Appl Microbiol Biotechnol (2013) 97:9727–9735
DOI 10.1007/s00253-013-5184-3

Enhancing RGI lyase thermostability by targeted single point mutations

Inês R. Silva · Dorte M. Larsen · Carsten Jers · Patrick Derkx · Anne S. Meyer · Jørn D. Mikkelsen

Received: 25 June 2013 / Revised: 7 August 2013 / Accepted: 10 August 2013 / Published online: 31 August 2013

Abstract Rhamnogalacturonan I lyase (RGI lyase) (EC 4.2.2.-) catalyzes the cleavage of rhamnogalacturonan I in pectins by β-elimination. In this study the thermal stability of a RGI lyase (PL 11) originating from *Bacillus licheniformis* DSM 13/ATCC14580 was increased by a targeted protein engineering approach involving single amino acid substitution. Nine individual amino acids were selected as targets for site-saturated mutagenesis by the use of a predictive consensus approach in combination with prediction of protein mutant stability changes and B-factor iteration testing. After extensive experimental verification of the thermal stability of the designed mutants versus the original wild-type RGI lyase, several promising single point mutations were obtained, particularly in position Glu434 on the surface of the enzyme protein. The best mutant, Glu434Leu, produced a half-life of 31 min at 60 °C, corresponding to a 1.6-fold improvement of the thermal stability compared to the original RGI lyase. Gly55Val was the second best mutation with a thermostability half-life increase of 27 min at 60 °C, and the best mutations following were Glu434Trp, Glu434Phe, and Glu434Tyr, respectively. The data verify the applicability of a combinatorial predictive approach for designing a small site saturation library for improving enzyme thermostability. In addition, new thermostable RGI lyases suitable for enzymatic upgrading of pectinaceous plant biomass materials at elevated temperatures were produced.

Keywords Protein engineering · Semi-rational · Consensus approach · *Bacillus licheniformis* · *Bacillus subtilis* expression

Introduction

Rhamnogalacturonan I lyases (RGI lyases) (EC 4.2.2.-) are endo-acting enzymes which catalyze the cleavage of α-(1,4) bonds between rhamnose and galacturonic acid in the rhamnogalacturonan I backbone chain of pectins by β-elimination (Mutter et al. 1996). The enzyme action results in the release of oligomers containing Δ4,5 unsaturated galacturonic acid moieties at the non-reducing end. RGI lyase activity was first identified as being part of the plant cell wall-degrading enzyme set of *Aspergillus aculeatus* (Kofod et al. 1994; Mutter et al. 1996). More recently, RGI lyase activity has also been found to be produced by *Bacillus subtilis* (Ochiai et al. 2007), *Cellvibrio japonicus* (McKie et al. 2001), *Erwinia chrysanthemi* (Laatu and Condemine 2003), as well as in a few plants (Naran et al. 2007).

As part of our quest for developing new enzyme processes for selective production of bioactive oligosaccharides from pectinaceous plant biomass streams (Holck et al. 2011; Michalak et al. 2012; Thomassen et al. 2011), we recently expressed a putative RGI lyase encoding gene (PL 11) from *Bacillus licheniformis* (DSM 13) in *Pichia pastoris* (Silva et al. 2011). The monocomponent enzyme protein product was found to be a relatively thermostable RGI lyase having a half-life of 15 min at 61 °C (Silva et al. 2011). However, an even higher thermal stability of the enzyme would be required for its genuine industrial use for upgrading of plant biomass to allow processing at elevated temperatures at which the

Inês R. Silva and Dorte M. Larsen contributed equally to the work.

Electronic supplementary material The online version of this article (doi:10.1007/s00253-013-5184-3) contains supplementary material, which is available to authorized users.

I. R. Silva · D. M. Larsen · C. Jers · 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

P. Derkx
Chr. Hansen A/S, Bøge Allé 10-12, 2970 Hørsholm, Denmark

substrate viscosity and the risk of contamination are lower, the enzymatic rates are higher, mixing is easier, and the substrate solubility is better.

Protein engineering has become an important tool for evolving enzymes to become more fit for industrial applications, e.g., to improve thermal stability (Eijsink et al. 2005). Unfortunately, the diversity of data available regarding the structural basis for enzyme thermostability makes it difficult to define a general strategy for enhancing the thermal stability of an enzyme (Gray et al. 2001). Only in case detailed knowledge about the structure–function–stability relations of an enzyme is available, a rational design can be accomplished immediately. If not, directed evolution or a semi-rational design approach must be employed (Bommarius et al. 2006). Directed evolution of enzymes is based on generating molecular diversity using random mutagenesis in a gene of interest followed by identification of the optimized proteins (usually by high-throughput screening) (Cole and Gaucher 2011). A drawback of this evolution method, however, is the generation of very large libraries, which only partially cover the ample sequence space possible for a protein and which moreover require considerable screening efforts. A semi-rational design strategy allows for the creation of targeted libraries, thereby reducing the library size (Lutz 2010). When it is desired to enhance the thermal stability of an enzyme, a semi-rational design strategy can be based on sequence alignment of homologous proteins for identifying amino acids related to thermostability, i.e., the so-called consensus approach (Davidson 2006). This consensus approach assumes that amino acids occurring more frequently in a certain position within the protein family contribute to the overall stability. Consequently, replacement of non-consensus amino acids by the consensus ones is presumed to enhance thermostability.

Alternatively, several computational techniques exist for predicting functional consequences (e.g., enzyme stability) of amino acid or peptide fragment substitutions in proteins (Bommarius et al. 2006). Two examples of in silico methods are prediction of protein mutant stability changes (PoPMuSiC) and the B-factor iterative test (B-FIT). The PoPMuSiC algorithm can 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 is based on crystallographic B-factor data that can give insight into the protein flexibility. Making a protein more rigid can improve stability, and therefore, amino acids in areas exhibiting high flexibility values (high B-factor) are potential targets for site saturation mutagenesis in optimization of thermostability (Dalby 2011; Reetz et al. 2006).

There is only one example in the literature describing the optimization of the thermostability of a lyase. A pectate lyase was thus improved via an alignment approach producing an enzyme with a 6 °C increase in T_m (apparent melting temperature) and a 23-fold increase in the half-life at 45 °C without compromising the enzyme's catalytic efficiency (Xiao et al. 2008).

Our hypothesis behind the present study was that increased RGI lyase thermostability of an already relatively thermostable enzyme could be evolved by replacing a specific amino acid reinforcing protein packing or protein–protein interactions. Our efforts were first directed toward identifying the most relevant amino acid for this replacement by creating a targeted library for site saturation mutagenesis by combining several methods, namely the consensus approach, PoPMuSiC, and B-FIT modeling. This approach was based on the tenet that the combination of the results of these different methods would allow only a minimal number of substitutions to be prioritized. The objective of this work, therefore, was to assess the applicability of this combinatorial approach for evolving RGI lyase thermostability. The process involved the following: (1) Identification of a satisfactory number of putative amino acids for replacement by this combinatorial prediction approach, construction of site-saturated mutant libraries of each, and a first screen evaluation of activity and thermal stability of each enzyme mutant product obtained. (2) Selection of the most thermostable mutants of the enzyme for a secondary screen to verify the identification of the best new thermostable RGI lyase mutants. (3) Reconfirmation of the thermal stabilities (half-life) of the best mutants after scaled-up enzyme production and purification.

Materials and methods

Prediction of candidates

A multiple sequence alignment approach was used as one of the methods for prediction of possible candidates (Davidson 2006). The RGI lyase wild type (WT), originally identified by blast searching in UniProtKB (Expasy Proteomics Server) as Q65KY4 in the *B. licheniformis* strain DSM 13/ATCC14580 and expressed after codon optimization in *P. pastoris* (Silva et al. 2011), was used as query in BLASTp (National Center for Biotechnology Information) with PSI-BLAST algorithm default settings. Forty-five sequences were selected for ClustalW2 multiple alignment using default parameters (Goujon et al. 2010) and edited in Jalview (Waterhouse et al. 2009). The frequencies of each amino acid in the RGI lyase WT and the consensus sequence, respectively, were calculated. For each position in the alignment, the frequency of the RGI lyase was divided by the frequency of the consensus. If this value was below 1, the amino acid was defined as not conserved and further compared with PoPMuSiC results as well as with the aligned amino acids in the thermostable putative RGI lyase (uniprot A5INB0) from *Thermotoga petrophila* RKU-1 (Zhaxybayeva et al. 2009) included in the multiple alignment. Amino acid positions scoring well in both methods were prioritized for mutagenesis. As a

second approach, the RGI lyase WT was 3D modeled using the HHpred prediction server (Soeding 2005). The model was used to study alpha-helix projections in Pymol version 1.3 for selection of candidates and for predictions of ddG values calculated by PoPMuSiC version 2.0 (Dehouck et al. 2009). B-FIT predictions were used on the basis of B-values from PDB structures of *B. subtilis* RGI lyase YesW (2Z8R, 2Z8S, 2ZUX). The final list of candidates was made by combining and comparing the results of B-FIT, PoPMuSiC, and the consensus approach. The candidate for alpha-helix stabilization was selected only from 3D model projections.

Sequences, strains, and cloning

The gene encoding the *B. licheniformis* RGI lyase WT (Uniprot Q65KY4) (Silva et al. 2011) was codon-optimized for expression in *B. subtilis* (sequence shown in Fig. S1; GenBank accession no. KF373119). Nine site saturation libraries, each in a single amino acid position, were made and inserted between *Nco*I and *Hind*III of pDP66K-PME (Øbro et al. 2009) in frame with the CGTase signal peptide (the signal peptide is shown in Fig. S1). Each of the nine site-saturated mutant libraries was provided in *Escherichia coli* glycerol stocks (Geneart AG, Regensburg, Germany). Purifications were performed using the Zyppy™ Plasmid Miniprep Kit (Zymo Research, Irvine, CA, USA).

For transformation of *B. subtilis* 1A976 (Zhang and Zhang 2011), a single colony was inoculated in 30 ml of Luria broth (LB) (10 g/l tryptone, 5 g/l yeast extract, and 10 g/l NaCl; pH 7.0) supplied with 5 μg/l erythromycin and cultured overnight at 37 °C at 150 rpm. The culture was used to inoculate 30 ml of growth medium (1 ml LB, 3 ml 10× bacillus salts (15 mM $(NH_4)_2SO_4$, 80 mM K_2HPO_4, 44 mM KH_2PO_4, 3.4 mM sodium citrate, 1 mM $MgSO_4$; pH 7.4), 3 ml 10 % L-arabinose, 23 ml dH_2O) to an optical density at 600 nm (OD_{600}) of 0.1. The culture was then grown for 1–1.5 h at 37 °C at 150 rpm until OD_{600} reached 0.16. One milliliter of culture was then transferred to a 12-ml falcon tube and induced with D-xylose (0.3 %) at 250 rpm, 37 °C for 2 h. One microgram of purified plasmid was added, and incubation continued at the same conditions for an additional hour. Cells were plated on LB supplied with 50 μg/ml kanamycin and 15 g/l agar and incubated overnight at 37 °C.

Protein synthesis

For the primary and secondary screens, the expression of the RGI lyase WT was performed in *B. subtilis* 1A976 50 μg/ml kanamycin in deep-well plates. The day after transformation, single colonies were picked and transferred to 125 μl LB at pH 7.5 supplied with 50 μg/ml kanamycin in a 96-low-well plate format covered with airpore sheets. The plates were incubated at 250 rpm, 37 °C for 7 h. Ten microliters of preculture was then transferred to 96-deep-well plates containing 1,100 μl LB (pH 7.5) supplied with 50 μg/ml kanamycin and incubated at 250 rpm at 37 °C.

For "larger scale" production in shake flasks, a single colony of each of the selected mutants from the secondary screen was precultured in 5 ml LB supplied with 50 μg/ml kanamycin (in falcon tubes) at 250 rpm, 37 °C for 16 h. Next, the OD_{600} was adjusted to 0.05 in 100 ml LB, and incubation was done in 500-ml shake flasks at 180 rpm, 37 °C for 24 h. Supernatants containing the secreted RGI lyase were recovered after centrifugation (15 min at 5,000×*g* and 4 °C) and sterile filtrated using 0.22-μm filters. Finally, the supernatants were concentrated to 3 ml using Vivaspin® 20 columns (GE Healthcare).

Purification of RGI WT and mutants

Histidine-tagged RGI lyase WT and its mutants were purified by affinity chromatography. The chromatography was carried out in 96-well filter plates (Polystyrene UnifilterMicroplate, Whatman®) loaded with Ni Sepharose 6 Fast Flow media according to the manufacturer's instructions (GE Healthcare, Uppsala, Sweden). The columns were equilibrated with binding buffer: 40 mM EPPS buffer (4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid), pH 7.4, containing 24 mM imidazole and 500 mM NaCl. Tagged enzymes were eluted with the eluting buffer (500 mM imidazole, 40 mM EPPS buffer (pH 7.4), and 500 mM NaCl).

After production in shake flasks, the WT and mutants were purified in a similar way, but His Spin Trap columns (GE Healthcare) were used instead of plates. In each case, 820 μl of protein solution was collected in the elution buffer. Hereafter, imidazole was removed with Vivaspin® 6 columns (GE Healthcare) using 10 mM EPPS buffer containing 10 % glycerol (pH 8.0) as a washing buffer. Finally, the proteins were quantified by measuring the absorbance at 280 nm and using the extinction coefficient 124,680 M^{-1} cm^{-1} from ProtParam (Gasteiger et al. 2005).

Primary and secondary screens

A standard setup was initially established using the RGI lyase WT. The purified enzyme was distributed into five distinct microtiter plates (60 μl in each), and each plate was incubated at 50 °C in a Biotek Synergy HT microplate reader (BioTek Instruments, Winooski, VT, USA) for 0, 10, 20, 25, and 30 min in duplicate and placed on ice. Residual enzyme activities were measured as described below.

The primary screen of the nine mutant libraries was performed as for the WT. For each library, *B. subtilis* was transformed and a deep-well plate was inoculated with 90 transformants (presumed to cover all 19 mutants) and 6 WT (in known positions) making up a total of 96 samples. The

biological parameter half-life ($t_{1/2}$) was used to evaluate the thermostability of the mutant variants. In both the primary and secondary screens, the first-order rate constant of the thermal denaturation, kd, was obtained from the slope of the plots of the natural logarithm of the residual activity versus time, and half-lives were calculated as ln2/kd. In the secondary screen, the winners of the first screen were picked to a new plate, protein was produced and purified, and the $t_{1/2}$ was measured at 50 °C in four replicates, thereby allowing a direct comparison of the mutants.

Evaluation of mutant enzymes after scale-up production

For thermal stability evaluation, the WT and mutant variants were incubated in an Eppendorf Thermomixer® at 60 °C for different times: 0, 10, 20, 25, and 30 min. The enzymes had a concentration of 0.012 mg/ml in 16 mM EPPS buffer (pH 8.0) and 500 mM NaCl. After incubation, 60-μl aliquots of enzyme were distributed into individual wells in three plates (three replicates) that were placed on ice immediately. The residual activities of the enzyme mutants, obtained after the various incubation times, were measured as described below.

Enzyme activity assays

RGI from potato was from Megazyme (Wicklow, Ireland). The enzyme and substrate were first preincubated separately at 40 °C for 3 min. The reaction mixture (200 μl) comprised 0.750 g/l of potato RGI, 10 mM EPPS buffer (pH 8.0), 2 mM $MnCl_2$, and purified RGI lyase (60 μl WT or mutant). The reactions were carried out for 10 min at 40 °C in microtiter plates. The solutions were pipetted using a MICROLAB® STARLET robotic workstation (Hamilton, Bonaduz, Switzerland), which allowed initiation of the reaction in all 96 wells at the same time. The enzyme activity was based on measuring initial reaction rates by assessing the initial linear absorbance increase at 235 nm, due to the double bond formed in the reaction product (Infinite M200 Tecan). One unit of enzyme activity was defined as the formation of 1 μmol of product per minute at pH 8.0 and 40 °C on 0.750 g of potato RGI per liter. The extinction coefficient used was 5,500 M^{-1} cm^{-1}.

Data deposition

The codon-optimized nucleotide sequence for expression of the *B. licheniformis* DSM13/ATCC14580 RGI lyase wild-type gene in *B. subtilis* has been deposited in the GenBank database (accession no. KF373119).

Results

Selection of candidates for mutagenesis

Nine sites for site-saturated mutagenesis were selected by combining and comparing the results from the consensus approach, PoPMuSiC prediction (ddG and solvent accessibility), B-FIT values (based on *B. subtilis* RGI lyase YesW), and alpha-helix projections (3D stucture) (Table 1).

From a multiple sequence alignment, Gly25, Asp158, Lys243, Gly326, Glu434, and Val541 were selected on the basis of a low occupancy of these amino acids (AAs) compared to the consensus AAs. Then, structural prediction and comparison to a suspected thermostable, putative RGI lyase from *T. petrophila* RKU-1 was performed since this strain is an extreme thermophile and has recently been shown to harbor genes encoding extremely thermostable hydrolytic enzymes (Haq et al. 2012a; Haq et al. 2012b; Santos et al. 2010). Positions were highly prioritized if a consensus amino acid was also found in the sequence at the same position. The selection was thus a compromise between different methods, and several of the selected positions had ratios (consensus approach) that were high compared to the criteria described in Davidson (2006) because the PoPMuSiC predictions with one or more substitutions at these positions with negative ddG values were also taken into account. The structural predictions concerned one or more substitutions with low ddG values (<0) having a relatively high solvent accessibility (see Table 1). The rationale for favoring solvent accessibility was that surface-exposed positions can accommodate more different amino acids and exhibit only few intraprotein interactions, and candidate positions would therefore likely contribute an additive stabilization effect (Perl and Schmid 2002). Finally, loop regions with high flexibility (high B-factor) in the homologous *B. subtilis* RGI lyase YesW were identified. For these regions, PoPMuSiC was also used as a guideline to decide on positions most likely to stabilize the loops leading to selection of Gly55 and Ala67 (Table 1). Finally, position Gly558 associated with the alpha-helix was chosen based on the 3D model of the RGI lyase. Mutation of Gly558 would have the potential to fill a cavity in the protein interior and, depending on the nature of the substitution, make additional interactions that could provide additional rigidity to the enzyme.

Selection of thermostable variants (primary and secondary screens)

The first goal was to establish a screening setup in a 96-well format to select mutants with higher thermostability compared to RGI lyase WT and to minimize the risk of selecting false-positives. Therefore, a preliminary screen containing only RGI lyase WT was performed. The average $t_{1/2}$ at 50 °C for

Table 1 Summary of all nine sites selected for site saturation mutagenesis in the RGI lyase WT

RGI lyase WT position	PoPMuSiC				Consensus		RGI L *T.p*	B-FIT	Exp. results	
	Acc (%)	ΣddG	Mutant	ddG	AA	Ratio	AA	B-factor	AA	Δ%
25Gly	96	−4.8	Asp	−0.62	Asp	0.88	Asp		None	
55Gly[a]	89	−2.5	Pro	−0.75	Ser	0.40	Pro	28.82 (Ala85)	Val	34.9
			Val	−0.56						
			Ser	−0.29						
67Ala[a]	69	−1.7	Pro	−0.75	Lys	0.78	Pro	31.36 (Lys96)	Pro	24.9
158Asp	26	−2.8	Lys	−0.65	Asn	0.29	Asn		Glu	22.6
			Thr	0.02					Ser	14.3
			Asn	0.24					Thr	13.9
			Glu	0.76						
243Lys	36	−6.6	Asp	−0.63	Asp	0.90	Asp		Gly	24.3
			Gly	0.44					Ala	23.3
			Ala	0.14						
			Pro	−0.75						
326Gly	62	−2.9	Ala	−0.48	Glu	0.31	Lys		Glu	23.7
			Glu	−0.15						
			Lys	−0.02						
434Glu	57	−2.3	Pro	−0.68	Pro	0.15	Pro		Tyr	31.4
			Tyr	−0.01					Leu	54.8
			Leu	0.03					Trp	33.2
			Trp	−0.06					Phe	32.5
			Phe	0.07						
541Val	37	−0.2	Pro	−0.16	Pro	0.78	Pro		Glu	25.1
			Tyr	0.37					Tyr	24.9
			Cys	0.68						
			Glu	0.80						
558Gly	1	−2.5	Trp	−0.56	Gly	1	Asp		None	
			Asp	−1.39						

The position of the site is shown together with the respective amino acid (AA). All sites apart from Gly558 were selected on the basis of the consensus approach combined with PoPMuSiC prediction and the aligned putative RGI lyase from *T. petrophila* (RGI L *T.p*). The predicted ddG values for the winners at 60 °C are also shown. The Acc (%) refers to the relative solvent accessibility. Gly558 was selected for stabilization of the alpha-helix. The values of thermostability increment relative to the WT (Δ%) for the (purified) enzymes at 60 °C after scale-up production are also stated

[a] Positions selected on the basis of high B-factor using B-FIT

the WT enzyme plate was 16 min with a coefficient of variation of 15 %. This indicated a sufficiently robust setup with a relatively low degree of variability applicable for mutant screening.

By including a WT in six defined wells, for each of the nine mutated position plates, it was possible to verify that the variability of the WT $t_{1/2}$ was between 10 and 15 % in the actual screen. The $t_{1/2}$ values of the mutants had a somewhat higher variation range because of enhanced thermostability and/or presence of mutants with low or abolished activities (Fig. 1).

The selection criterion (SC) was defined as a $t_{1/2}$ being 20 % higher than the average WT $t_{1/2}$ in the plate. This was considered a robust criterion because it was at least 5 % higher than the inherent variability of the assay and took into account differences due to growth and process conditions for each specific plate. The outcome of the first screen was the selection of 53 different mutants in eight of the nine selected positions. For position Gly558, no mutants were selected due to the half-life being equal to or lower than the selection criterion.

To eliminate any plate-to-plate variation, the selected mutants were transferred to a new plate and analyzed simultaneously in a second screen. The WT (known positions) had an activity of 0.089±0.017 U/ml (with no incubation) and a $t_{1/2}$ at 50 °C of 18 min with a variability of 13 %. Seventy-two

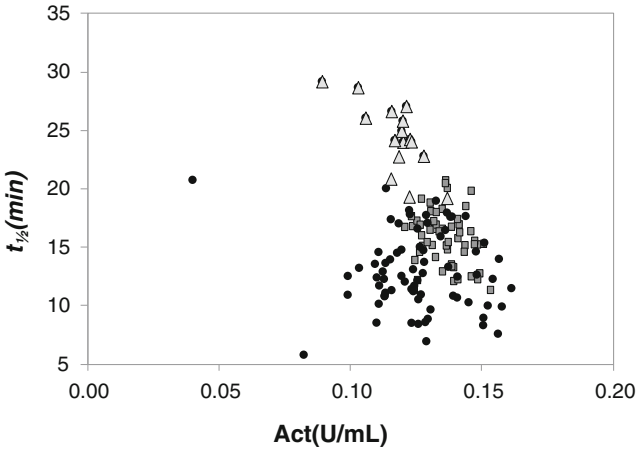

Fig. 1 The activity at 40 °C without prior heat incubation plotted versus the $t_{1/2}$ (min). The 96 WT samples (square), E434 mutant variants (*circle*), and the winners (*triangle*) of the primary screen at the E434 position are shown. E434 was selected to illustrate the higher $t_{1/2}$ variability of the mutants (55 %) compared to the WT (15 %). The inactive mutants are not shown (no activity). It is possible to verify the existence of variants with higher thermostability by comparison to the WT data

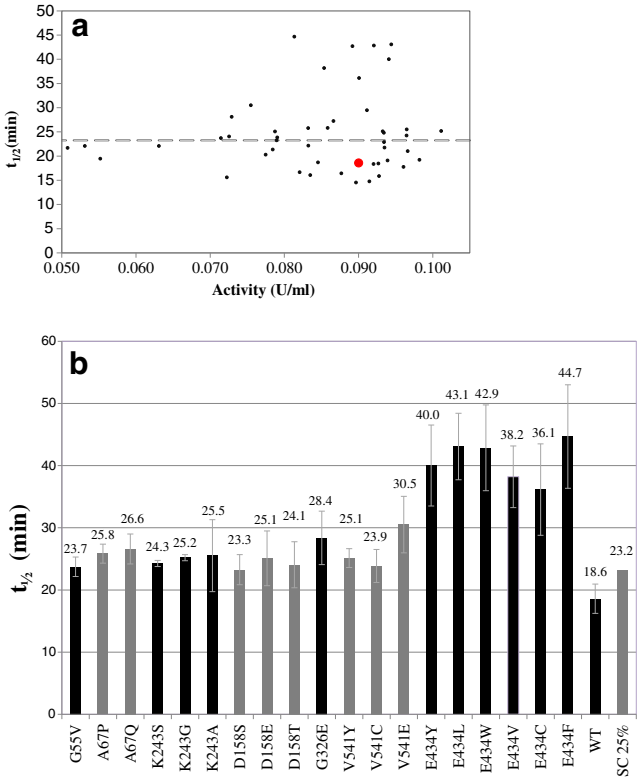

Fig. 2 Results of the secondary screen: **a** The activity of mutants chosen in the first screen at 40 °C (no incubation) versus the $t_{1/2}$ (min) was plotted. The mutants are shown in *black circles* and the WT in *red circle*. The *gray line* represents a half-life value 25 % higher than the WT (SC=23 min). **b** Plot of the half-life of the 19 winner mutants and the respective mutated residue. The half-life of the WT and SC can also be seen

percent of the mutants had a $t_{1/2}$ 20 % higher than the WT, 60 % had a $t_{1/2}$ 25 % higher, and as expected, there were no inactive mutants (Fig. 2a). Mutants displaying $t_{1/2}$>25 % were selected for further analysis.

Even though some of the selected mutants had a slightly lower activity than the WT enzyme, the lowest was 0.072 U/ml, and they were all within the range of uncertainty for the WT activity. With Gly25 being an exception, thermostable mutants were found for all the remaining seven positions. These mutants were sequenced. However, several had the same mutation which resulted in 19 unique mutants (Fig. 2b). The three best of these mutants and their respective half-life increments relative to the WT were Gly326Glu (53 %), Val541Glu (64 %), and Glu434Phe (132 %). In position Glu434 in particular, several promising mutations were obtained (Fig. 2b).

The 16 best performing mutant enzymes were then selected for scale-up, purification, and further analysis. The selection factors were elevated half-life values and good activity levels (at least in the range of the WT).

Significance of imidazole

In the screening setup, the elution buffer was not removed prior to thermal stability measurement. The presence of imidazole did not affect the activity of the WT and mutants, but it influenced thermostability. It also precluded protein quantification and brought inaccuracy to the activity measurements (high background). Therefore, protein purification was followed by buffer change to remove imidazole and NaCl. When the purified WT enzyme was incubated at 50 °C in a thermomixer, 60 % of the activity was still retained after 4 h. In contrast, the same enzyme incubated with 300 mM of imidazole at 50 °C had a $t_{1/2}$ of only 6 min. At 60 °C the $t_{1/2}$ was 20 min, i.e., 1–2 min higher than that at WT 50 °C in the screening (Fig. 2b versus Fig. 3). The half-life was thus generally higher for the WT when the imidazole was removed (Fig. 3 and Table 1). Likewise, all the mutants seemed to benefit from the removal of imidazole since the half-lives at 60 °C were similar to those observed at 50 °C in the initial screen. While imidazole appeared to have a negative effect in all cases, incubation with 500 mM NaCl appeared to enhance thermostability. In the case of WT and E434L, the $t_{1/2}$ increased from 13 to 20 min and from 23 to 31 min, respectively, when the NaCl concentration was changed from 150 to 500 mM. When the amount of salt (NaCl) was increased to 2 M, the half-lives decreased, however: 13 and 19 min, respectively (data not shown).

Identification of the best mutations among the purified enzyme mutants

The analysis of purified enzymes showed lower variability: The replicates of the half-life measurements had a variation of 7 % for the WT and was between 1 and 10 % for the mutants.

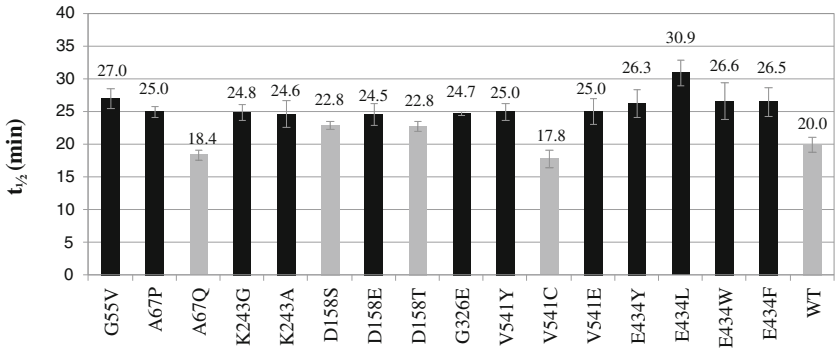

Fig. 3 Results after scale-up production. Plot of half-life values of the 16 mutants selected from the secondary screen and of the WT. Mutants shown in *black* have a $t_{1/2}$ at least 20 % higher than the WT

The mutants were considered relevant if the half-life was higher than 24 min (20 % higher than the WT).

In general, the Glu434 mutations appeared to be most stabilizing with several mutants displaying significantly increased thermostability at 60 °C (Fig. 3 and Fig. S2 in the Supplementary material). The mutant Glu434Leu had a half-life of 30.9 min, corresponding to a ~1.6-fold increase in thermal stability at 60 °C (equal to a 55 % increase) compared to the WT. Gly55Val was the second best mutation with an increase of 35 % (27.0 min). The next best mutations were Glu434Trp, Glu434Phe, and Glu434Tyr with a half-life of 26.6 (33 %), 26.5 (33 %), and 26.3 min (31 %), respectively (Fig. 3). There were also other mutations with a $t_{1/2}$ higher than 24 min: Val541Glu (25.0 min), Ala67Pro (25.0 min), Val541Tyr (25.0 min), Lys243Gly (24.8 min), Gly326Glu (24.7 min), Lys243Ala (24.6 min), and Asp158Glu (24.5 min) (Fig. 3 and Table 1).

In this fashion, we identified 12 mutants which were significantly more thermostable than the already quite thermostable WT RGI lyase without compromising enzyme activity. The specific activity of the WT at 40 °C (with no incubation) was thus 7 U/mg, and the average activity of the mutants was not statistically different from this activity (data not shown).

Discussion

Spontaneous mutations that stabilize proteins above the threshold required for their function in their natural context are rare and usually not maintained during evolution since they do not confer a selective advantage (Wintrode and Arnold 2000). Improvement of enzyme stability by directed evolution demands a high degree of diversity resulting in large libraries and consequently requiring significant screening efforts (Bommarius et al. 2006; Foit et al. 2009).

Our strategy was to increase the chance of finding stabilizing mutations using semi-rational approaches allowing a reduction in library size. In this way, we made a more functional library with a reduced number of variants that were inactive or with lower thermostability. This led to a reduced number of mutants to screen while increasing the probability of finding improved mutants.

The methodology allowed us to select nine amino acid positions to which we applied site saturation mutagenesis. All the mutations were located at the surface except for position Gly558. Mutagenesis at the surface can often be stabilizing when exposed residues are not involved in strong interactions with other residues and hence are often additive (Schmid 2011).

Position Gly558 is located at the core of the protein in the α-helix. The mutagenesis of this position did not render any variant with increased thermostability in the first screen. In fact, some of the variants had a half-life similar to the WT, but most of the mutations were destabilizing (mutants with no or low activity). The protein core is often well packed, and it can be difficult to stabilize by single mutations (Richards 1977). The protein interior is essentially formed by non-polar residues involved in a number of interactions with neighboring residues, and when these are substituted, the probability of destabilizing the existing conformation is high (Loladze et al. 2002).

PoPMuSiC has been applied successfully to improve the thermostability of other proteins, e.g., a feruloyl esterase A from *Aspergillus niger* where the single mutations Asp93Gly and Ser18Phe increased the half-life from 8 min (WT) to 9.4 and 60.5 min at 50 °C, respectively (Zhang and Wu 2011). The majority of positions chosen according to the PoPMuSiC server results had a ΣddG more negative than the average sum for the 45 % of AA where stabilizing mutations were predicted (−1.68) (Table 1). While PoPMuSiC is recognized as the best performing predictor available, its predicting power did not allow a correct ranking of beneficial substitutions, since with the exception of Ala67Pro and Gly326Glu, the mutations predicted as most favorable by PoPMuSiC (lowest ddG) were not the best variants. This was also the case with the consensus approach in that only in one of the seven (Gly326Glu) consensus AAs was the best substitution. It should be noted however that the ratio of WT AA was relatively high in several cases (>0.75 in four positions). Comparison to the potentially thermostable RGI lyase from *T. petrophila* also, only in one case, "predicted" the best substitution (Ala67Pro). Nevertheless, the

combination of the methods for selection of sites to target in combination with site saturation mutagenesis proved a successful strategy.

The use of B-fitter in combination with PoPMuSiC identified Gly55 and Ala67 as candidates due to high flexibility. The mutations Gly55Val and Ala67Pro improved the thermostability, and Gly55Val was the second best mutation overall with a half-life 35 % higher than WT. This method was previously applied to increase the protein stability of a lipase from *B. subtilis* enhancing the inactivation temperature from 48 to 93 °C (Reetz, et al. 2006). Gly is the residue that has the highest conformational entropy while Pro has the lowest, which could indicate that the mutations increase thermostability by decreasing the entropy of unfolding (Li et al. 2005). The position that had the highest number of mutants with significant improvement of thermostability was Glu434. Mutation of the acidic amino acids for hydrophobic residues: tyrosine, phenylalanine, tryptophane, and leucine, increased the thermal robustness of the enzyme. It has been reported that substitution of polar residues at the surface of the protein by uncharged or less polar groups can be a stabilization factor (Schmid 2011). The mutation of a charged residue to a hydrophobic one should lead to a reduction in desolvation penalty. When glutamic acid was mutated for any of the hydrophobic residues, probably the gain in hydrophobic interactions offsets the loss of bonding to water. This suggests that these amino acids could establish stronger interactions with the protein core, thereby increasing protein stability. Charged residues are able to make a number of electrostatic interactions that can affect stability. Examination of the structure revealed the presence of another negatively charged residue (Asp380) within 5.6 Å of the glutamic acid that could possibly induce electrostatic repulsion. It is commonly assumed that electrostatic interactions can be screened by high salt concentrations (Kohn et al. 1997). Since the stability enhancement of E434L over WT was unaffected in the range of 0.15–2 M NaCl, we therefore found it unlikely to be part of the mechanism for the increased stability observed in the mutants. Therefore, the improved thermostability of E434L was probably due to enhanced non-polar interactions. A similar situation is seen for the streptococcal protein Gβ1 domain that is strongly stabilized by changing threonine to isoleucine (Malakauskas and Mayo 1998).

Mutagenesis of ubiquinin reveals stabilization when threonine is mutated to leucine or isoleucine (Loladze and Makhatadze 2005). Also, the higher thermostability when Lys243 was mutated to glycine or alanine could be explained by the gain in hydrophobic interactions. Compared to the exposed lysine, the substitution with glycine or alanine would result in more buried residues. In spite of these results, some of the selected mutations seemed to confer RGI lyase stability by interacting strongly with the solvent, for example, at position 158 the substitution of aspartic acid for glutamic acid. The importance of a stronger polar interaction with the solvent, for protein stabilization, was also illustrated when Gly326 and Val541 were also replaced by glutamic acid. Interestingly, the stability was also increased when Val541 was changed for tyrosine. While testing the impact of electrostatic repulsion, we noticed that the salt concentration conveyed a big impact on the protein stability. When the concentration of NaCl was increased from 150 to 500 mM, the enzyme showed increased thermostability. In solutions with low ionic strength, the salt neutralizes charges at the protein surface eliminating unfavorable electrostatic interactions, the "salting-in" effect (Perl and Schmid 2002). At high salt concentrations, the salt interacts with water resulting in lower solvent availability to interact with the protein. In this way, the protein–protein interactions are enhanced (hydrophobic contacts) resulting in lower protein stability, the "salting-out" effect (Koide et al. 2001; Timasheff 1992). Overall, our methodology based on a targeted library decreased the number of variants to analyze and allowed the identification of 12 beneficial mutations that enhanced the thermostability of the RGI lyase. The most pronounced increase in thermal stability, increasing the RGI lyase half-life at 60 °C from 20 to 31 min (1.6-fold higher), took place with the mutation Glu434Leu. Other mutations, Gly55Val and Glu434Trp, enhanced significantly RGI lyase stability by 35 and 33 %, respectively. Although the increments made by the remaining mutations were smaller, the fact that they occurred at six different positions located at the protein surface might enable combination of the mutations with an additive effect on protein thermostability.

Acknowledgments This study was partially supported by the Danish Strategic Research Council's Committee on Food and Health (FøSu) Project "Biological Production of Dietary Fibres and Prebiotics" (no. 2101-06-0067).

References

Bommarius AS, Broering JM, Chaparro-Riggers JF, Polizzi KM (2006) High-throughput screening for enhanced protein stability. Curr Opin Biotechnol 17:606–610

Cole MF, Gaucher EA (2011) Utilizing natural diversity to evolve protein function: applications towards thermostability. Curr Opin Chem Biol 15:399–406

Dalby PA (2011) Strategy and success for the directed evolution of enzymes. Curr Opin Struct Biol 21:473–480

Davidson AR (2006) Multiple sequence alignment as a guideline for protein engineering strategies. Methods Mol Biol 340: 171–181

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

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

Foit L, Morgan GJ, Kern MJ, Steimer LR, von Hacht AA, Titchmarsh J, Warriner SL, Radford SE, Bardwell JCA (2009) Optimizing protein stability *in vivo*. Mol Cell 36:861–871

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, Totowa, NJ, pp 571–607

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

Gray K, Richardson T, Kretz K, Short J, Bartnek F, Knowles R, Kan L, Swanson P, Robertson D (2001) Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. Adv Synth Catal 343: 607–617

Haq IU, Hussain Z, Khan MA, Muneer B, Afzal S, Majeed S, Akram F (2012a) Kinetic and thermodynamic study of cloned thermostable endo-1,4-β-xylanase from *Thermotoga petrophila* in mesophilic host. Mol Biol Rep 39:7251–7261

Haq IU, Khan MA, Muneer B, Hussain Z, Afzal S, Majeed S, Rahid N, Javed MM, Ahmad I (2012b) Cloning, characterization and molecular docking of a highly thermostable β-1,4-glucosidase from *Thermotoga petrophila*. Biotechnol Lett 34:1703–1709

Holck J, Hjernø K, Lorentzen A, Vigsnæs 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

Kofod LV, Kauppinen S, Christgau S, Andersen LN, Heldt Hansen HP, Dorreich K, Dalbøge H (1994) Cloning and characterization of 2 structurally and functionally divergent rhamnogalacturonases from *Aspergillus aculeatus*. J Biol Chem 269:29182–29189

Kohn WD, Kay MC, Hodges RS (1997) Salt effects on protein stability: two-stranded α-helical coiled-coils containing inter- or intrahelical ion pairs. J Mol Biol 267:1039–1052

Koide A, Jordan MR, Horner SR, Batori V, Koide S (2001) Stabilization of a fibronectin type III domain by the removal of unfavorable electrostatic interactions on the protein surface. Biochemistry 40: 10326–10333

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

Laatu M, Condemine G (2003) Rhamnogalacturonate lyase RhiE is secreted by the out system in *Erwinia chrysanthemi*. J Bacteriol 185:1642–1649

Li WF, Zhou XX, Lu P (2005) Structural features of thermozymes. Biotechnol Adv 23:271–281

Loladze VV, Makhatadze GI (2005) Both helical propensity and side-chain hydrophobicity at a partially exposed site in α-helix contribute to the thermodynamic stability of ubiquitin. Proteins 58:1–6

Loladze VV, Ermolenko DN, Makhatadze GI (2002) Thermodynamic consequences of burial of polar and non-polar amino acid residues in the protein interior. J Mol Biol 320:343–357

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

Malakauskas SM, Mayo SL (1998) Design, structure and stability of a hyperthermophilic protein variant. Nat Struct Biol 5:470–475

McKie VA, Vincken JP, Voragen AGJ, van den Broek LAM, Stimson E, Gilbert HJ (2001) A new family of rhamnogalacturonan lyases contains an enzyme that binds to cellulose. Biochem J 355:167–177

Michalak M, Thomassen LV, Roytio H, Ouwehand AC, Meyer AS, Mikkelsen JD (2012) 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. Enzym Microb Technol 50:121–129

Mutter M, Colquhoun IJ, Schols HA, Beldman G, Voragen AGJ (1996) Rhamnogalacturonase B from *Aspergillus aculeatus* is a rhamnogalacturonan α-L-rhamnopyranosyl-(1→4)-α-D-galactopyranosyluronide lyase. Plant Physiol 110:73–77

Naran R, Pierce ML, Mort AJ (2007) Detection and identification of rhamnogalacturonan lyase activity in intercellular spaces of expanding cotton cotyledons. Plant J 50:95–107

Øbro J, Sorensen T, Derkx P, Madsen CT, Drews M, Willer M, Mikkelsen JD, Willats WGT (2009) High-throughput screening of *Erwinia chrysanthemi* pectin methylesterase variants using carbohydrate microarrays. Proteomics 9:1861–1868

Ochiai A, Itoh T, Kawamata A, Hashimoto W, Murata K (2007) Plant cell wall degradation by saprophytic *Bacillus subtilis* strains: gene clusters responsible for rhamnogalacturonan depolymerization. Appl Environ Microbiol 73:3803–3813

Perl D, Schmid FX (2002) Some like it hot: the molecular determinants of protein thermostability. ChemBioChem 3:39–44

Reetz MT, Carballeira JD, Vogel A (2006) Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angew Chem Int Ed 45:7745–7751

Richards FM (1977) Areas, volumes, packing, and protein structure. Annu Rev Biophys Bioeng 6:151–176

Santos CR, Squina FM, Navarro AM, Ruller R, Prade R, Murakami MT (2010) Cloning, expression, purification, crystallization and preliminary X-ray diffraction studies of the catalytic domain of a hyperthermostable endo-1,4-β-D-mannanase from *Thermotoga petrophila* RKU-1. Acta Crystallogr F66:1078–1081

Schmid FX (2011) Lessons about protein stability from *in vitro* selections. ChemBioChem 12:1501–1507

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 21:951–960

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

Timasheff SN (1992) Solvent effects on protein stability. Curr Opin Struct Biotechnol 2:35–39

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191

Wintrode PL, Arnold FH (2000) Temperature adaptation of enzymes: lessons from laboratory evolution. Adv Protein Chem 55:161–225

Xiao Z, Bergeron H, Grosse S, Beauchemin M, Garron M, Shaya D, Sulea T, Cygler M, Lau PCK (2008) Improvement of the thermostability and activity of a pectate lyase by single amino acid substitutions, using a strategy based on melting-temperature-guided sequence alignment. Appl Environ Microbiol 74:1183–1189

Zhang SB, Wu ZL (2011) Identification of amino acid residues responsible for increased thermostability of feruloyl esterase A from *Aspergillus niger* using the PoPMuSiC algorithm. Bioresource Technol 102:2093–2096

Zhang X, Zhang YP (2011) Simple, fast and high-efficiency transformation system for directed evolution of cellulase in *Bacillus subtilis*. Microb Biotechnol 4:98–105

Zhaxybayeva O, Swithers KR, Lapierre P, Fournier GP, Bickhart DM, DeBoy RT, Nelson KE, Nesbø CL, Doolittle WF, Gogarten JP, Noll KM (2009) On the chemiric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. PNAS 106: 5865–5870