Thermostability enhancement of an endo-1,4-β-galactanase from Talaromyces stipitatus by site-directed mutagenesis

ARTICLE *in* APPLIED MICROBIOLOGY AND BIOTECHNOLOGY · DECEMBER 2014

Impact Factor: 3.34 · DOI: 10.1007/s00253-014-6244-z

7 AUTHORS, INCLUDING:

Dorte møller larsen
Danish Veterinary and Food Administration
9 PUBLICATIONS **99** CITATIONS
SEE PROFILE

Anders Thygesen
Technical University of Denmark
44 PUBLICATIONS **939** CITATIONS
SEE PROFILE

Anne S Meyer
Technical University of Denmark
264 PUBLICATIONS **8,362** CITATIONS
SEE PROFILE

Jorn Dalgaard D Mikkelsen
Technical University of Denmark
132 PUBLICATIONS **4,467** CITATIONS
SEE PROFILE

Appl Microbiol Biotechnol
DOI 10.1007/s00253-014-6244-z

Thermostability enhancement of an endo-1,4-β-galactanase from *Talaromyces stipitatus* by site-directed mutagenesis

Dorte M. Larsen · Christian Nyffenegger · Maria M. Swiniarska · Anders Thygesen · Mikael L. Strube · Anne S. Meyer · Jørn D. Mikkelsen

Received: 17 September 2014 / Revised: 14 November 2014 / Accepted: 17 November 2014

Abstract Enzymatic conversion of pectinaceous biomasses such as potato and sugar beet pulp at high temperatures is advantageous as it gives rise to lower substrate viscosity, easier mixing, and increased substrate solubility and lowers the risk of contamination. Such high-temperature processing requires development of thermostable enzymes. *Talaromyces stipitatus* was found to secrete endo-1,4-β-galactanase when grown on sugar beet pectin as sole carbon source. The mature protein contained 353 AA and the MW was estimated to 36.5 kDa. It was subjected to codon optimization and produced in *Pichia pastoris* in 2 l scale yielding 5.3 g. The optimal reaction condition for the endo-1,4-β-galactanase was determined to be 46 °C at pH 4.5 at which the specific activity was estimated to be 6.93 μmol/min/mg enzyme with half-lives of 13 and 2 min at 55 and 60 °C, respectively. For enhancement of the half-life of TSGAL, nine single amino acid residues were selected for site-directed mutagenesis on the basis of semi-rational design. Of these nine mutants, G305A showed half-lives of 114 min at 55 °C and 15 min at 60 °C, respectively. This is 8.6-fold higher than that of the TSGAL at 55 °C, whereas the other mutants displayed moderate positive to negative changes in their half-lives.

Keywords Protein engineering · Semi-rational design · Multiple alignment · GH53 · Half-life

Introduction

A high thermal stability of enzymes is required to allow biocatalytic processing of plant biomass at elevated temperatures. Higher temperatures, i.e., temperatures above 50 °C are preferred for such processing due to the decreased substrate viscosity, improved substrate solubility, and lowered risk of microbial contamination. Elevated processing temperatures, and hence thermostable enzymes, are particularly pertinent in advanced enzymatic processing of pectinaceous plant biomass streams for production of, e.g., bioactive oligosaccharides where high substrate concentrations and efficient mixing are desirable for achieving high enzymatic conversion rates and high product yields (Holck et al. 2011; Thomassen et al. 2011a; Ravn and Meyer 2014). Endo-1,4-β-galactanases (E.C.3.2.1.89) catalyze the hydrolysis of the β-1,4-galactan in type 1 arabinogalactan side chains of "hairy" rhamnogalacturonan I of pectin in an endo-fashion. Endo-1,4-β-galactanases may be employed to release prebiotic compounds from these rhamnogalacturonan I side chains in pectinaceous agroindustrial substrates such as, e.g., potato pulp (Michalak et al. 2012). Endo-1,4-β-galactanases are categorized in glycoside hydrolase family 53, GH53 (Cantarel, et al. 2009) and belong to clan GH-A the so called $(\beta/\alpha)_8$-barrel fold superfamily performing cleavage via the catalytic retaining mechanism mediated by two carboxylic acids located at the ends of the β-strands 4 (acid/base) and 7 (nucleophile), respectively (Naumoff 2011). Some fungal endo-1,4-β-galactanases have been analyzed and crystallized (Ryttersgaard et al. 2002; Le Nours et al. 2003; Otten et al. 2013; Sakamoto et al. 2013), but only two relatively thermostable endo-1,4-β-galactanases from *Myceliophthora*

Electronic supplementary material The online version of this article (doi:10.1007/s00253-014-6244-z) contains supplementary material, which is available to authorized users.

D. M. Larsen · C. Nyffenegger · A. Thygesen · M. L. Strube · A. S. Meyer · J. D. Mikkelsen (✉)
Center for Bioprocess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, 2800 Kongens Lyngby, Denmark
e-mail: jdm@kt.dtu.dk

M. M. Swiniarska
Dako Denmark A/S, Produktionsvej 42, 2600 Glostrup, Denmark

Published online: 02 December 2014

thermophila and *Humicola insolens* both having optimum at 65 °C have been characterized so far (Le Nours et al. 2003). The currently available data do not provide any clues as to which traits of the fungal endo-1,4-β-galactanases that provide for the thermal stability or whether the thermal stability relates to particular sequence motifs, single amino acids, or structural traits of these enzymes.

Native enzymes can be optimized by molecular evolution to become more thermostable and thus able to catalyze processes at elevated temperature regimes (Bornscheuer et al. 2012; Davids et al. 2013). Design of enzyme stability is facilitated by the availability of a 3D structure, as it provides more precise knowledge about the positioning of each amino acid such as distance to the catalytic site and solvent accessibility. Using random mutagenesis, the probability of finding beneficial mutations is extremely low and therefore requires expensive, laborious high-throughput screening methods and selection among a vast number of no-go data. Alternatively, a semi-rational approach with more targeted mutations generally produces a smaller number of mutants (Davids et al. 2013; Lutz 2010), but its success depends on the strategy by which mutations are chosen (Davids et al. 2013; Silva et al. 2013).

The objective of this research was to discover, characterize, and clone a novel gene encoding an endo-1,4-β-galactanase from *Talaromyces stipitatus* (*Penicillium stipitatum*) a mesophile, non-pathogenic, and saprophytic fungus belonging to the Ascomycetes and, in addition, to optimize this endo-1,4-β-galactanase which is potentially suitable for the partial degradation of potato galactan. *T. stipitatus* is known for its ability to secrete thermostable enzymes (Gruber et al. 1998; Li et al. 1998) and is moreover phylogenetically relatively close to *Thermomyces* species (Houbraken et al. 2014). An endo-1,4-β-galactanase expressed by *T. stipitatus* has not been described previously.

The hypothesis behind the work undertaken was that a "funnel" approach involving sequence alignment comparisons and subsequent use of *in silico* methods, namely, prediction of protein mutant stability changes (PoPMuSiC) and the B-factor iterative test (B-FIT), respectively, with additional structural considerations using a 3D homology model to narrow mutation targets, would provide for a feasible strategy to improve thermal stability of a new endo-1,4-β-galactanase from *T. stipitatus*. Sequence alignment was composed of TSGAL, two relatively thermostable endo-1,4-β-galactanases and the well characterized and the crystallized endo-1,4-β-galactanase from *Aspergillus aculeatus* (Ryttersgaard et al. 2002).The PoPMuSiC algorithm is developed to predict the changes in folding free energy (ddG) for all possible single mutations in a protein (Dehouck et al. 2009; Gilis and Rooman 2000; Kwasigroch et al. 2002), whereas the B-FIT test relies on crystallographic B-factor data of the protein flexibility. We here report the discovery of a new endo-1,4-β-galactanase from *T. stipitatus* (TSGAL) by molecular identification and isolation hereof by primer strategies, and the subsequent thermal stability optimization using a new rational "funnel" approach to target a set of nine proposed mutations.

Materials and methods

Fungal cultivation and culture broth preparation

T. stipitatus ATCC 10500 (CBS 375.48) was provided by Technical University of Denmark, Department of Systems Biology, Center for Microbial Biotechnology. With a scalpel, a small plug (approx. 5 mm^2) was transferred to 50 ml liquid media: 3 g/l $NaNO_3$, 1 g/l K_2HPO_4, 0.5 g/l KCl, 0.5 g/l $MgSO_4 \cdot 7H_2O$, 0.01 g $ZnSO_4 \cdot 7H_2O$, 0.05 g $CuSO_4 \cdot 5H_2O$, and 20 g/l high molecular sugar beet pectin provided by DuPont, Nakskov, DK, dissolved in 96 % ethanol (1 ml used for 1 g before autoclaving), 50 ml of H_2O, 5 leca beads (FOG byggemarked, Lyngby, DK) so the surface was almost covered in a 250-ml blue cap flask. The lids were not tightened fully to allow escape of CO_2 and availability of O_2. The flasks were incubated for 5 days in darkness at 25 °C. After incubation, the mycelium biomass was separated from the liquid fraction by centrifugation (10,000 g for 20 min) and frozen at −80 °C for later DNA extractions. Forty milliliters of liquid culture were sterile filtrated through 17 mm and 0.45 μm filter and kept on ice. Relative galactanase activity was measured using azurine cross-linked (AZCL) galactan from potato (Megazyme, Bray, Ireland). AZCL was prepared as detailed in the supplier's procedures.

Identification

The sequence of the endo-1,4-β-galactanase from *A. aculeatus* (Uniprot P48842) (Le Nours et al. 2003) was submitted in blast from http://www.uniprot.org/uniprot/P48842 (Magrane and UniProt 2011) using default settings. Protein sequences (Q8X168, Q9Y7F8, P48842, I8I846, B8NNI2, Q5B153, B5MGR3, B6H1V0, B6QVV5, B8MTD9, Q4WJ80, and B0XRP3) were selected ranging from 67 to 99 % sequence homology, all originated from the Ascomycota division. Sequences including P48842 were submitted to iCODEHOP (Boyce et al. 2009) using default parameters. The following primers were successful in amplification of the gene: forward primers: F1:5′CCCTGACCTA CCGCGGNGCNGAYAT3′; F2:5′GTCGAGATCGTCTCCA THGGNAAYGARA3′. reverse primers: R1:5′GAYT TYGAYHWGATGGGCGTCTCCT; 3′;R2:5′GARA CNRAYTGGCCCGTCACGTGC 3′. PCR reactions were performed in a reaction volume of 25 μL containing 1 × Dream Taq™ Buffer (Thermo Scientific), 0.2 mM of each dNTP, 0.5 μM of each primer, approx. 10 ng DNA, 0.16 mM

$MgCl_2$, 0.8 U of Dream Taq™ DNA Polymerase (Thermo Scientific) and nuclease-free water. R1, F1, R2, and F2 were combined in reactions using the following PCR program: denaturation for 30 s at 95 °C, then 30 cycles consisting of 94 °C for 1 min, 60 °C for 40 s (decreasing by 1 °C/cycle), and 72 °C for 2 min. This was followed by 30 cycles of 96 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 8 min. Elongation time was adapted to the largest expected fragment. Amplification products were analyzed by electrophoresis in a 1 % (*w*/*v*) agarose gel in 1 × TBE (Gibco, Life technologies) and stained with ethidium bromide in order to visualize the DNA bands.

DNA fragment was purified using gel purification (GE Healthcare) and cloned using the CloneJET PCR cloning kit (Thermo Scientific). Positive selected clones were verified as described in the manual. Plasmids were purified using Zyppy Plasmid Miniprep kit (Zymo Research) as described in the manual. Sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany) using the primers pJET1.2_Forward 5′ CGACTCACTATAGGGA GAGCGGC 3′ and PJET1.2_Reverse 5′AAGAACATCG ATTTTCCATGGCAG 3′.

Flanking regions of the gene were identified using genome walking. The known sequence part was used for the design of specific primers, which were then used in combination with a semi-degenerated primer each, as described in Guo and Xiong (2006). Primers for amplification of upstream and downstream regions are summarized in Table S1. All reactions were combined with the semi 2 primer: 5′ GCCTTAAGGCCTANGARMSNCC NAG 3′.

The first PCR was performed with approximately 10 ng of genomic DNA from *T. stipitatus*, whereas the following rounds were performed with 1 μl from the former PCR reaction as template. Each PCR reaction contained the same amount of chemicals as used for CODEHOP PCR (see above). Every PCR program was designed according to a touchdown set up: 30 s at 95 °C, followed by 30 cycles of 1 min at 94 °C, annealing temperatures starting at 60 °C for 1 min (decreasing by 1 °C/cycle), and 1 min at 72 °C for extension. This step was followed by 30 cycles of 96 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 8 min.

Introns and exons were predicted using AUGUSTUS (Stanke et al. 2006). Signal peptide was predicted using SignalP 4.1 Server (Petersen et al. 2011). Prediction for protein molecular parameters was performed using ProtParam (Gasteiger et al. 2003). Pairwise sequence alignments were performed using SIM alignment tool (Duret et al. 1996).The encoding gene of TSGAL was proof read using the iProof TM High-Fidelity PCR Kit, according to the suppliers manual (Bio-Rad Laboratories, Hercules, CA). The flanking sequences were used for primer design. Proof_Fw 5′ TGTGTACGGAAATTGCTA CTTTAGG 3′ and proof_Rv 5′ GTCGTATCTCATTCGC ATCC 3′. See supplementary material for sequence information.

Semi-rational construction of TSGAL mutants

Selection of promising sites to be mutated in order to increase thermostability was done as follows. TSGAL was aligned with 1HJQ from *H. insolens*, 1HJS from *M. thermophila*, and 1FHL from *A. aculeatus* using CLC workbench and ClustalW2 (Goujon et al. 2010). The alignment resulted in 59 residues that were identical in the thermostable galactanase variants 1HJQ and 1HJS but different in TSGAL. These residues were subjected to a more detailed investigation and selection process. The structure for TSGAL was homology modeled on the basis of 1FHL using the HHpred server (Soeding 2005) and the position of the 59 residues identified. Amino acids close to or at the surface were prioritized and their possibility for stabilizing the enzyme by, e.g., restriction of the conformational space, helix stabilization, removal of desolvation penalties, removals of steric clashes, or additional intramolecular interactions assessed. Twenty sites were predicted to be promising candidates for mutagenesis. This set was narrowed further down through B-FIT, using the B-factor values from 1FHL from *A. aculeatus*, and predictions of ddG values, calculated by PoPMuSiC version 2.0 (Dehouck et al. 2009) and amino acids chosen, for which both tools predicted a stabilizing effect. Q171S was solely selected on the basis of its high B-factor in 1FHL and the position of serine in 1HJS with an advantageous prediction value in PoPMuSiC. S319P was selected due to high B-factor in 1FHL and supported by the observation, that 1HJQ has a proline at the corresponding position. The mutants were constructed using the QuickChange® Site-Directed Mutagenesis Kit (Strategene, CA) according to the manual. Primers were designed using the web program https://www.genomics.agilent.com/primerDesignProgram.jsp (Novoradovsky et al. 2005) listed in Table S2.

Production of recombinant endo-1,4-β-galactanases

The sequence encoding the TSGAL was codon optimized for *Pichia pastoris* and cloned into the pJ912-vector system by DNA2.0 (Menlo Park, CA, USA). The vector was linearized as described previously (Silva et al. 2011). The transformation and verification of clones were performed as described in the Invitrogen protocol. The TSGAL mutants and TSGAL were expressed in 50-ml flasks and induced by methanol. Expression of the TSGAL was

upscaled in a 5-l Sartorius Biostat Aplus fermenter with a working volume of 2 l (Silva et al. 2011).

Purification of endo-1,4-β-galactanases

The purification of TSGAL and the mutants were performed using columns loaded with Ni Sepharose 6 Fast Flow (GE Healthcare) following the manual of the manufacturer. The protein was eluted in 20 mM phosphate buffer pH 7.5, 100 mM NaCl, 400 mM imidazole. The imidazole was removed by repeated diafiltration using Vivaspin® 6 10000 MWCO separators (GE Healthcare). The buffer was composed of 50 mM phosphate buffer and 100 mM NaCl and exchanged six times after each centrifugation (10 min at 3000 g). The buffer used for the final dilution was 100 mM Na citrate, 100 mM NaCl pH 4.5. Protein concentrations were measured at 280 nm using Gene5™ TAKE3 module version 1.09 according to the manufacturer's instructions (Biotek ® Instruments Inc.).

Enzymatic assays

The reducing assay was based on the p-hydroxybenzoic acid hydrazide methods and was used to determine the pH and temperature optimum of TSGAL as detailed by Michalak et al. 2012 (Michalak et al. 2012). Substrate solution of 200 μl (20 g/l potato galactan) purchased from Megazyme International Ltd. (Bray, Co. Wicklow, Ireland) and 200 μl of enzyme solution was preheated to reaction temperature. The final substrate concentration was 10 g/l and the enzyme/substrate ratio (E/S) (*v*/*w*) was 0.15 % (15 μg/ml enzyme in reaction) and mixed at 600 rpm (Eppendorf thermomixer) for 10 min. pH of substrate and enzyme solutions were adjusted with 100 mM citrate buffer. The reaction was prepared as described in Michalak et al. 2012 (2012). The statistical design was a central composite factorial design analyzed with multiple linear regression using the lm-procedure in R (R Core Team 2014).

For measurements of half-lives, the enzymes were incubated at 55 and 60 °C in water bath. An enzyme solution (5 μg/ml of the protein) buffered in 100 mM Na citrate buffer at pH 4.5, 100 mM NaCl was preheated in 15 ml falcon tubes at the respective temperature for 5 min. Enzyme solution of 200 μl was added to 200 μl of substrate containing 20 g/l of azo-potato-galactan (Megazyme), which was preheated for 5 min to 45 °C. Thermal deactivation at 60 °C was measured for periods of 0, 2.5, and 5 min (10 min was also used for mutant G305A). At 55 °C, it was measured for periods of 0, 5, 10, and 15 min (25 min was also used for mutant G305A). The reaction was performed at 45 °C for 5 min in a thermomixer (Eppendorf) at 850 rpm. The reaction was terminated by adding 1 ml of 96 % ethanol. The solution was centrifuged at 15,000 g for 10 min and 200 μl were transferred to microtiter plates (two replicates) for absorbance reading at 590 nm. Half-life was determined based on a semilogarithmic plot of activity versus time as shown in Figure S2. Half-lives and specific activities (ΔOD/min/mg protein) were compared using weighted least square analysis followed by Dunnet's test.

Data deposition

The codon-optimized nucleotide sequence for expression of the *T. stipitatus* (ATCC10500) endo-1,4-betagalactanase gene, TSGAL, in *P. pastoris* has been deposited in the GenBank database, with the ID: KM675738.

Results

Expression of endo-1,4-β-galactanase and sequence alignment

T. stipitatus was found to secrete endo-1,4-β-galactanase (TSGAL) in minimal media containing sugar beet pectin as sole carbon source. A partial sequence of 815 bp encoding an endo-1,4-β-galactanase was identified with the designed CODEHOP primers. The DNA flanking sequences were prolonged downstream with 339 bp and upstream with 361 bp using a genome walking technique (Guo and Xiong 2006). The open reading frame of *TSGAL* consisted of 1,062 bp with a start and a stop codon, two exons at 136–468 bp and 526–1251 bp, and an intron from 467–525 bp. TSGAL contained an N-terminal signal peptide encoded by 57 bp. The putative protein was deduced to be 353 AA in length, including the N-terminal signal peptide of 19 AA (see Figure S1). The molecular mass of the mature protein was calculated to be 36.5 kDa.

Sequence similarity

The *TSGAL* gene showed 92 and 79 % identity to mRNA sequences from putative arabinogalactan endo-1,4-beta-galactosidases from *T. stipitatus* ATCC 10500 (uniprot B8MTD9) and *Penicillium marneffei* (uniprot B6QVV5), respectively. Both sequences were obtained from genome sequencing projects. The amino acid sequence of TSGAL (without SP) showed 96.7 and 87.1 % identity to these putative endo-1,4-β-galactosidases protein sequences. TSGAL showed 70.1 and 69.5 % sequence identity to the endo-1,4-β- galactanases from *Aspergillus tubingensis* and *Aspergillus niger*, respectively. Additionally, it showed 62.0 % and 59.6 % sequence similarity to the endo-1,4-β-

galactanases from the thermophilic organisms *Myceliophthora termophila* (1HJS) and *H. insolens* (1HjQ) (Le Nours et al. 2003). Sequences from the respective 3D structures have been used. TSGAL was more related to the endo-1,4-β-galactanases exhibiting temperature optimum at 65 °C compared to the *Aspergilli* galactanases having optima around 50 °C (data not shown). Further information on the sequence relationship is shown in Figure S3, where a ClustalW multiple alignment of TSGAL of *T. stipistatus* and GH53 endo-1,4-β-galactanases from *A. aculeatus* (P48842), *P. chrysogenum* (B5MGR3), *H. insolens* (P83691), and *M. thermophila* (P83692) are shown.

Expression of TSGAL in *P. pastoris*

The gene sequence encoding TSGAL was codon optimized for heterologous expression in *P. pastoris* and cloned in frame with a C-terminal 6×histidin tag. The gene was expressed successfully in *P. pastoris* and produced in 2 l yielding 5.3 g in total. According to the SDS-PAGE, the recombinant protein exhibits a band with the calculated size of 37 kDa. Glycosylated forms of the enzyme could be visualized by silver staining revealing a protein smear with molecular weights exceeding 37 kDa (data not shown).

Statistical design was implemented to estimate the optimal reaction conditions for TSGAL using potato galactan as the substrate. The optimal reaction condition for the endo-1,4-β-galactanase was determined to be 46 °C at pH 4.5 as depicted in Fig. 1, at which the specific activity was 6.93 μmol/min/mg enzyme.

Single-site mutant enzyme activity

Nine single sites were selected for site-directed mutagenesis on the basis of multiple alignment, structural considerations (using a 3D homology model of TSGAL) as well as B-FIT and PoPMuSiC (see Materials and Methods for more details). The locations of the mutated sites are shown in Fig. 2.

K18A and Q171S

K18 and Q171 were selected primary due to the high B-factor value of the respective amino acids in the structure of the homologue 1FHL and were replaced by alanine and serine, the amino acid at the corresponding position in 1HJQ and 1HJS. Additionally, PoPMuSiC predicted a favorable ddG for Q171S (see Table 1). K18 is placed at the C-terminus in one of the α-helices in TSGAL and an alanine substitution could probably stabilize this helix as alanine is believed to have a high helix propensity (Spek et al. 1999; Myers et al. 1997).Unfortunately, a decrease in half-lives of both K18A

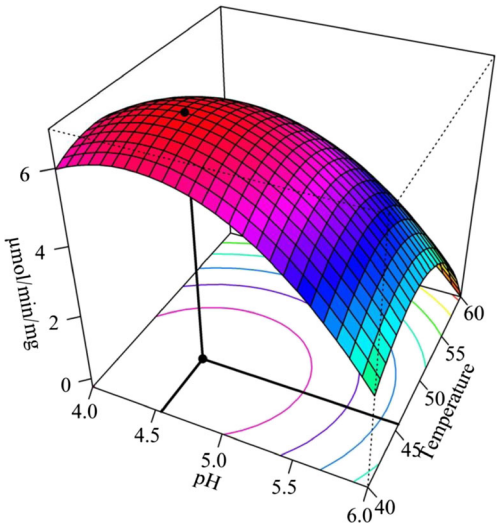

Fig. 1 Response surface plot of the optimal condition for TSGAL's activity on potato galactan pH 4.0–6.0 and 40–60 °C. The optimal condition is pH 4.5 and 46 °C

(7.9±0.5 min) and Q171S (8.4±1.4 min) was observed compared to the half-life of the wild type enzyme TSGAL, which was 13.3±1.6 min at 55 °C (see Fig. 3).

N71A, G116D, and G287S

The residues were predicted from the homology structure to be associated with helix secondary structure elements. Both N71 and G287 are located at the C-terminus of the helix spanning residues 59–71 and 271–287, respectively. The N71A mutant was suggested to have two beneficial effects: (i) removal of the hydrophilic asparagine, thereby removing

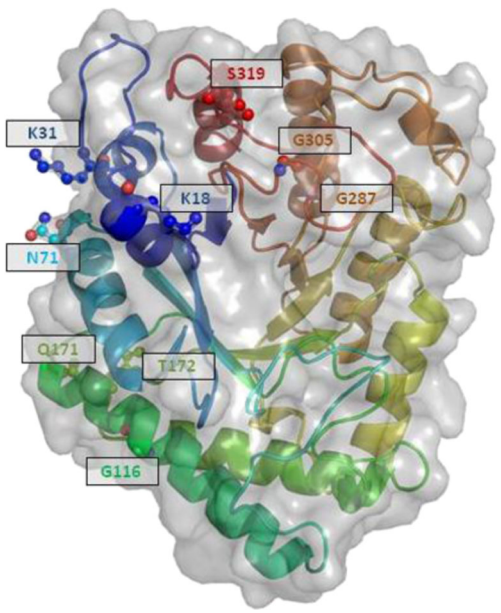

Fig. 2 3D homology model of TSGAL on the basis of the structure of *A. aculeatus* (1FHL). The nine mutated sites are represented as spheres

Table 1 TSGAL mutations showing the result from the multiple alignment, B_FIT, and PoPMuSiC and their respective half-lives at 55 and 60 °C

TSGAL		Alignment		B-fitter	PoPMuSiC	Half-lives [min]	
Mutation	1HJQ	1HJS	1FHL	B-factor (1FHL)	ddG	55 °C	60 °C
TSGAL						13.3±1.6	2.1±0.2
K18A	Ala	Ala	Glu	58	−0.08	7.9±0.5*	n.m.
K31P	Pro	Pro	Ala	83	−0.87	20.6±3.7	2.5±0.1
N71A	Ala	Ala	Ala	79	−0.57	16.5±1.0*	2.1±0.3
G116D	Asp	Asp	Glu	21	−0.15	17.0±1.4	2.5±0.3
Q171S	Asn	Ser	Ala	53	−0.73	8.4±1.4	n.m.
T172P	Pro	Pro	Thr	61	1.12	20.7±2.6*	2.3±0.1
G287S	Ser	Ser	Ala	39	−0.63	13.8±2.2	n.m.
G305A	Ala	Ala	Ala	54	0.25	114.4±8.1*	15.3±1.8*
S319P	Pro	Gln	Tyr	58	0.95	11.0±0.4	n.m.

A negative value of ddG in PoPMuSiC predictions gives an indication of stability improvements. A high B-factor value corresponds to high flexibility

n.m. not measured

*Significantly different from TSGAL

an eventually desolvation penalty (Schmid 2011) and (ii) replacement of asparagine by an amino acid with a higher helix propensity (Blaber et al. 1993). Both B-fitter and PoPMuSiC classified this mutant as being beneficial. Mutating G287 to serine was expected to have a stabilizing effect due to the removal of the helix terminating glycine and the restriction of the conformational space by introducing an amino acid with a (more bulky) side chain. While the results from B-FIT suggested a small chance for favorable effects when mutating this residue, PoPMuSiC predicted a significant stabilization. G116 was expected to be a promising site to be subjected to mutation due to its location in the middle of the predicted α-helix spanning residues 103–124. Glycine is known as a helix breaker, and its removal was expected to stabilize the helix significantly (Lyu et al. 1991). By replacing G116 with the hydrophilic residue asparagine, the risk of protein precipitation is expected to be minimal. While PoPMuSiC predicted a slight stabilization associated with this mutant, B-FIT did not classify this mutant as being beneficial. When comparing the measured half-live times, the mutant G287S did not differ from the wild-type (see Table 1). The half-lives of N71A and G116D were increased with 17 and 28 % at 55 °C, respectively. At 60 °C, it did not seem that mutation N71A had an effect on the thermostability (see Fig. 3). However, the half-life of G116D at 60 °C was increased by 19 %, e.g., different from TSGAL. According to Fig. 3, there were only minor differences in the specific activities of the mutants compared to TSGAL.

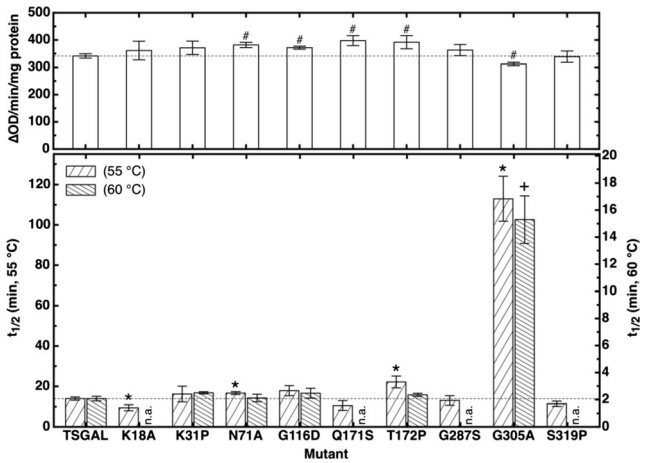

Fig. 3 Specific activities at 45 °C of all variants are shown in the top. *Number sign* significantly different from TSGAL. Below the half-lives of TSGAL and its mutants are shown measured at 55 and 60 °C. Values are means ± SD. *Asterisk* significantly different from TSGAL at 55 °C. *Plus sign* significantly different from TSGAL at 60 °C

Proline substitutions

K31P, T172P, and S319P were based on the theory of stabilizing the enzyme by restricting the conformational space when introducing the "stiff" residue proline leading to an entropic stabilization. Both 1HJQ and 1HJS contain proline residues at the corresponding positions. Apart from having a high B-factor value, K31 was selected in order to remove the desolvation free energy penalty of the hydrophobic lysine residue. This residue had the highest B-factor of all sites selected in this study (A31 in 1FHL) and PoPMuSiC predicted a favorable ddG value (see Table 1). K31P exhibited increased half-lives of 55 % at 55 °C and 20 % at 60 °C, respectively (see Fig. 3). The increase in $t_{1/2}$ at both temperatures compared with the TSGAL was found to be statistically significant. The choice of the T172P mutant was supported by the high B-factor of T172 in 1FHL. This mutation gave an increase in half-life of 56 % at 55 °C. In addition to the

abovementioned considerations, S319P was selected due to the high B-factor of the corresponding residue T319 in 1FHL. In this mutant, neither the half-life nor the specific activity of S319P was affected significantly (see Table 1 and Fig. 3).

G305A

This mutation was selected in order to stabilize the enzyme entropically by replacing glycine with alanine, thereby restricting the conformational space. Alanine was chosen due to its presence in 1HJQ and 1HJS at the corresponding position. This mutant was selected although neither the B-factor nor the ddG value suggested this mutation to be beneficial. The G305 to Ala mutation gave the highest stabilization of TSGAL. The mutated enzyme exhibited half-lives of 114.4±8.1 min at 55 °C and 15.3±1.8 min at 60 °C were observed. This corresponds approximately to a ninefold increase. There was only a minor effect on the specific activity (see Fig. 3).

Discussion

In this paper, we report a novel endo-1,4-β-galactanase (TSGAL) secreted by *T. stipitatus* when grown on sugar beet pectin as sole carbon source. The enzyme has a potential to release prebiotics from agricultural side streams such as potato pulp and sugar beet pectin as shown previously for homologous enzymes (Michalak et al. 2012; Thomassen et al. 2011b). TSGAL contains all the key residues from Clan-A enzymes and its homology model shares the same overall structure as the *A. aculeatus* endo-1,4-β-galactanase (Ryttersgaard et al. 2002).

The methylotropic yeast *P. pastoris* has been used widely as a host for expression of fungal enzymes (Cregg et al. 2000). TSGAL and its mutants were successfully expressed in *P. pastoris* with an indication of glycosylation. Often, the enzymes seem to be highly glycosylated, whereby the stability or activity can be affected (Silva et al. 2014; Melgaard and Svendsen 1994). For elucidation of any effects of the glycosylation, TSGAL could be co-expressed in a prokaryotic host or the glycosylation site N-X-S/T (TSGAL: N71PSD and N132YTM) could be mutated before expression in *P. pastoris* In addition, it is not verified whether the native enzyme identified in AZCL assay is the same as the one that were identified by iCODEHOP. Purification with additional protein sequence determination of the native enzyme would be required for full validation.

According to the literature, there are no fungal thermostable galactanases reported with low pH optima available. Acid-tolerant and thermostable endo-1,4-β-galactanases are useful when supplied to animal feed as described in the recent US patent (De Maria et al. 2013). TSGAL has an optimum at pH 4.5 whereas the pH optimum for the thermostable endo-1,4-β-galactanases from *M. termophila* (1HJS) and *H. insolens* (1HJQ) are 8.5 and 7.0, respectively (Le Nours et al. 2003). In addition, the pH optimum of TSGAL is also lower than that of the endo-1,4-β-galactanases from *Emericella nidulans* with a pH optimum at 5.0 (Michalak et al. 2012). Since TSGAL did not exhibit high thermostability above 60 °C, but exhibits relatively high similarity to homologous enzymes from thermophiles, its comparison with the thermostable variants make this enzyme a relevant candidate for further enhancement of stability by site-directed mutagenesis. We therefore selected a small number of sites in order to create single mutants. The selection process involved alignment with two thermostable endo-1,4-β-galactanases, assessment of structural features using a predicted homology structure, B-factor values from the well-known endo-1,4-β-galactanase from *A. aculeatus* and prediction using PoPMuSiC.

Site-directed mutagenesis of five out of nine amino acid residues in TSGAL showed increased half-lives ranging from 17 to 760 % at 55 °C. Four out of these five mutants (K31P, G116D, T172P, and G305A) were also found to have increased thermostability at 60 °C. K31P and T172P were both selected due to the location of residue 31 and 172 in a region with theoretically high flexibility (B-factor values from 1FHL) in order to restrict the conformational space. However, in some cases, proline can cause elimination of crucial non-covalent interactions or by limiting the number of dihedral angle combinations, thereby destabilizing or even disrupting the protein (Eijsink et al. 2005). Along with other examples (Silva et al. 2013; Watanabe et al. 1994; Yi et al. 2011), this work proves that carefully selected mutations with proline insertion may increase the thermostability. The increase of the stabilization is probably due to an "entropic stabilization" where the chain entropy of the unfolded state is decreased (Eijsink et al. 2004; Matthews et al. 1987).

Mutations N71A and G287S have also increased half-live times. Both sites are located at the C-terminus of α-helix. Two explanations or a combination of the two for the stabilizing effect in mutant N71A are likely. Stabilization is conducted either by removal of a free energy desolvation penalty or helix stabilization by inserting a residue with higher helix propensity. To get a deeper insight in the molecular basis of stabilization, site saturated mutants at this position could reveal whether mutations to other hydrophobic amino acids lead to the same effect (desolvation penalty elimination) or amino acids with high helix propensity lead to stabilization (helix stabilization). Another explanation for the increased

half-life could be the removal of asparagine that can undergo deamidation, which is temperature dependent. The deamidation of asparagine leads to succinimide formation that is unstable in aqueous solution (Daniel et al. 1996). In two studies of α-amylases, the risk of having deamidation was removed with success by replacement of asparagines (Declerck et al. 2000; Rahimzadeh et al. 2012).

G305A exhibited half-lives of 114±8 min and 15.3± 1.8 min at 55 and 60 °C, respectively. It was chosen due to structural considerations, non-supported by predictions from B-fitter or PoPMuSiC. Interestingly, this mutation led to the most significant stabilization. Similar results were obtained in mutants of a neutral protease from *Bacillus subtilis* where both G147A (in a helix) and G189A (in a loop) were constructed as single and double mutants, respectively. The single mutants showed increased thermostability, and the mutations were additive. The increased stability of G189A is explained by an increased internal packing of the protein (Margarit et al. 1992).The stabilization of TSGAL by G305A is probably due to an "entropic stabilization" (rigidification). Glycine substitution with alanine reduces the conformational space particularly of the unfolded protein thereby shifting the equilibrium between the unfolded and folded state towards the latter (Eijsink et al. 2004; Matthews et al. 1987).

G305A and G116D are also expected to have additive effects on thermostability, too. Another promising idea could be to mutate residues in the loop, where G305 is located, since this loop might be crucial to convey overall stability to the enzyme. PoPMuSiC did not predict G305A or T172P as being more stabilizing although their half-lives were increased. One explanation for the prediction is that the dataset that was used to train and validate the PoPMuSiC model did not embrace GH53-related proteins. Another explanation could be that the quality of the predicted 3D homology model of TSGAL is not reliable enough to do the prediction.

The end-point of our work is to recover value-added components from renewable biomass including pectinaceous streams, which is an important aspect for developing a viable, competitive bio-economy. The currently available commercial enzymes are able to catalyze the degradation of most of the plant cell wall materials available, but there has so far been very little incentive to develop thermally stable enzymes for enzymatic processing of high-viscosity pectinaceous streams. The work presented verifies that creation of a targeted library for site saturation mutagenesis is a valid approach for fast improvement of thermal robustness of enzymes. The data obtained also illustrates the amazing feature that even single-point mutations can drastically change the thermostability of enzymes. Enzyme functionality is pivotal to biology and biotechnology and the use of targeted combinatory methods for fast development of novel enzymes makes it possible to develop both an improved understanding of the functionality of enzymes and speedy evolution of enzymes for new applications.

This is the first report on the characterization of the endo-1,4-β-galactanases from *T. stipitatus*. Furthermore, the half-life of the enzyme was proven to be enhanced by rational design of selected mutant candidates showing significant increase of half-lives in T172P, G116D, K31P, and G305A. The study leads to further work with identification of additional single mutations, mutant combinations, and assessment of their thermodynamics. Additionally, it would be worth to test the G305A mutant in the release of potential prebiotics from potato pulp in larger scale.

Acknowledgments This study was supported by the Danish Strategic Research Council's Committee on Food and Health (FøSu) project "Biological Production of Dietary Fibers and Prebiotics" no. 2101-06-0067. Thanks to Jens Christian Frisvad from DTU, Department of Systems Biology, Center for Microbial Biotechnology for providing *Talaromyces stipitatus*.

References

Blaber M, Zhang X, Matthews B (1993) Structural basis of amino acid alpha-helix propensity. Science 260:1637–1640

Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K (2012) Engineering the third wave of biocatalysis. Nature 485:185–194

Boyce R, Chilana P, Rose TM (2009) iCODEHOP: a new interactive program for designing COnsensus-DEgenerate Hybrid Oligonucleotide Primers from multiply aligned protein sequences. Nucleic Acids Res 37:W222–W228

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37:D233–D238

Cregg JM, Cereghino JL, Shi JY, Higgins DR (2000) Recombinant protein expression in *Pichia pastoris*. Mol Biotechnol 16:23–52

Daniel R, Dines M, Petach H (1996) The denaturation and degradation of stable enzymes at high temperatures. Biochem J 317:1–11

Davids T, Schmidt M, Boettcher D, Bornscheuer UT (2013) Strategies for the discovery and engineering of enzymes for biocatalysis. Curr Opin Chem Biol 17:215–220

De Maria L, Svendsen A, Borchert TV, Christensen LL, Larsen S, Ryttersgaard C (2013) Galactanase variants (Patent: US8367390 B2)

Declerck N, Machius M, Wiegand G, Huber R, Gaillardin C (2000) Probing structural determinants specifying high thermostability in *Bacillus licheniformis* alpha-amylase. J Mol Biol 301:1041–1057

Dehouck Y, Grosfils A, Folch B, Gilis D, Bogaerts P, Rooman M (2009) Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0. Bioinformatics 25:2537–2543

Duret L, Gasteiger E, Perriere G (1996) LALNVIEW: a graphical viewer for pairwise sequence alignments. Comput Appl Biosci 12:507–510

Eijsink VGH, Bjork A, Gaseidnes S, Sirevag R, Synstad B, van den Burg B, Vriend G (2004) Rational engineering of enzyme stability. J Biotechnol 113:105–120

Eijsink V, Gaseidnes S, Borchert T, van den Burg B (2005) Directed evolution of enzyme stability. Biomol Eng 22:21–30

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel R, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788

Gilis D, Rooman M (2000) PoPMuSiC, an algorithm for predicting protein mutant stability changes. Application to prion proteins. Protein Eng 13:849–856

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38:W695–W699

Gruber K, Klintschar G, Hayn M, Schlacher A, Steiner W, Kratky C (1998) Thermophilic xylanase from *Thermomyces lanuginosus*: high-resolution X-ray structure and modeling studies. Biochemistry (N Y) 37:13475–13485

Guo H, Xiong J (2006) A specific and versatile genome walking technique. Gene 381:18–23

Holck J, Hjerno K, Lorentzen A, Vigsnaes LK, Hemmingsen L, Licht TR, Mikkelsen JD, Meyer AS (2011) Tailored enzymatic production of oligosaccharides from sugar beet pectin and evidence of differential effects of a single DP chain length difference on human faecal microbiota composition after in vitro fermentation. Process Biochem 46:1039–1049

Houbraken J, de Vries RP, Samson RA (2014) Modern taxonomy of biotechnologically important *Aspergillus* and *Penicillium* species. Adv Appl Microbiol 86:199–249

Kwasigroch JM, Gilis D, Dehouck Y, Rooman M (2002) PoPMuSiC, rationally designing point mutations in protein structures. Bioinformatics 18:1701–1702

Le Nours J, Ryttersgaard C, Lo Leggio L, Ostergaard P, Borchert T, Christensen L, Larsen S (2003) Structure of two fungal beta-1,4-galactanases: searching for the basis for temperature and pH optimum. Protein Sci 12:1195–1204

Li D, Yang Y, Peng Y, Shen C (1998) Purification and characterization of extracellular glucoamylase from the thermophilic *Thermomyces lanuginosus*. Mycol Res 102:568–572

Lutz S (2010) Beyond directed evolution—semi-rational protein engineering and design. Curr Opin Biotechnol 21:734–743

Lyu P, Wang P, Liff M, Kallenbach N (1991) Local effect of glycine substitution in a model helical peptide. J Am Chem Soc 113:3568–3572

Magrane M, UniProt Consortium (2011) UniProt Knowledgebase: a hub of integrated protein data. Database-the Journal of Biological Databases and Curation

Margarit I, Campagnoli S, Frigerio F, Grandi G, Defilippis V, Fontana A (1992) Cumulative stabilizing effects of glycine to alanine substitutions in *Bacillus subtilis* neutral protease. Protein Eng 5:543–550

Matthews BW, Nicholson H, Becktel WJ (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc Natl Acad Sci U S A 84:6663–6667

Melgaard M, Svendsen I (1994) Different effects of N-glycosylation on the thermostability of highly homologous bacterial (1,3-1,4)-beta-glucanases secreted from yeast. Microbiology-Uk 140:159–166

Michalak M, Thomassen LV, Roytio H, Ouwehand AC, Meyer AS, Mikkelsen JD (2012) Expression and characterization of an endo-1,4-beta-galactanase from *Emericella nidulans* in *Pichia pastoris* for enzymatic design of potentially prebiotic oligosaccharides from potato galactans. Enzyme Microb Technol 50:121–129

Myers JK, Pace CN, Scholtz JM (1997) A direct comparison of helix propensity in proteins and peptides. Proc Natl Acad Sci U S A 94:2833–2837

Naumoff DG (2011) Hierarchical classification of glycoside hydrolases. Biochemistry-Moscow 76:622–635

Novoradovsky A, Zhang V, Ghosh M, Hogrefe H, Sorge JA, Gaasterland T (2005) Computational Principles of Primer Design for Site Directed Mutagenesis. Technical Proceedings of the 2005 NSTI Nanotechnology Conference and Trade Show, Volume 1 : 532–533

Otten H, Michalak M, Mikkelsen JD, Larsen S (2013) The binding of zinc ions to *Emericella nidulans* endo-beta-1,4-galactanase is essential for crystal formation. Acta Crystallographica Section F-Structural Biology and Crystallization Communications 69:850–854

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Rahimzadeh M, Khajeh K, Mirshahi M, Khayatian M, Schwarzenbacher R (2012) Probing the role of asparagine mutation in thermostability of *Bacillus* KR-8104 alpha-amylase. Int J Biol Macromol 50:1175–1182

Ravn HC, Meyer AS (2014) Chelating agents improve enzymatic solubilization of pectinaceous co-processing streams. Process Biochem 49:250–257

Ryttersgaard C, Lo Leggio L, Coutinho P, Henrissat B, Larsen S (2002) *Aspergillus aculeatus* beta-1,4-galactanase: substrate recognition and relations to other glycoside hydrolases in clan GH-A. Biochemistry (N Y) 41:15135–15143

Sakamoto T, Nishimura Y, Makino Y, Sunagawa Y, Harada N (2013) Biochemical characterization of a GH53 endo-beta-1,4-galactanase and a GH35 exo-beta-1,4-galactanase from *Penicillium chrysogenum*. Appl Microbiol Biotechnol 97:2895–2906

Schmid F (2011) Lessons about protein stability from in vitro selections. Chembiochem 12:1501–1507

Silva IR, Jers C, Otten H, Nyffenegger C, Larsen DM, Derkx PMF, Meyer AS, Mikkelsen JD, Larsen S (2014) Design of thermostable rhamnogalacturonan lyase mutants from *Bacillus licheniformis* by combination of targeted single point mutations. Appl Microbiol Biotechnol 98:4521–4531

Silva IR, Larsen DM, Jers C, Derkx P, Meyer AS, Mikkelsen JD (2013) Enhancing RGI lyase thermostability by targeted single point mutations. Appl Microbiol Biotechnol 97:9727–9735

Silva IR, Larsen DM, Meyer AS, Mikkelsen JD (2011) Identification, expression, and characterization of a novel bacterial RGI Lyase enzyme for the production of bio-functional fibers. Enzyme Microb Technol 49:160–166

Soeding J (2005) Protein homology detection by HMM-HMM comparison. Bioinformatics (Oxford) 21:951–960

Spek EJ, Olson CA, Shi ZS, Kallenbach NR (1999) Alanine is an intrinsic alpha-helix stabilizing amino acid. J Am Chem Soc 121:5571–5572

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B (2006) AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res 34:W435–W439

Thomassen LV, Larsen DM, Mikkelsen JD, Meyer AS (2011a) Definition and characterization of enzymes for maximal biocatalytic solubilization of prebiotic polysaccharides from potato pulp. Enzyme Microb Technol 49:289–297

Thomassen LV, Vigsns LK, Licht TR, Mikkelsen JD, Meyer AS (2011b) Maximal release of highly bifidogenic soluble dietary fibers from industrial potato pulp by minimal enzymatic treatment. Appl Microbiol Biotechnol 90:873–884

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

Yi Z-L, Pei X-Q, Wu Z-L (2011) Introduction of glycine and proline residues onto protein surface increases the thermostability of endoglucanase CelA from *Clostridium thermocellum*. Bioresour Technol 102:3636–3638