perturb $K_t^{0.5}$, such as V22A, K48V, H82A, T85K, and H419N (Table 2).

Combining several mutations, each of which confers an incremental improvement in thermal stability, can be a way to achieve significant gains in protein thermal stability [17, 18]. The L186Y mutation alone increased $K_t^{0.5}$ by 3.2 ± 0.4 °C, and its additivity (within experimental error) to hits 4B9, 2B10, and 5F11 was demonstrated, with $K_t^{0.5}$ increasing by 3.8 ± 0.3, 4.2 ± 0.8, and 3.0 ± 0.1 °C, respectively, thereby generating variants with combined $K_t^{0.5}$ increases of 6.8 ± 0.1 to 8.8 ± 0.5 °C (Table 1). Simple additivity toward thermal stability can be anticipated for mutations not in direct contact [23, 30]; for example, an early directed-evolution thermal stability study showed eight of nine mutations individually contributing to the thermal stability of an esterase [10].

Kinetic parameter k_{cat}^{X2} decreased in the mutant xylosidases, ranging from a minimal change for hit 4B9 (8 % decrease, $\Delta K_t^{0.5} = 3.0$ °C) to a significant decrease for L186Q (80 % decrease, $\Delta K_t^{0.5} = 6.7$ °C), and parameters K_m^{X2}, K_m, K_i^{xylose} and K_{ss}^{X2} increased varyingly and in tandem for the mutants except L186Q, where K_m^{X2} was decreased by ~1/2 (Tables 4, 5). While in general, a higher $K_t^{0.5}$ would allow thermal compatibility with other thermally stable enzymes in a cocktail, the effects of concurrent changes in k_{cat}, K_m, K_i, and K_{ss} during selection on the efficiency of industrial saccharification are process-specific. The effect of a lowered k_{cat} would be mitigated under saccharification reactor conditions owing to the increased turnover of the thermally stabilized enzymes at higher temperatures. This was demonstrated here, where XylBH43 and L186Kwere tested at their respective $K_t^{0.5}$ values for turnover efficiency of natural substrate X2, and the determined k_{cat}^{X2} values were

similarly ~260 s^{-1} (Table 4). GH43 xylosidases are considered attractive candidates for enzyme cocktails since catalysis occurs without an enzyme-bound intermediate that is susceptible at high substrate loadings to transferase-type activity (e.g., the transxylosylation of X2 yielding xylotriose [14]), conditions that may occur in commercial saccharification. However, GH43 xylosidases are prone to substrate and product inhibition. As a result, whereas an increased K_m is deleterious for processes requiring high turnover at low substrate concentrations, the observed concurrent deterioration of K_m can both be ameliorated at higher substrate concentrations and be a positive attribute owing to the in-tandem decrease in product (K_i) and substrate (K_{ss}) inhibition. The effect of increased K_i on 4NPX hydrolysis was tested at 300 mM xylose and resulted in the convergence of turnovers for XylBH43 and L186Y (Table 5; Fig. 4a), while the effect of increased K_{ss} was tested at up to 120 mM natural substrate X2, wherein the turnovers of XylBH43 and L186Y, K, and H converged (Table 4; Fig. 4b).

Acknowledgments This work was supported by funding from the United States Department of Agriculture, CRIS 5325-41000-049-00 (C.H., V.J.C., C.C.L., K.W.) and CRIS 3620-41000-118-00D (D.B.J. and J.D.B.). The mention of firm names or trade products does not imply that they are endorsed or recommended by the US Department of Agriculture over other firms or similar products not mentioned. The USDA is an equal opportunity provider and employer.

References

1. Arnold FH, Wintrode PL, Miyazaki K, Gershenson A (2001) How enzymes adapt: lessons from directed evolution. Trends Biochem Sci 26:100–106
2. Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. Proc Natl Acad Sci 103:5869–5874
3. Cantarel B, Coutinho P, Rancurel C, Bernard T, Lombard V, Henrissat B (2008) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37:D233–D238
4. Crameri A, Raillard S-A, Bermudez E, Stemmer WPC (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391:288–291
5. Ebringerová A, Hromádková Z, Heinze T (2005) Hemicellulose. Adv Polym Sci 186:1–67
6. Eijsink VGH, Gåseidnes S, Borchert TV, Van den Burg B(2005) Directed evolution of enzyme stability. Biomol Eng 22:21–30
7. Fan Z, Yuan L, Jordan DB, Wagschal K, Heng C, Braker JD (2010) Engineering lower inhibitor affinities in β-xylosidase. Appl Microbiol Biotechnol 86:1099–1113
8. Gao D, Uppugundla N, Chundawat SP, Yu X, Hermanson S, Gowda K, Brumm P, Mead D, Balan V, Dale BE (2011) Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. Biotechnol Biofuels 4:5
9. Gilbert HJ (2010) The biochemistry and structural biology of plant cell wall deconstruction. Plant Physiol 153:444–455
10. Giver L, Gershenson A, Freskgard P-O, Arnold FH (1998) Directed evolution of a thermostable esterase. Proc Natl Acad Sci 95:12809–12813
11. Hoseki J, Okamoto A, Takada N, Suenaga A, Futatsugi N, Konagaya A, Taiji M, Yano T, Kuramitsu S, Kagamiyama H (2003) Increased rigidity of domain structures enhances the stability of a mutant enzyme created by directed evolution. Biochemistry 42(49):14469–14475
12. Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focused directed evolution. Protein Eng Des Sel 23(12):903–909
13. Jordan DB, Li X-L, Dunlap CA, Whitehead TR, Cotta MA (2007) Structure-function relationships of a catalytically efficient β-D-xylosidase. Appl Biochem Biotechnol 141:51–76
14. Jordan DB, Wagschal K (2010) Properties and applications of microbial β-D-xylosidases featuring the catalytically efficient enzyme from *Selenomonas ruminantium*. Appl Microbiol Biotechnol 86:1647–1658
15. Jordan DB, Wagschal K, Fan Z, Yuan L, Braker JD, Heng C (2011) Engineering lower inhibitor affinities in β-D-xylosidase of *Selenomonas ruminantium* by site-directed mutagenesis of Trp145. J Ind Microbiol Biotechnol 38:1821–1835
16. Jordan DB, Wagschal K, Grigorescu AA, Braker JD (2012) Highly active β-xylosidases of glycoside hydrolase family 43 operating on natural and artificial substrates. Appl Microbiol Biotechnol. 97: 4415–4428. doi:10.1007/s00253-00012-04475-00254
17. Matsumura M, Yasumura S, Aiba S (1986) Cumulative effect of intragenic amino-acid replacements on the thermostability of a protein. Nature 323:356–358
18. Matthews BW, Nicholson H, Becktel WJ (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc Natl Acad Sci 84:6663–6667
19. McKee LS, Peñab MJ, Rogowskia A, Jackson A, Lewis RJ, York WS, Krogh KBRM, Viksø-Nielsen A, Skjøt M, Gilbert HJ, Marles-Wright J (2012) Introducing endo-xylanase activity into an exo-acting arabinofuranosidase that targets side chains. Proc Natl Acad Sci 109:6537–6542
20. Pattanaik S, Werkman J, Kong Q, Yuan L (2010) Site-directed mutagenesis and saturation mutagenesis for the functional study of transcription factors involved in plant secondary metabolite biosynthesis. Methods Mol Biol 643:47–57
21. Reetz MT, Carballeira JD (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. Nat Protoc 2(4):891–903
22. Reetz MT, Carballeira JD, Vogel A (2006) Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angew Chem Int Ed 45:7745–7751
23. Shortle D, Meeker AK (1986) Mutant forms of *Staphylococcal nuclease* with altered patterns of guanidine hydrochloride and urea denaturation. Proteins 1:81–89
24. Stemmer WPC (1994) DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. Proc Natl Acad Sci 91:10747–10751
25. Wagschal K, Franqui-Espiet D, Lee CC, Kibblewhite-Accinelli RE, Robertson GH, Wong DWS (2007) Genetic and biochemical characterization of an α-L-arabinofuranosidase isolated from a compost starter mixture. Enzyme Microb Technol 40:747–753
26. Wagschal K, Franqui-Espiet D, Lee CC, Robertson GH, Wong DWS (2005) Enzyme-coupled assay for β-xylosidase hydrolysis of natural substrates. Appl Environ Microbiol 71(9):5318–5323
27. Wagschal K, Jordan DB, Braker JD (2012) Catalytic properties of β-D-xylosidase XylBH43 from *Bacillus halodurans* C-125 and mutant XylBH43-W147G. Process Biochem 47:366–372
28. Wagschal K, Lee CC (2012) Microplate-bases active/inactive 1 screen for biomass degrading enzyme library purification and gene discovery. J Microbiol Methods 89:83–85
29. Watanabe K, Masuda T, Ohashi H, Mihara H, Suzuki Y (1994) Multiple proline substitutions cumulatively thermostabilize

Bacillus cereus ATCC7064 oligo-1,6-glucosidase: Irrefragable proof supporting the proline rule. Eur J Biochem 226:277–283
30. Wells JA (1990) Additivity of mutational effects in proteins. Biochemistry 29(37):8509–8517
31. Wells JA, Vasser M, Powers DB (1985) Cassette mutagenesis: an efficient method for generating multiple mutations at defined sites. Gene 34:315–323
32. Wolfenden R, Snider MJ (2001) The depth of chemical time and the power of enzymes as catalysts. Acc Chem Res 34:938–945
33. Wulf H, Mallin H, Bornscheuer UT (2012) Protein engineering of a thermostable polyol dehydrogenase. Enzyme Microb Technol 51(4):217–224
34. Yoshida S, Hespen CW, Beverly RL, Mackie RI, Cann IKO (2010) Domain analysis of a modular α-L-arabinofuranosidase with a unique carbohydrate binding strategy from the fiber-degrading bacterium *Fibrobacter succinogenes* S85. J Bacteriol 192(20):5424–5436
35. Yuan L, Kurek I, English J, Keenan R (2005) Laboratory-directed protein evolution. Microbiol Mol Biol Rev 69:373–392
36. Zhao H, Arnold FH (1997) Optimization of DNA shuffling for high fidelity recombination. Nucleic Acids Res 25(6):1307–1308