Computational and Structural Biotechnology Journal 13 (2015) 256–264

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Effect of mutations on the thermostability of *Aspergillus aculeatus* β-1,4-galactanase

Søs Torpenholt [a,1], Leonardo De Maria [b,2], Mats H.M. Olsson [a], Lars H. Christensen [b], Michael Skjøt [b,1], Peter Westh [c], Jan H. Jensen [a], Leila Lo Leggio [a,*]

[a] *Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark*
[b] *Novozymes A/S, Smørmosevej 25, 2880 Bagsværd, Denmark*
[c] *NSM, Research Unit for Functional Biomaterials, University of Roskilde, Universitetsvej 1, 4000 Roskilde, Denmark*

ARTICLE INFO

Article history:
Received 30 October 2014
Received in revised form 27 March 2015
Accepted 31 March 2015
Available online 9 April 2015

Keywords:
β-1,4-galactanase
Thermostability
Computational prediction
Protein design
GH53

ABSTRACT

New variants of β-1,4-galactanase from the mesophilic organism *Aspergillus aculeatus* were designed using the structure of β-1,4-galactanase from the thermophile organism *Myceliophthora thermophila* as a template. Some of the variants were generated using PROPKA 3.0, a validated pK_a prediction tool, to test its usefulness as an enzyme design tool. The PROPKA designed variants were D182N and S185D/Q188T, G104D/A156R. Variants Y295F and G306A were designed by a consensus approach, as a complementary and validated design method. D58N was a stabilizing mutation predicted by both methods. The predictions were experimentally validated by measurements of the melting temperature (T_m) by differential scanning calorimetry. We found that the T_m is elevated by 1.1 °C for G306A, slightly increased (in the range of 0.34 to 0.65 °C) for D182N, D58N, Y295F and unchanged or decreased for S185D/Q188T and G104D/A156R. The T_m changes were in the range predicted by PROPKA. Given the experimental errors, only the D58N and G306A show significant increase in thermodynamic stability.

Given the practical importance of kinetic stability, the kinetics of the irreversible enzyme inactivation process were also investigated for the wild-type and three variants and found to be biphasic. The half-lives of thermal inactivation were approximately doubled in G306A, unchanged for D182N and, disappointingly, a lot lower for D58N. In conclusion, this study tests a new method for estimating T_m changes for mutants, adds to the available data on the effect of substitutions on protein thermostability and identifies an interesting thermostabilizing mutation, which may be beneficial also in other galactanases.

© 2015 Torpenholt et al. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. β-1,4-galactanases

Endo-β-1,4-galactanases (EC 3.2.1.89) are glycoside hydrolases which hydrolyse the β-1,4-O-glycosidic linkages in β-1,4-galactan and arabinogalactan type I, parts of pectin found in the non-woody plant cell wall [1]. In the carbohydrate active enzyme database (CAZY) [2] they are classified in GH53. β-1,4-Galactanases have received some attention in industrial applications as demonstrated by a number of patents in the 1990s/2000s (eg [3,4]) and recent renewed interest in the academic literature for their potential in producing prebiotic oligosaccharides [5]. In several bacteria, for example the thermophile *Geobacillus stearothermophilus*, *Erwinia chrysanthemi*, *Bacillus subtilis* and *Bifidobacterium breve* [6–10], β-1,4-galactanase genes have been shown to be part of gene clusters involved in galactan utilization, consisting, in addition to the endo-galactanase, of a GH42 β-galactosidase, a galactooligosaccharide transport system, and a transcriptional regulator.

Both fungal and bacterial β-1,4-galactanases have been characterized biochemically and crystallographically. Fungal galactanases tend to release shorter products and are able to hydrolyse to some extent small oligosaccharides and small chromogenic substrates [11,12] while the investigated bacterial enzymes tend to release larger oligosaccharides, have a more extended substrate binding groove and cannot act on small substrates, presumably because of non-productive substrate

---

Abbreviations: AaGal, β-1,4-galactanase from *Aspergillus aculeatus*; AZCL-galactan, azurine-crosslinked galactan; CAZY, carbohydrate active enzyme database; DSC, differential scanning calorimetry; MtGal, β-1,4-galactanase from *Myceliophthora thermophila*; *Tm*, melting temperature; TsGal, *Talaromyces stipitatus* galactanase; WT, wild type.

* Corresponding author. Tel.: +45 35320295.
E-mail address: leila@chem.ku.dk (L. Lo Leggio).

[1] Novozymes A/S, Smørmosevej 25, 2880 Bagsværd, Denmark.

[2] Novo Nordisk A/S, Novo Nordisk Park, 2760 Måløv Denmark.

http://dx.doi.org/10.1016/j.csbj.2015.03.010

2001-0370/© 2015 Torpenholt et al. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

binding [6,11,12]. Several 3D structures have been elucidated for GH53 β-1,4-galactanases, among which a bacterial galactanase [11], while for fungal galactanases these represent both mesophilic and thermophilic enzymes [13–15].

1.2. Engineering thermostability

Protein engineering and site-directed mutagenesis are used routinely to establish biological function and role of amino-acid residues in proteins, like the stability. The protein stability can be changed significantly by single or multiple mutations of specific amino acids, sometimes resulting in a beneficial effect. This type of work is however expensive and time-consuming so rapid prediction of how mutations might affect the stability of a protein is desirable. An important parameter in determining the protein stability, including thermostability, is the determination of the pK_a value of all the ionizable residues in a protein, which however, experimentally, is also challenging and time consuming. Computational *pKa* prediction programs have therefore been developed [16]. One of these programs is PROPKA [17], that utilize a very fast empirical approach of *pKa* prediction. PROPKA is today one of the most widely used programs due to its ease of use and speed, at the same time yielding accurate results as compared to other programs [18]. Details of the method and implementation in PROPKA 3.0, the version used in this study, can be found elsewhere [19–21].

1.3. Aim of this study

Within the scope of this article, we pursue two goals. First of all, we investigate if PROPKA can be applied to design thermostable proteins and predict the magnitude of the gain in thermostability upon mutation. We use as model system the mesophilic *Aspergillus aculeatus* β-1,4-galactanase (AaGal, GenBank ID AAA32692.1, PDB IDs: 1FHL, 1FOB; [13]) using as template for mutation *Myceliophthora thermophila* (also known as *Thielavia heterothallica*) β-1,4-galactanase (MtGal, GenBank ID AAE73520.1, PDB ID: 1HJS [14]). We pursue therefore also the secondary goal of obtaining a more thermostable AaGal. A consensus approach [22,23] was used as a complementary design tool to obtain more thermostable variants, which were also evaluated by PROPKA. The thermostability of these variants together with that of the wild type enzyme was experimentally determined by their melting temperature, T_m, and thermal irreversible inactivation rate.

2. Material and methods

2.1. Cloning, expression and purification of enzymes

Cloning, expression and purification of WT AaGal is described in detail in [24]. The enzyme was more than 95% pure as judged from the SDS-PAGE gel.

Variants of AaGal were made using the original expression vector [24] as template. Mutations were introduced by PCR using a method essentially as the Quick Change Site Directed Mutagenesis Kit (Roche Applied Science, Indianapolis, IN). The primers used for mutagenesis are shown in Table 1. G104D/A156R was generated by consecutive rounds of PCR. All constructs were sequence verified prior to transformation into *Aspergillus oryzae*. The expression plasmids were transformed into *Aspergillus* as described in [25]. For each of the constructs 4–6 strains were isolated, purified and cultivated in microtiter plates. Expression was determined using the AZCL-galactan assay as described in Section 2.3 and the best producing strain was fermented in shake flasks.

Table 1
Primers used to construct the described variants.

D182N	atgatccatttggacAATggctggagctggg	Gtccaaatggatcatgatcttgggg
S185D, Q188T	ttggacgacggctggGATtgggatACCcagaactactttacga	Ccagccgtcgtccaaatggatcatg
G104D	tccacgaccgatctcGATactttgaaatggc	gagatcggtcgtggaccagccggaa
A156R	tattcgaacatcggcAGActgctgcactcg	Gccgatgttcgaatagctgctcgtt
D58N	agcgacggcagctacAATctggactacaa	gtagctgccgtcgctggggttcacc
Y295F	gatggtctgggagtgTTCtattgggagcc	cactcccagaccatcagtagtggcc
G306A	tggatcggcaatgctGCCttgggttcgag	agcattgccgatccaggctggctcc

The variants of AaGal were purified by hydrophobic interaction as essentially described in [14]. Ammonium sulphate (1 M), sodium chloride (1 M) and buffer (50 mM sodium acetate buffer, pH 4.5) were added to the culture filtrate and the pH adjusted to 4.5. The culture filtrate was then loaded on a Toyopearl Butyl 650 M from Tosoh. A gradient was applied from this loading buffer to 100% buffer B (50 mM sodium acetate buffer, pH 4.5) over ten column volumes, during which the enzymes eluted. The variants D58N, D182N, G104D/A156R were purified further by anion exchange using Q-sepharose 16/10 with a gradient of 0 to 1 M sodium chloride in 50 mM sodium acetate buffer, pH 4.5. 5 mL fractions were collected and the ones with galactanase activity (AZCL-galactan assay) were pooled and checked for purity by SDS-PAGE. Fractions with estimated 95% or more purity were used in further work.

2.2. Protein concentration and electrophoresis

The protein concentration was determined by absorbance measurements at 280 nm using a ND-1000 spectrophotometer from Fischer scientific. The purity was judged by SDS-PAGE on polyacrylamide gels (4–20% from Biorad or 4–12% from Nupage) after heat denaturation in the presence of DTT. Gels were stained/destained with Instant Blue from Expedeon.

2.3. Enzyme assays

2.3.1. AZCL galactan

This semiquantitative assay was used to identify expressing transformants and galactanase containing fractions. The assay was carried out in a micro titer plate. 20 μL sample (culture broth or elution fractions from purification, or MilliQ water as blank) and 200 μL 4 mg/mL AZCL-galactan (lupin) (from Megazyme International Ireland.) in 0.1 M sodium acetate buffer, pH 4.0, containing 0.01% Triton X-100, was added to each well and incubated for 10 min at room temperature with shaking and then 5 min at room temperature without shaking, to let the insoluble particles settle. The absorbance at 595 nm was then measured on 100 μL of the supernatant without stopping the reaction.

2.3.2. Azo-galactan

Azo-galactan is partly debranched potato galactan dyed with Remazol Brilliant Blue R (about one dye molecule per 20 sugar residues) and was purchased from Megazyme International Ireland. The assay was carried out essentially using the manufacturer's procedure (described in Megazyme booklet AGALP 11/99) with 10 μL pre-equilibrated enzyme solution (0.1 mg/mL in 50 mM sodium acetate buffer pH 4.5), 500 μL pre-equilibrated substrate solution (2% w/v azo-galactan in 50 mM sodium acetate buffer, pH 4.5) and 490 μL pre-heated buffer (50 mM sodium acetate buffer pH 4.5) in Eppendorf tubes. The reaction was incubated for 10 min at 40 °C and stopped by addition of 2.5 mL 95% ethanol with vigorous stirring, which will precipitate high molecular weight galactan. After incubation at room temperature for 10 min, the tubes were centrifuged for 10 min at 1800 g, and the A_{590} of the supernatant measured. A reaction mixture to which10 μL buffer was added instead of enzyme solution was used as blank. The A_{590} is correlated to the extent of small molecular weight dyed galactan fragments released, which is linearly correlated to the amount of reducing sugars released. All azo-galactan assay measurements were made in triplicates.

2.4. Irreversible thermal inactivation rate

The rate of irreversible thermal inactivation was determined at 55 °C by incubation of 0.1 mg/mL enzyme. Aliquots were taken out at appropriate intervals and immediately cooled on ice. Activity was measured using the azo-galactan assay. A is the activity remaining after heating for time *t*, while A_0 is the original activity for each variant without incubation at 55 °C. The inactivation rate was determined by non-linear regression of the A vs time (min) curves to the equation of a biphasic deactivation process (see Results for additional details).

2.5. Differential scanning calorimetry (DSC)

All protein samples were buffer changed into 50 mM sodium acetate buffer, pH 4.5 using a PD-10 (Sephadex™ G-25) column or a 10.000MWCO slide-A-lyser dialysis cassette from Pierce and diluted to 0.2 or 0.5 mg/mL with 50 mM sodium acetate buffer, pH 4.5, degassed for 10 min and equilibrated for 10 min, before each DSC experiment. The DSC experiments were performed using a nano DSC III from TA instruments with cell volumes of 300 μL. The heating scan rate was 60 °C/h. The calorimetric reversibility of the thermally induced transition was checked by reheating the protein solution several times in the calorimetric cell after cooling at 120 °C/h. The heating temperatures were chosen so they covered the endothermic peak (see below for details).

Measurements were carried out in triplicate for WT and in duplicate for the variants. Raw data was analysed with NanoAnalyze Data Analysis version 2.0.1 from TA Instruments by subtracting the buffer/buffer scan from the protein/buffer scan of each experiment and normalized regarding protein concentration. The data were fitted directly to the equation of an irreversible unfolding model enabling calculation of an apparent melting temperature (T_m) (see below for more details).

2.6. Identification of mutation sites by PROPKA

In order to screen possible mutations we used PROPKA to identify stabilizing motifs that were present in the thermophilic template but not in the mesophilic variant. PROPKA predicts protein pK_a values, and can therefore be used to examine the electrostatic part of the protein folding free energy, $\Delta G^{uf\rightarrow f}$. This is related to shifts in protein pK_a values because the interactions that stabilize the folded protein compared to the unfolded protein are the same interactions that shifts the pK_a values of the residues in the unfolded state to its folded state. The total electrostatic protein unfolding free energy can be made into a thermodynamic cycle where the unfolding free energy is calculated as

$$\Delta G^{f\rightarrow uf}(pH) = \Delta G^{f\rightarrow uf}_{neutral} + \Delta\Delta G^{f\rightarrow uf}_{elstat}\left(pH_{ref}\right) + \Delta\Delta G^{f\rightarrow uf}_{elstat}\left(pH_{ref}, pH\right). \tag{1}$$

The first term is the unfolding free energy when all ionizable residues are in their neutral form, the second term is the unfolding contribution when ionizable residues are charged to a reference pH value, pH_{ref}, and the third term is the unfolding contribution when changing the charge of the ionizable residues according to an actual pH value. In this particular case we use a pH_{ref} where all acids are uncharged and all bases are fully charged, which makes the contribution from the second term in the equation simple, and focus on the last term. The last term in the equation is the pH dependent term and can be calculated from the difference in charge between the folded and unfolded protein, $\Delta Q(pH)$ [26], as

$$\Delta\Delta G^{f\rightarrow uf}\left(pH_{ref}, pH\right) = 2.3RT\int_{pH_{ref}}^{pH}\Delta Q(pH)\,dpH. \tag{2}$$

The protein pK_a values determine the integrand charge difference, T is the temperature, and R is the gas constant. This charge difference between the folded and unfolded protein is calculated for all protein pK_a values by PROPKA.

PROPKA relates a protein pK_a value to its 3D protein structure. The pK_a values of all the ionizable residues in a protein are calculated by determining perturbations, ΔpK_a, to a reference water pK_a value, pK_a^{model}. Perturbations from desolvation ($\Delta pK_a^{\mathrm{DS}}$), protein dipole interactions ($\Delta pK_a^{Q\mu}$), and charge-charge interactions (ΔpK_a^{QQ}) based on the position and orientation of groups surrounding each ionizable residue are included. Thus, the protein pK_a value is expressed as

$$pK_a^{protein} = pK_a^{model} + \Delta pK_a^{model\rightarrow protein} = pK_a^{model} + \Delta pK_a^{DS} + \Delta pK_a^{Q\mu} + \Delta pK_a^{QQ}. \tag{3}$$

The empirical equations used in this expression are described in [17] and its later modifications in [19].

The calculated $\Delta G^{f\rightarrow uf}$, an indication of the change in thermostability, can then be converted to T_m using Gibbs' relationship.

$$\Delta G^{f\rightarrow uf} = \Delta H - T_m \cdot \Delta S. \tag{4}$$

We use the approximate approach of Robertson and Murphy [27] to relate the protein entropy, ΔS, with the number of protein residues. Once we have an estimate for ΔS, we can obtain ΔH from the calculated $\Delta G^{f\rightarrow uf}$ at ambient temperature (298 K), and thereby estimate T_m as the temperature where $\Delta G^{f\rightarrow uf} = 0$.

The $\Delta G^{f\rightarrow uf}$ of the template MtGal was obtained from protein pK_a values predicted with PROPKA using the published crystal structure (PDB code 1HJS and 1.87 Å resolution, [14]). All configurations in the crystallographic asymmetric unit and side-chain alternative conformations (4 × 2 for MtGal) were taken into account. These pK_a values were used to identify stabilizing groups of residue in MtGal. The groups (including their interaction partners) were then subsequently 'moved combinatorially' into AaGal by *in silico* modelling and the total $\Delta G^{f\rightarrow uf}$ of the mutated protein was re-evaluated. Mutants with significant elevated $\Delta G^{f\rightarrow uf}$ were constructed by site-directed mutagenesis.

2.7. Identification of mutation sites by a consensus approach

The multiple sequence homology alignment was performed by first collecting all the sequences of AaGal homologues of 40% or higher sequence identity. Subsequently, a multiple sequence alignment of all the so obtained sequences was created using ClustalW2 [28], the alignment columns were automatically scanned and the positions where the original AaGal amino acid was under represented identified. Mutants of AaGal were constructed to make a specific position more alike the majority of known homologues.

3. Results

3.1. Generation of mutants

The temperature optima for activity are 50 °C at pH 3.5 for AaGal and 65 °C at pH 6.5 for MtGal, despite a high sequence identity of 57% and correspondingly high structural similarity [14]. Residues that potentially contribute to a higher melting temperature, T_m, in MtGal were identified with PROPKA. Additional mutations were identified by a consensus method. The mutations in AaGal identified with PROPKA were D182N, S185D/Q188T, G104D/A156R, while Y295F and G306A were identified by consensus. D58N was predicted to be stabilizing by both methods. These variants were constructed by site-directed mutagenesis. All variants had at least half of wild type activity as measured by the azo-galactan assay, see Table 2.

3.2. Determination of thermodynamic stability

The thermal stability of a protein is typically defined by its melting temperature, T_m, which is the temperature where the transition between the folded and unfolded form of a protein is most pronounced, defined as the maximum point of the heat capacity curve. The PROPKA predicted change in T_m of each single variant is between 0.37 and 1.1 °C, listed in Table 2. Since the small changes in T_m constitute an experimental challenge, DSC was chosen as a technique to measure thermostability. The heat capacity is measured as a function of temperature. All enzymes showed one unfolding peak in the DSC experiments.

In many proteins an irreversible denaturation upon heating is observed. The reversibility of denaturation was therefore checked by reheated scan; a heated sample is cooled and then reheated with the same rate. The end-temperature of each repeated scan was gradually increased, to include the entire range of the endothermic peak. The original curves were reproduced until T_m was reached. This indicates a partially reversible denaturation of AaGal. The irreversible denaturation is most likely a result of aggregation during the heating process. Aggregates were observed in recovered samples.

The irreversible denaturation is believed to follow a reversible step described by the Lumry and Eyring model [29]:

$$N \underset{k_2}{\overset{k_1}{\rightleftharpoons}} U \xrightarrow{k_3} I.$$

The first step is a reversible unfolding of the native catalytically active protein, followed by an irreversible alteration of the unfolded protein to give the final irreversible state. If the rate constant k_3 of the second irreversible step U → I is much higher than the rate constant k_2 of the reversible step of U → N then most of the U molecules will be converted to I, reducing the Lumry and Eyring model to:

$$N \rightarrow I.$$

The denaturation process can then be considered as a one-step process following first-order kinetics. The excess heat capacity determined by DSC can for this process be described as Eq. (5) [30,31]:

$$C_p^{ex} = \frac{\Delta H E_a}{RT_m^2} exp\left(\frac{\Delta H(T-T_m)}{RT_m^2}\right) \times exp\left(-exp\left(\frac{\Delta H(T-T_m)}{RT_m^2}\right)\right). \qquad (5)$$

Thus by fitting the DSC curves to this equation it was possible to obtain the T_m value. The T_m is an apparent T_m-value, which can here be taken as an informative value because irreversible aggregation was shown to be essentially limited to after the melting point. In our study the pH, buffer, scan rate, equilibration time, heating range were kept constant, enabling a comparison of the apparent T_m measured for wild type AaGal and its variants.

The recorded T_m-values are listed in Table 2 along with the PROPKA calculated change in T_m upon mutation. The effect on thermostability of individual mutation was found to be rather limited. The variants Y295F and G306A designed by consensus, the variant D182N designed by PROPKA, and the variant D58N identified by both methods are stabilized compared to WT AaGal. The gain in stability is in the range predicted by PROPKA.

A typical example of recorded denaturation curve is displayed in Fig. 1 along with its curve fit in Fig. 2. The asymmetric appearance of the endothermic peak is characteristic for an irreversible denaturation process.

Table 2
Predicted T_m, experimental T_m and relative activities of AaGal WT and variants. Activity was measured by the azo-galactan assay.

Variant	Activity (%)	DSC	PROPKA	DSC
		T_m (°C)	ΔT_m (°C)	ΔT_m (°C)
WT	100	60.87 ± 0.15	–	–
D58N	67	61.52 ± 0.08	0.75	0.65
Y295F	51	61.21 ± 0.34	0.41	0.34
D182N	88	61.40 ± 0.26	1.1	0.53
G306A	87	62.00 ± 0.01	0.37	1.1
S185D/Q188T	67	60.98 ± 0.27	0.57	0.11
G104D/A156R	82	60.69 ± 0.03	1.1	−0.18

3.3. Thermal denaturation kinetics

Although thermodynamic stability is most easily related to theoretical calculations, the kinetic stability of proteins is often of higher practical significance in applications. Therefore the kinetic stability of stabilized variants was further investigated. The rate of irreversible thermal inactivation was determined for WT, D58N, D182N and G306A β-1,4-galactanase, variants with an observed increase in T_m of 0.5 or above, by measuring the residual activity at 55 °C after incubation. The resulting A vs time (min) curves are shown in Fig. 3. Inspection of semi-logarithmic plots such as the one displayed as example in the inset of Fig. 3 for the WT enzyme, shows that both for WT and variants the data cannot be fit by a single first order process. Rather the data display a characteristic pattern of a biphasic denaturation process, which can be interpreted as an intermediate as in [33]. The biphasic denaturation process can be described as:

$$N \xrightarrow{k_1} X \xrightarrow{k_2} D.$$

N is the native enzyme form, X is an intermediate, D the denatured enzyme form, k_1 and k_2 the rate constants. The specific activity (A) at time (t) can for this process be described as in [33]:

$$A = \left(A_1 - \frac{A_2 k_1}{k_1 - k_2}\right) \exp(-k_1 t) + \left(\frac{A_2 k_1}{k_1 - k_2}\right) \exp(-k_2 t). \qquad (6)$$

A_1 and A_2 are the specific activities of the native enzyme N and the intermediate X, respectively. The rate constants were obtained by

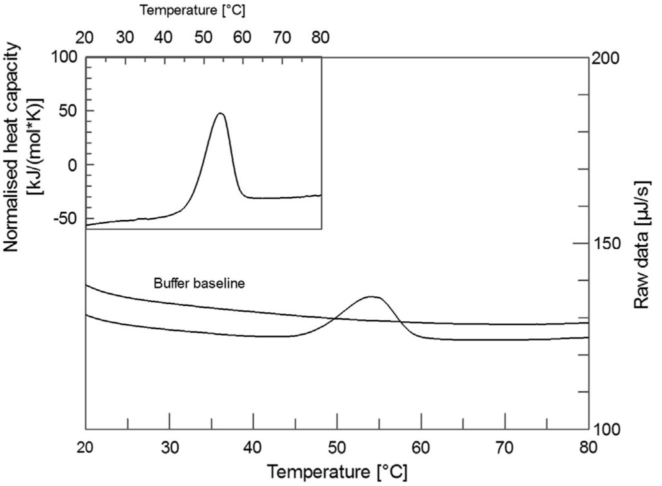


Fig. 1. Schematic DSC curves of 0.5 mg/ml WT AaGal in 50 mM sodium acetate buffer, pH 4.5 (sample) and buffer alone. Raw data [μ cal/s] was measured as a function of temperature [°C] with buffer/buffer and sample/buffer in the cells. The inset shows the appearance of a normalized DSC curve after subtraction of the buffer/buffer curve from the sample/buffer curve and protein concentration normalization. This figure was made with Grafit 7.0.0 [32].

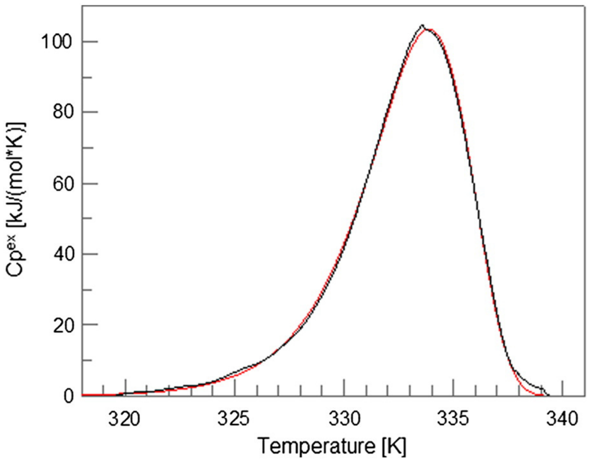


Fig. 2. Fit of a DSC thermogram to an irreversible model. The fit line is given in red and the experimental data in black. The experimental data given for 0.5 mg/mL WT AaGal in 50 mM sodium acetate buffer, pH 4.5 has been normalized according to the enzyme concentration, subtracted the buffer scan values and integrated baseline. The fit was made with Grafit 7.0.0 [32].

non-linear fitting of the data to Eq. (6) in Grafit 7.0.0 [32]. They are shown in Table 3, together with the half-lives calculated by:

$$t_{½} = \frac{\ln(2)}{k}. \tag{7}$$

The inactivation rate constants k_1 and k_2 were 0.0390 min^{-1}, 0.193 min^{-1}, 0.0405 min^{-1}, 0.0158 min^{-1} and 0.0514 min^{-1}, 0.231 min^{-1}, 0.0510 min^{-1}, 0.0291 min^{-1} corresponding to half-lives $t_{½,1}$ and $t_{½,2}$ of 17.8 min, 3.6 min, 17.1 min, 43.9 min and 13.5 min, 3.0 min, 13.6 min, 23.8 min for WT, D58N, D182N and G306A, respectively, also listed in Table 3.

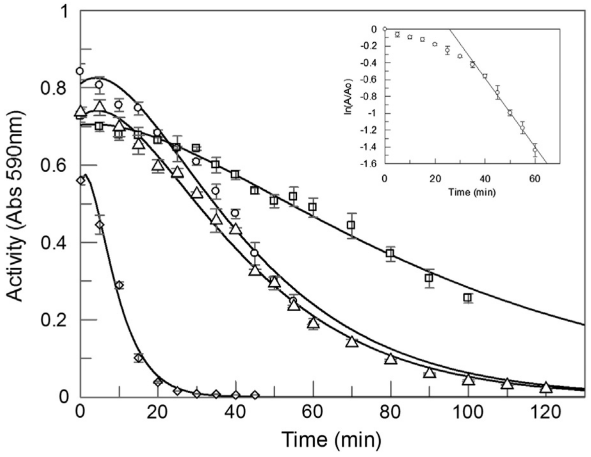


Fig. 3. Thermal inactivation of galactanases. WT AaGal is marked with circles, D182N with triangles, G306A with squares and D58N with diamonds. The enzymes were heated at 55 °C for variable times prior to activity determination by the azo-galactan method (see Materials and Methods section). Specific activity can be expressed as change in A_{590} since the same enzyme amount was used in each assay. In the inset, we show an example of semi-logarithmic plot for WT AaGal, clearly showing that the data cannot be reasonably modelled by a single first order rate constant (to visualize this, the latter part of the plot has been fit to a straight line, clearly showing that this is not a satisfactory model for all data). In the inset, A/A_0 is the residual activity (activity at time *t* divided by activity at time 0). The figure was made with Grafit 7.0.0 [32].

Table 3
Rate constants (k) and half-lives ($t_{½}$) of thermal inactivation of AaGal and variants.

Variant	k_1 (min^{-1})	k_2 (min^{-1})	$t_{½,1}$ (min)	$t_{½,2}$ (min)
WT	0.0390 ± 0.0050	0.0514 ± 0.0077	17.8	13.5
D58N	0.193 ± 0.029	0.231 ± 0.041	3.6	3.0
D182N	0.0405 ± 0.0028	0.0510 ± 0.0041	17.1	13.6
G306A	0.0158 ± 0.0012	0.0291 ± 0.0032	43.9	23.8

4. Discussion

4.1. Rational behind the designed variants

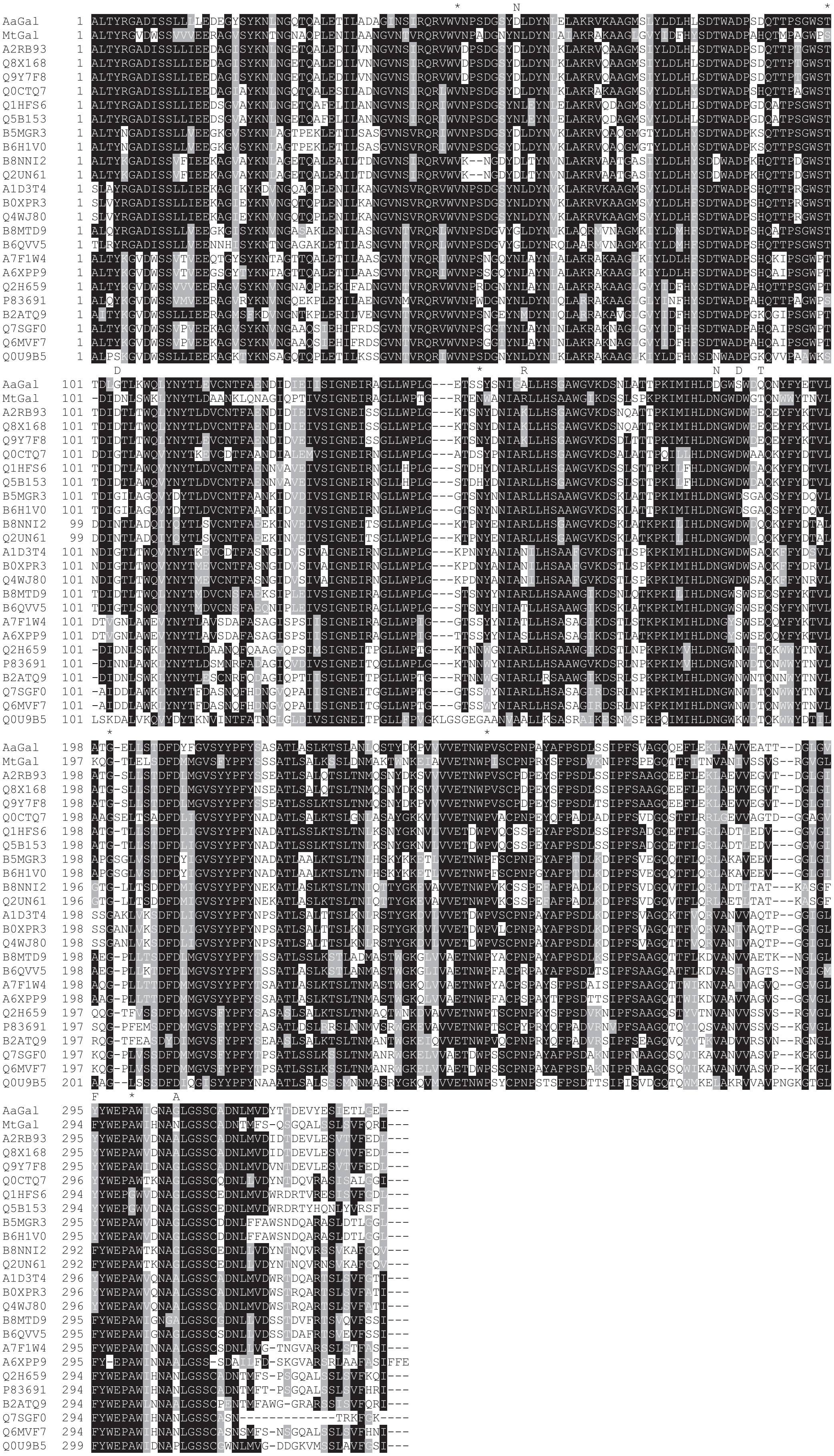
Elevated thermostability is often a desirable protein property, since many industrial applications require enzymes with activity at high temperatures. A great interest has therefore been in isolating proteins from thermophilic or hyperthermophilic organisms, to use directly for these applications or as template protein for protein engineering study. In this study we used the thermophilic Mtgal as template for PROPKA to design new variants of the mesophilic Aagal.

A consensus approach was chosen as complementary protein engineering tool. Mutations in AaGal generated with PROPKA were D182N, S185D/Q188T and G104D/A156R, while Y295F, G306A with the consensus approach, while D58N was identified by both. All mutated residues are shown in the alignment in Fig. 4. The target point residues of D182N, S185D/Q188T, G104D/A156R, D58N and Y295F are all represented in the MtGal. This is not the case for G306A. Gly306 in MtGal is an asparagine, also observed in β-1,4-galactanase from *Humicola insolens*, another thermophilic fungus, but is an Ala in the majority of sequences used for the consensus strategy which do not have a Gly, as shown in Fig. 4.

4.2. Predicted structural impact of the mutations

In order to illustrate the structural impact the mutations might have, the structures of AaGal and MtGal (used as a PROPKA design template) were superimposed, as displayed in Fig. 5. The superimposition reveals that the stabilizing effect of the mutation D182N and D58N might come from removal of the repulsive force between two aspartates (Fig. 5A) and (Fig. 5D), formation of a hydrogen bond in S185D/Q188T (Fig. 5B), formation of a coulombic interaction in G104D/A156R (Fig. 5C). Although suggested by a consensus strategy, the stabilizing effect for Y295F (Fig. 5E) is hard to rationalize in terms of structure, as the mutation is expected to remove a hydrogen bond between Oη inTyr295 and Oγ in Ser43 in the core of the enzyme. However this mutation is also predicted to be slightly stabilizing by PROPKA. Also hard to fully rationalize in terms of structure is the stabilizing effect for G306A suggested by the consensus strategy. A general thermostabilizing effect of mutation of Gly, a flexibility inducing residue, to non-Gly, is often ascribed in the literature to a beneficial entropic effect (the entropy of the unfolded state is reduced) as first proposed by Matthews [35], so this could be an explanation. On the other hand there is no lack in the literature of studies showing mutation from non-Gly to Gly as strongly stabilizing, for example, a single Ser to Gly substitution increases the T_m of a *Clostridium thermocellum* endoglucanase by 7 °C [36]. An additional His to Gly thermostabilizing mutation was later identified for the same protein [37].

In AaGal, Gly306 is located in a loop region approximately 10 Å away from the catalytic active site residue Glu246 at the surface of the protein, see Fig. 5F. The loop also contains Cys310 involved in a disulphide bond to Cys253. One possible specific effect of mutation of Gly306 to an alanine could be to strengthen this disulphide bond by subtle structural alterations, thus improving the overall stability.



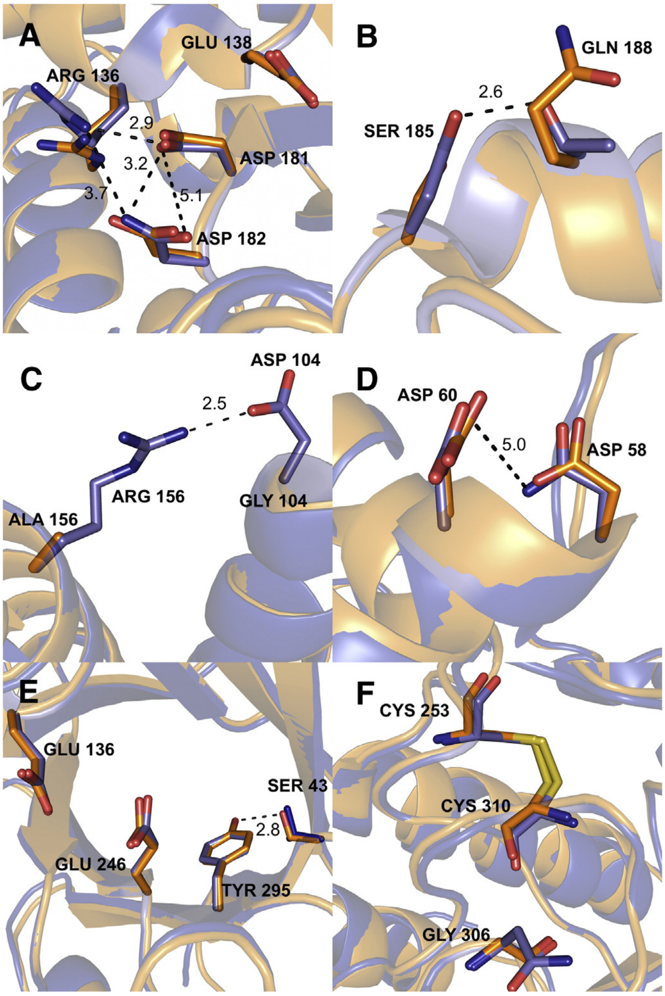


Fig. 5. Structural environment of the generated variants. A) D182N, B) S185D/Q188T, C) G104D/A156R, D) D58N, E) Y295F and F) G306A. AaGal (PDB code 1FOB) is coloured orange, MtGal (used as design and structural template, PDB code 1HJS) in blue). The black dashed lines show selected putative hydrogen bonds. The labelling is according to AaGal. The figure was made with Pymol [38].

4.3. Analysis of the change in stability upon mutation

DSC measurements on the wild-type enzyme and the four single mutants D182N, D58N, G306A and Y295F showed a slight increase in the T_m for the variants in correspondence with the PROPKA predicted values, see Table 2. The standard deviations were however in the range of 0.03 to 0.34 °C, emphasizing the difficulties in measuring very low ΔT_m in proteins. Taking into account these standard deviation we estimate that the increase in G306A is definitely significant, that the increase in T_m observed for D58N is probably significant, but that the increase observed for D182N and Y295F are within experimental error. These results do however show that PROPKA 3.0 might be a valuable tool in predicting the magnitude of the change in T_m upon mutation before experiments are taken on. Also, that the higher thermal stability of MtGal compared to AaGal obviously cannot be assigned to one specific residue, but is a combined effect of many small contributions.

The double mutants S185D/Q188T and G104D/A156R showed no change or decrease in T_m upon mutation, suggesting that the predicted hydrogen bond is not formed between Asp185 and Thr188 in S185D/Q188T, and neither is the predicted salt bridge between Asp104 and Arg156 in G104D/A156R, presumably due to unforeseen interactions with neighbouring residues. This emphasizes the importance of accurate predictions of mutant structures when redesigning enzymes, restricting the use of PROPKA 3.0 alone for this specific application.

There are ways to improve these predictions, such as relaxing the structure around the mutations with molecular dynamics, but the computational effort and general uncertainties increase as these techniques become more complex making it less of an attractive option for screening. A compromise can be reached by using scoring functions and only consider rotation around dihedral angles. This reduces the dimension of the problem, but it again requires accurate energy expressions for the protein structure and it is not obvious that it actually refines the geometry.

4.4. Analysis of the thermal inactivation of AaGal

WT AaGal and its variants G306A, D182N and D58N, with a ΔT_m of at least 0.5 °C were also compared with respect to their thermostabilities by recording time courses of the irreversible decay of enzyme activity at 55 °C. The inactivation temperature of 55 °C was chosen to obtain practically useful inactivation rates. T_m was recorded to be in the range of 60.87 to 62.00 °C. The inactivation time courses were obtained for the enzymes under identical conditions. The decrease of azo-galactan activity (A) for each variant was plotted as a function of preincubation time. The so obtained activity vs time curves displayed a biphasic deactivation pattern, see Fig. 3, more clearly seen in semilogarithmic plots (shown in the inset of Fig. 3 for WT).

The rate constant k_1 at 55 °C is approximately 2.5 times lower and k_2 is approximately 1.8 times lower of G306A compared to that of WT AaGal, see Table 3, (k_1 is 0.0390 min^{-1} and 0.0158 min^{-1}, k_2 is 0.0514 min^{-1} and 0.0219 min^{-1} corresponding to half-lives of $t_{½,1}$ of 17.8 min and 43.9 min and $t_{½,2}$ of 13.5 min and 23.8 min, for WT and G306A, respectively), implying that the stability indeed has been increased upon mutation. The mutation of Asp182 to an asparagine did however not change the thermal irreversible inactivation rate, implying that this mutation do not alter the thermostability of the enzyme. The mutation of Asp58 to an asparagine, was observed to increase the rate constants k_1 and k_2 from 0.0390 min^{-1} and 0.0.0514 min^{-1} to 0.193 min^{-1} and 0.231 min^{-1} and thereby lowering the thermostability of the enzyme significantly.

5. Conclusions

In this study four single point mutations were identified giving a very small increase in T_m, one designed by PROPKA and three by a consensus approach. Given the experimental errors, the increases only seem significant for the D58N and G306A variants. However, only the G306A variant gives also rise to an increase in the lifetime of the enzyme at an elevated temperature. Interestingly one of the mutations which was mildly stabilizing in terms of T_m, D58N, shows a drastically reduced half-life, highlighting the importance of kinetic as well as thermodynamic stability for practical applications.

Some important lessons were learnt in the course of this study. Variants involving a double mutation, where the introduced residues were supposed to interact, failed to produce the desired effect, probably because the relatively simplistic approach used here does not correctly

Fig. 4. Multiple sequence alignment of fungal β-1,4-galactanases. The sequences are labelled using the UniProtKB database identifier (http://www.uniprot.org/) except for AaGal and MtGal (UniProtKB P48842 and P83692, respectively), which are indicated with their name. The signal peptide is not included in the sequences. Residues mutated in AaGal in this study (D58, G104, A156, D182, S185, Q188, Y295 and G306) are indicated by the new residue type above the mutated residue in the sequence of AaGal. Asterisks indicate residues 50, 100, 150, 200, 250 and 300 in the AaGal sequence. The alignment was generated with Clustal W2 [28] and visualized with Boxshade [34] with a 0.5 threshold.

predict the variant structures. Without applying more sophisticated approaches to structure prediction of variants, which are in themselves time consuming, we suggest that a combinatorial approach exploiting the often additive effect of independent amino-acid substitutions maybe effective, as PROPKA performed satisfactorily in predicting the magnitude of the effect on the thermostability of individual substitutions. Although there is a danger that a substitution with adverse effect will cancel out the effect of a favourable substitution, we would recommend that a good strategy would be to predict and experimentally produce combinations of spatially separated triple mutations, giving predicted increase in T_m of at least 2 °C. A sparse matrix approach could be used to reduce the number of variants to be produced and tested and still obtain informative results.

In this study we do identify a mutation, G306A, that gives a moderate but significant increase in stability. This single mutation can increase significantly the lifetime of AaGal, and allow a decrease of enzyme load in applications. Interestingly, a recent study published during processing of this manuscript, reports significant thermostabilization of a *Talaromyces stipitatus* galactanase (TsGal) by the G305A mutation [39]. TsGal shares high sequence identity (67%) with AaGal, and the mutation is in a structurally equivalent loop between the last strand and helix in the $(\beta\alpha)_8$ barrel. However the thermostabilizing mutation in TsGal is not at an equivalent amino acid, but its neighbour, since the WT sequences of the two enzymes are GNGALG for TsGal and GNAGLG for AaGal (with the mutated Gly residues underlined in the respective sequences). This loop seems therefore to be a stability 'hotspot' for GH53 galactanases, which may be related to a conserved disulphide bond in the vicinity of the loop, as discussed in the previous section.

It remains to be established if Ala is the most favourable residue at this hotspot or whether the requirement is simply for a non-Gly. Yet it would seem that there is a sequence bias towards Ala. Since the corresponding residue is an Asn in some already thermostable enzymes like MtGal, if this is a specific effect of a substitution to Ala, insertion in an already thermostable background like MtGal could further improve the stability, and conversely it would be interesting to study the effect of a G306N substitution in AaGal.

Much knowledge of protein thermostability is gained in systems which are fully reversible and with simple first order inactivation kinetics. Unfortunately, the majority of industrially interesting enzymes does not fit these simple models. This is also the case here, as we show that the AaGal unfolding is partly irreversible, and that the kinetics of unfolding are biphasic. There is a strong need of additional experimental data available in the public domain, including the apparently negative results, where predicted effects were not obtained, to build up a solid data basis against which to test theoretical approaches.

Acknowledgement

The authors acknowledge funding from the NABIIT program (from the Danish Council for Strategic Research, grant number 2106-07-0030 to JHJ) and the Program Commission on Sustainable Energy and Environment (grant number 2104-07-0028 to PW). Kristina Mortensen (NOVOZYMES), Johnny Christensen (NOVOZYMES) and Dorthe Boelskifte (University of Copenhagen) are thanked for their technical assistance, Pernille Galberg (University of Copenhagen) for scientific discussions and Tobias Tandrup (University of Copenhagen) for help with final formatting of Fig. 5.

References

[1] Carpita NC, Gibeaut DM. Structural models of primary-cell walls in flowering plants consistency of molecular with the physical-properties of the walls during growth. Plant J 1993;3:1–30.
[2] Lombard V, Golaconda Ramulu H, Drula E, Coutinho P, Henrissat B. The Carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 2014;42:D490–5.
[3] Bjørnvad M, Clausen I, Schülein M, Bech L, Østergaard P et al. (2000) Bacterial galactanases and uses thereof. WO0047711 patent.
[4] De Maria L, Svendsen A, Borchert TV, Christensen LLH, Larsen S et al. (2004) Galactanase variants. WO2004056988 patent.
[5] Michalak MT, Thomassen LV, Roytio H, Ouwehand AC, Meyer AS, et al. Expression and characterization of an endo-1,4-β-galactanase from *Emericella nidulans* in *Pichia pastoris* for enzymatic design of potentially prebiotic oligosaccharides from potato galactans. Enzyme Microb Technol 2013;50:121–9.
[6] Tabachnikov O, Shoham Y. Functional characterization of the galactan utilization system of *Geobacillus stearothermophilus*. FEBS J 2013;280:950–64.
[7] Shipkowski S, Brenchley J. Bioinformatic, genetic, and biochemical evidence that some glycoside hydrolase family 42 β-galactosidases are arabinogalactan type I oligomer hydrolases. Appl Environ Microbiol 2006;72:7730–8.
[8] Delangle A, Prouvost A, Cogez V, Bohin J, Lacroix J, et al. Characterization of the *Erwinia chrysanthemi* Gan locus, involved in galactan catabolism. J Bacteriol 2007;189:7053–61.
[9] Daniel R, Haiech J, Denizot F, Errington J. Isolation and characterization of the lacA gene encoding β-galactosidase in *Bacillus subtilis* and a regulator gene, lacR. J Bacteriol 1997;179:5636–8.
[10] O'Connell Motherway M, Fitzgerald GF, van Sinderen D. Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. Microb Biotechnol 2011;4:403–16.
[11] Ryttersgaard C, Le Nours J, Lo Leggio L, Jørgensen CT, Christensen LLH, et al. The Structure of Endo-β-1,4-galactanase from *Bacillus licheniformis* in complex with two oligosaccharide products. J Mol Biol 2004;341:107–17.
[12] Torpenholt S, Le Nours J, Christensen U, Jahn M, Withers S, et al. Activity of three β-1,4-galactanases on small chromogenic substrates. Carbohydr Res 2011;346:2028–33.
[13] Ryttersgaard C, Lo Leggio L, Coutinho PM, Henrissat B, Larsen S. *Aspergillus aculeatus* β-1,4-Galactanase: Substrate Recognition and Relations to Other Glycoside Hydrolases in Clan GH-A. Biochemistry 2002;41:15135–43.
[14] Le Nours J, Ryttersgaard C, Lo Leggio L, Østergaard PR, Borchert TV, et al. Structure of two fungal β-1,4-galactanases: Searching for the basis for temperature and pH optimum. Protein Sci 2003;12:1195–204.
[15] Otten H, Michalak M, Mikkelsen J, Larsen S. The binding of zinc ions to *Emericella nidulans* endo-β-1,4-galactanase is essential for crystal formation. Acta Crystallogr F 2013;69:850–4.
[16] Alexov E, Mehler EL, Baker N, Baptista AM, Huang Y, et al. Progress in the prediction of pK(a) values in proteins. Proteins Struct Funct Bioinf 2011;79:3260–75.
[17] Li H, Robertson AD, Jensen JH. Very fast empirical prediction and rationalization of protein pKa values. Proteins Struct Funct Bioinf 2005;61:704–21.
[18] Davies M, Toseland C, Moss D, Flower DR. Benchmarking pKa prediction. BMC Biochem 2006;7:18.
[19] Olsson MHM, Sondergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pK(a) predictions. J Chem Theory Comput 2011;7:525–37.
[20] Rostkowski M, Olsson MHM, Sondergaard CR, Jensen JH. Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. BMC Struct Biol 2011;11:6.
[21] Olsson MHM. Improving the desolvation penalty in empirical protein pK(a) modeling. J Mol Model 2012;18:1097–106.
[22] Lehmann M, Pasamontes L, Lassen S, Wyss M. The consensus concept for thermostability engineering of proteins. BBA 2000;1543:408–15.
[23] Godoy-Ruiz R, Ariza F, Rodriguez-Larrea D, Perez-Jimenez R, Ibarra-Molero B, et al. Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments. J Mol Biol 2006;362:966–78.
[24] Christgau S, Sandal T, Kofod LV, Dalbøge H. Expression cloning, purification and characterization of a β-1,4-galactanase from *Aspergillus aculeatus*. Curr Genet 1995;27:135–41.
[25] Huge-Jensen B, Andreasen F, Christensen T, Christensen M, Thim L, et al. *Rhizomucor miehei* triglyceride lipase is processed ans secreted from transformed *Aspergillus oryzae*. Lipids 1989;24:781–5.
[26] Yang AS, Honig B. On the pH-dependence of protein stability. J Mol Biol 1993;231:459–74.
[27] Robertson AD, Murphy KP. Protein structure and the energetics of protein stability. Chem Rev 1997;97:1251–67.
[28] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. Clustal W and clustal X version 2.0. Bioinformatics 2007;23:2947–8.
[29] Lumry R, Eyring H. Conformation changes of proteins. J Phys Chem 1954;58:110–20.
[30] Sanchez-Ruiz JM, Lopezlacomba JL, Cortijo M, Mateo PL. Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. Biochemistry 1988;27:1648–52.
[31] Sanchez-Ruiz JM. Theoretical analysis of Lumry–Eyring models in differential scanning calorimetry. Biophys J 1992;61:921–35.
[32] Leatherbarrow RJ. GraFit version 7. Horley, UK: Erithacus Software Ltd.; 2009.
[33] Violet M, Meunier JC. Kinetic study of the irreversible thermal denaturation of *Bacillus licheniformis* α-amylase. Biochem J 1989;263:665–70.
[34] Hofmann K, Baron M. B BOXSHADE version 3.21. http://sourceforge.net/projects/boxshade/.
[35] Matthews BW, Nicholson H, Becktel WJ. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc Natl Acad Sci U S A 1987;84:6663–7.

[36] Anbar M, Lamed R, Bayer EA. Thermostability enhancement of *Clostridium thermocellum* cellulosomal endoglucanase Cel8A by a single glycine substitution. ChemCatChem 2010;2:997–1003.
[37] Yi ZL, Pei XQ, Wu ZL. Introduction of glycine and proline residues onto protein surface increases the thermostability of endoglucanase CelA from *Clostridium thermocellum*. Bioresour Technol 2011;102:3636–8.
[38] DeLano WL. The PyMOL molecular graphics system, version 1.2r1. Schrödinger: LLC; 2002.
[39] Larsen D, Nyffenegger C, Swiniarska M, Thygesen A, Strube M, et al. Thermostability enhancement of an endo-1,4-β-galactanase from *Talaromyces stipitatus* by site-directed mutagenesis. Appl Microbiol Biotechnol 2014. http://dx.doi.org/10.1007/s00253-014-6244-z.