

RESEARCH ARTICLE

Open Access

# Mutagenesis of *Trichoderma reesei* endoglucanase I: impact of expression host on activity and stability at elevated temperatures

Harshal A Chokhawala<sup>1,2</sup>, Christine M Roche<sup>1,2</sup>, Tae-Wan Kim<sup>1</sup>, Meera E Atreya<sup>1,4</sup>, Neeraja Vegesna<sup>3</sup>, Craig M Dana<sup>1,2</sup>, Harvey W Blanch<sup>1,2\*</sup> and Douglas S Clark<sup>1,2\*</sup>

## Abstract

**Background:** *Trichoderma reesei* is a key cellulase source for economically saccharifying cellulosic biomass for the production of biofuels. Lignocellulose hydrolysis at temperatures above the optimum temperature of *T. reesei* cellulases (~50°C) could provide many significant advantages, including reduced viscosity at high-solids loadings, lower risk of microbial contamination during saccharification, greater compatibility with high-temperature biomass pretreatment, and faster rates of hydrolysis. These potential advantages motivate efforts to engineer *T. reesei* cellulases that can hydrolyze lignocellulose at temperatures ranging from 60–70°C.

**Results:** A B-factor guided approach for improving thermostability was used to engineer variants of endoglucanase I (Cel7B) from *T. reesei* (TrEGl) that are able to hydrolyze cellulosic substrates more rapidly than the recombinant wild-type TrEGl at temperatures ranging from 50–70°C. When expressed in *T. reesei*, TrEGl variant G230A/D113S/D115T (G230A/D113S/D115T Tr\_TrEGl) had a higher apparent melting temperature (3°C increase in  $T_m$ ) and improved half-life at 60°C ( $t_{1/2}$  = 161 hr) than the recombinant (*T. reesei* host) wild-type TrEGl ( $t_{1/2}$  = 74 hr at 60°C, Tr\_TrEGl). Furthermore, G230A/D113S/D115T Tr\_TrEGl showed 2-fold improved activity compared to Tr\_TrEGl at 65°C on solid cellulosic substrates, and was as efficient in hydrolyzing cellulose at 60°C as Tr\_TrEGl was at 50°C. The activities and stabilities of the recombinant TrEGl enzymes followed similar trends but differed significantly in magnitude depending on the expression host (*Escherichia coli* cell-free, *Saccharomyces cerevisiae*, *Neurospora crassa*, or *T. reesei*). Compared to *N.crassa*-expressed TrEGl, *S. cerevisiae*-expressed TrEGl showed inferior activity and stability, which was attributed to the lack of cyclization of the N-terminal glutamine in Sc\_TrEGl and not to differences in glycosylation. N-terminal pyroglutamate formation in TrEGl expressed in *S. cerevisiae* was found to be essential in elevating its activity and stability to levels similar to the *T. reesei* or *N. crassa*-expressed enzyme, highlighting the importance of this ubiquitous modification in GH7 enzymes.

**Conclusion:** Structure-guided evolution of *T. reesei* EGI was used to engineer enzymes with increased thermal stability and activity on solid cellulosic substrates. Production of TrEGl enzymes in four hosts highlighted the impact of the expression host and the role of N-terminal pyroglutamate formation on the activity and stability of TrEGl enzymes.

**Keywords:** *Trichoderma reesei*, Endoglucanase I, B-factor, Cellulase, Thermostability, Cell-free synthesis, *Saccharomyces cerevisiae*, *Neurospora crassa*, Pyroglutamate, Glutaminyl cyclase

\* Correspondence: blanch@berkeley.edu; clark@berkeley.edu

<sup>1</sup>Energy Biosciences Institute, University of California, Berkeley, CA 94720, USA

<sup>2</sup>Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

Full list of author information is available at the end of the article

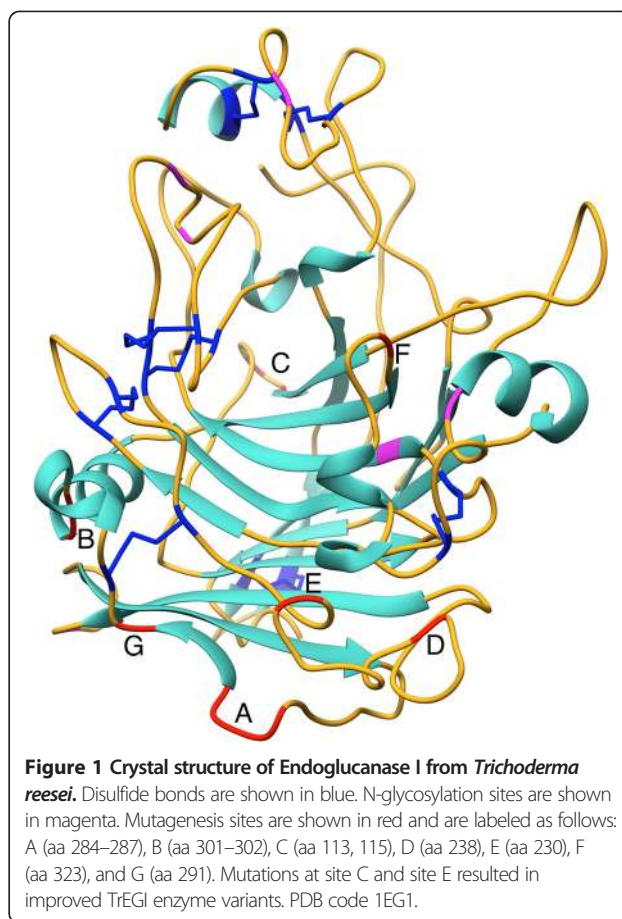
## Background

Efficient saccharification of cellulosic biomass to fermentable sugars is necessary for economical production of second generation biofuels. Several enzymes (exoglucanases, endoglucanases, oxidative enzymes, and  $\beta$ -glucosidases) act in concert to hydrolyze cellulose to glucose. The industrially employed filamentous fungus *Trichoderma reesei* has been evolved for cellulase production and is capable of secreting large quantities of cellulolytic enzymes at a relatively low cost [1,2]. These include two exoglucanases, Cel7A and Cel6A (CBH I and CBH II), which represent 50% and 20% of the total cellulase content, respectively, four endoglucanases, Cel7B (EG1), Cel5A (EGII), Cel12A (EGIII), and Cel45A (EGV) that represent approximately 15%, 10%, 1%, and <1% of the total cellulase content, respectively, and several lytic polysaccharide monooxygenases [3,4].

Lignocellulose hydrolysis at relatively high temperatures (above 50°C, which is near the optimum for many fungal cellulases) may offer potential advantages, including reduced solution viscosity at high-solids loadings (>20 wt%), lower risk of microbial contamination during saccharification, greater compatibility with high-temperature pretreatments, and potentially faster rates of hydrolysis [5]. The short half-lives of *T. reesei* cellulases at temperatures above 50°C, together with very low expression levels of thermophilic cellulases (typically less than 100 mg/L, compared to over 150 gm/L for *T. reesei* cellulases) [6], motivates the development of thermostable *T. reesei* cellulases that can hydrolyze lignocellulose efficiently at temperatures beyond 50°C. *T. reesei* endoglucanase I (TrEGI) is known to randomly cleave internal cellulosic bonds, thereby creating shorter cellulosic chains. Recent studies on optimizing the components of cellulase systems have underscored the importance of having TrEGI comprise a high fraction of the cellulase mixture (25–35%) in order to efficiently hydrolyze pretreated lignocellulosic biomass [7,8].

TrEGI has a bimodular structure with a 375-amino-acid (aa) catalytic domain (CD, Figure 1) attached to a 35-aa carbohydrate-binding module (CBM) via a 26-aa linker. TrEGI has 11 disulfide bonds (8 in the CD [highlighted in blue in Figure 1] and 3 in the CBM), 6 N-glycosylation sites in the CD (highlighted in magenta in Figure 1), and a linker region rich in serine and threonine that are potential O-glycosylation sites. It was selected for engineering enhanced thermostability because of its importance in cellulose hydrolysis and the availability of its crystal structure [9].

Although protein engineering has been used to improve the thermostability of numerous cellulases [10–13], there are no reports on improving the thermostability or increasing the optimal temperature of TrEGI. In this study, we employed a B-factor guided approach [14] to engineer thermostable variants of TrEGI using *E. coli* cell-free protein synthesis due its ease and throughput (Figure 2). Many thermostable mutants of TrEGI identified using

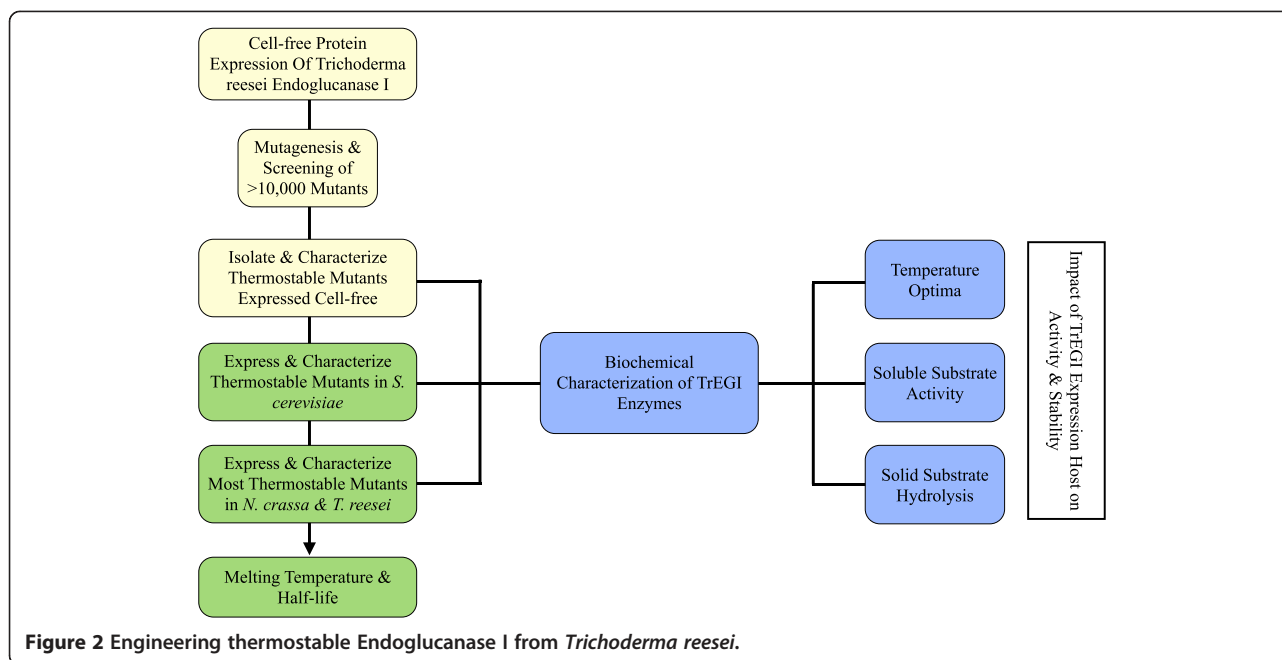


cell-free protein synthesis were subsequently expressed in the fungal hosts *S. cerevisiae*, *N. crassa*, and *T. reesei* to better mimic the properties of the native TrEGI enzyme in terms of folding and glycosylation (Figure 2). TrEGI variants, particularly G230A/D113S/D115T, were more stable compared to their recombinant wild-type versions when both the variant and wild-type were expressed in *S. cerevisiae*, *N. crassa*, or *T. reesei*, indicating that the positive influence of these mutations on stability translates across different expression hosts. However, the *S. cerevisiae*-expressed enzymes were inferior in terms of activity and stability compared to the same enzymes expressed in *T. reesei* or *N. crassa*, which we attribute to the lack of cyclization of the N-terminal glutamine of TrEGI expressed in our *S. cerevisiae* system.

## Results and discussion

### B-factor guided mutagenesis

Our efforts to improve the thermostability of TrEGI entailed a B-factor guided approach (B-FIT method) [14]. The B-factor is a measure of the average deviation of a given residue from its mean position in the crystal structure. Residues having the highest average B-factors (obtained from the crystal structure) correspond to the most flexible sites in the



protein and provide targets for mutagenesis to improve thermostability. This strategy is based on the premise that the thermostability of a mesophilic enzyme can be enhanced by increasing its rigidity at the most flexible sites in its structure. Examples of B-factor guided mutagenesis to increase thermostability include a *Bacillus subtilis* lipase ( $\Delta T_{50}^{60}$  of 42°C) [14] and a *Pseudomonas fluorescens* esterase ( $\Delta T_{50}^{60}$  of 9°C) [15], where  $\Delta T_{50}^{60}$  corresponds to the change in temperature at which the enzyme loses half its activity after an hour of incubation.

Using the crystal structure of the TrEGl catalytic domain (Figure 1), we identified the 20 amino acids with the highest B-factors. These amino acids were grouped in seven sites (Figure 1, in red) comprising one or more amino acids [Site A (aa 284–287), Site B (aa 301–302), Site C (aa 113, 115), Site D (aa 238), Site E (aa 230), Site F (aa 323), Site G (aa 291)], and were chosen for site-saturation mutagenesis. Amino acids with high B-factors that were in close proximity to the N- or C-terminus, disulfide bridges, or N-glycosylation sites in the TrEGl crystal structure were not selected because they could interfere with disulfide-bond formation or lead to false positives due to compensation for non-glycosylation in a bacterial expression system.

A modification of the cell-free protein synthesis protocol developed in our laboratory [16] was used to express TrEGl in an active, soluble form (Additional file 1: Figure S1). The stability of the cell-free expressed TrEGl is comparable to TrEGl expressed in *E. coli* [17], but is considerably reduced compared to the native enzyme expressed in *T. reesei* [18]. Although cell-free

protein synthesis yielded an inferior TrEGl enzyme, we proceeded to use cell-free expression as a tool to rapidly screen mutants of TrEGl, with the assumption that the effect of mutations imparting thermostability will be conserved upon expression in fungal hosts such as *T. reesei*.

The protocol for screening and selection of mutant TrEGl enzymes is shown in (Additional file 1: Figure S2). Mutagenesis and screening of 11,000 mutants at sites A–G generated ~500 mutants with higher activity toward carboxymethyl cellulose (CMC) after pre-incubation at 50°C for 45 min than the wild-type TrEGl expressed using cell-free synthesis. Sequencing the 70 mutants that showed highest activities after the heat treatment revealed that all but six mutants had a mutation at site E (amino acid 230, glycine). Alanine, arginine, serine, threonine, leucine, lysine, glutamic acid, glutamine, and methionine substitutions at this position all imparted improved thermostability. In addition, we found mutations at site C (amino acids D113, D115) that gave rise to more stable variants of TrEGl. Amino acid changes at this site for these TrEGl mutants were D113S/D113T and D115L/D115G. Even though the D113L/D115G mutant had greater thermostability at 50°C, it was not pursued further because it had low specific activity relative to wild-type TrEGl expressed using cell-free synthesis. Since the single mutants (D113S, and D115T) were not isolated during the screening process that covered >95% of all possible amino acid combinations at site C [19], it is highly likely that mutations at both these sites need to be present simultaneously to afford increased stability.

**Activity measurements of round I variants with improved thermostability**

The activities of some of these thermostable TrEGI mutants toward CMC, based on the amount of reducing sugar released before and after exposure to 50°C for various times, are summarized in Figure 3. Combining mutations at G230 with mutations at sites D113 and D115 resulted in triple mutant enzymes (G230X/D113S/D115T, X = K, A, S, R, E, L, T, M; Figure 3) that were more active at 50°C than either of the engineered parents (Figure 3). All the triple mutants showed higher specific activity at 50°C on CMC compared to the wild-type TrEGI expressed using cell-free synthesis at 50°C; mutants G230K/D113S/D115T and G230T/D113S/D115T exhibited ~5 fold improvement in specific activity.

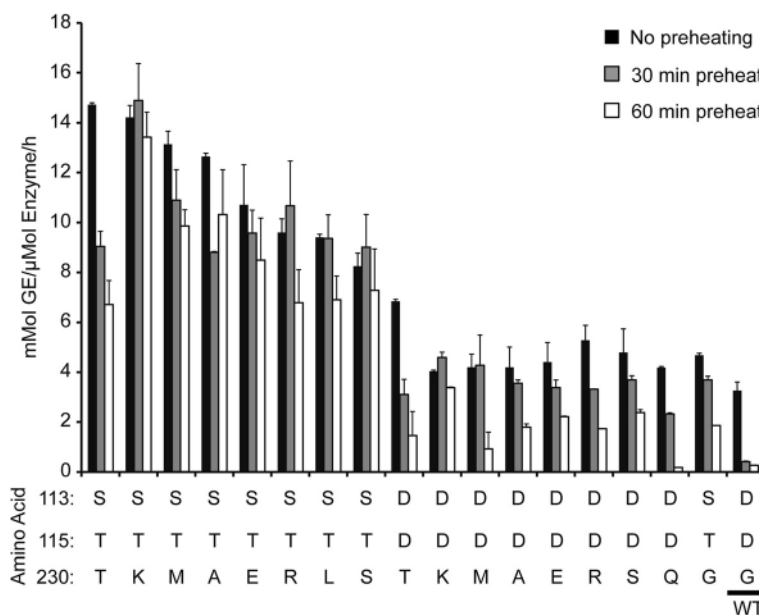
The engineered TrEGI enzymes also showed higher activity on solid substrates: Avicel and ionic-liquid pretreated Avicel (IL-Avicel) (Additional file 1: Figure S3). All of the engineered enzymes showed ~2-fold higher activity for Avicel hydrolysis (Additional file 1: Figure S3A) and ~5-fold improvement for IL-Avicel hydrolysis (Additional file 1: Figure S3B, G230X/D113S/D115T, X = K, A, S, R) at 50°C relative to the wild-type TrEGI expressed by cell-free synthesis. Many of these TrEGI mutants were active against both substrates at 55°C, conditions under which the wild-type TrEGI expressed by cell-free synthesis was essentially inactive.

**Characterization of *Trichoderma reesei* endoglucanase I variants expressed in *S. cerevisiae***

Considering the impact of proper folding and glycosylation on the stability, activity, and structure of proteins

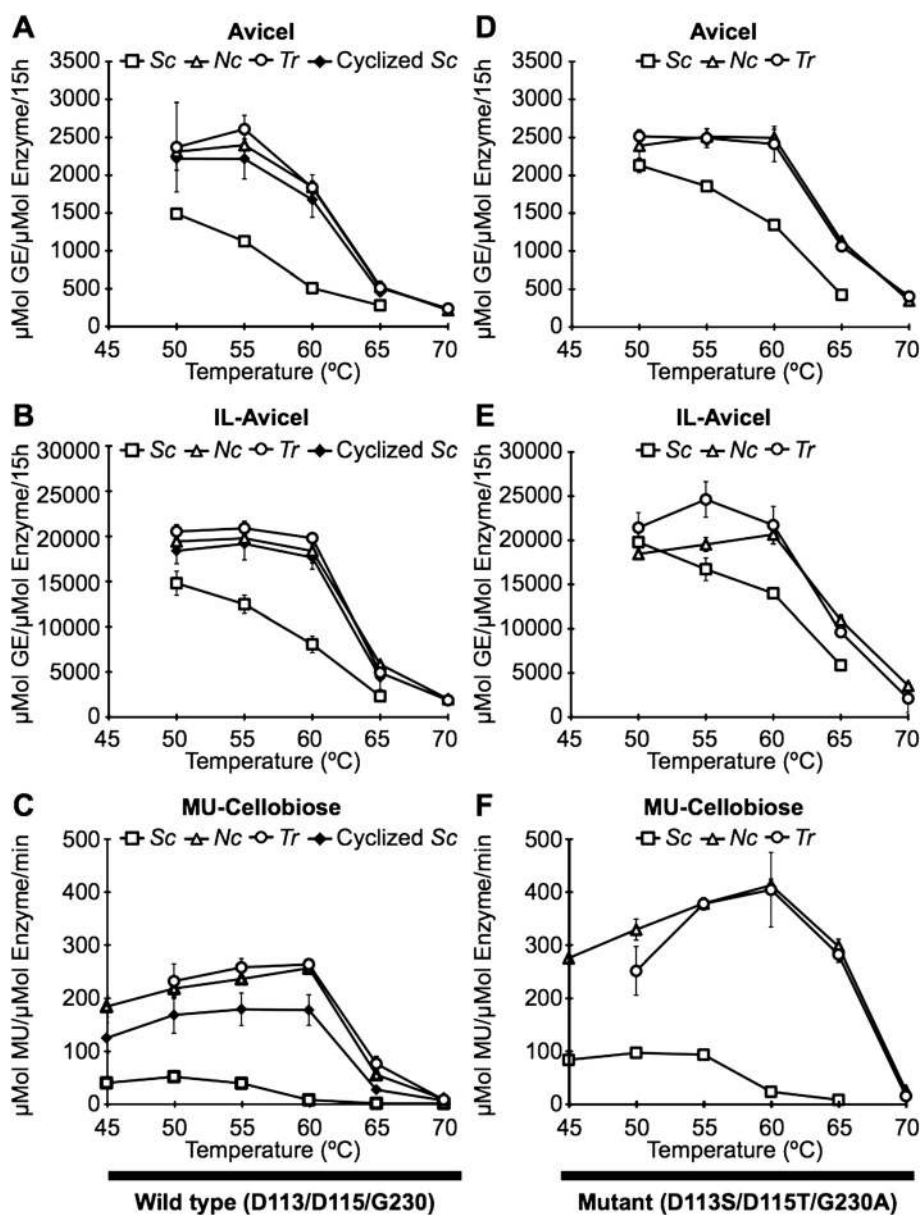
[20-24], we sought to express TrEGI enzymes in a fungal host. Because recombinant expression of proteins in filamentous fungi is typically tedious, time-consuming, and not amenable to high-throughput, we used *S. cerevisiae* [25] as an intermediate fungal expression host to evaluate the stability and activity of the engineered TrEGI triple mutants (enzymes expressed in *S. cerevisiae* are designated Sc\_TrEGI). Of all the triple mutants tested (G230X/D113S/D115T, X = K, A, S, R, E, L, T, M), only five of the TrEGI mutants were expressed in active and soluble forms (G230X/D113S/D115T, X = A, E, R, L, T). The enzymes were secreted from yeast in a hyperglycosylated state (MW of 70–200 KDa) as observed previously [26,27], and thus were treated with an endoglycosidase (PNGase F) that cleaves N-glycans, reducing the molecular weight to 54–56 kDa (TrEGI MW is 46 kDa based on aa sequence, with glycans contributing 8–10 kDa), which nearly matches the molecular weight of the natively-expressed TrEGI enzyme (50–55 kDa) [26].

Many of these TrEGI mutants exhibited higher hydrolytic activities on soluble (MU-cellobiose) and solid substrates (Avicel and IL-Avicel) compared to Sc\_TrEGI (Additional file 1: Figure S4) at 50–65°C. The mutant G230A/D113S/D115T Sc\_TrEGI (Figure 4D and 4E) is > 2.5-fold better than Sc\_TrEGI (Figure 4A and 4B) at hydrolyzing Avicel at 60°C and IL-Avicel at 65°C. In addition, the activity of G230A/D113S/D115T Sc\_TrEGI at 60°C against Avicel and IL-Avicel is similar to that of Sc\_TrEGI at 50°C on these substrates. Hence, B-factor guided engineering of TrEGI provided variants with



**Figure 3** Activity of *T. reesei* endoglucanase I mutants at 50°C on CMC. Error bars represent one standard deviation. TrEGI proteins were expressed using cell-free protein synthesis and were used without any purification.





**Figure 4** Temperature activity profile of wild-type *T. reesei* endoglucanase I (A-C) and G230A/D113S/D115T mutant (D-F) on Avicel (A,D), IL-Avicel (B,E), and MU-Cellobiose (C,F) expressed in *S. cerevisiae*, *N. crassa*, *T. reesei*, and *S. cerevisiae* followed by N-terminal cyclization of the wild-type TrEGI with glutaminy cyclase. Error bars represent one standard deviation. Assays were carried out with purified TrEGI proteins (>95% pure based on SDS-PAGE analysis).

hydrolytic activities on Avicel and IL-Avicel at 60°C similar to the activity of Sc\_TrEGI toward these same substrates at 50°C.

All of the engineered triple mutants expressed in *S. cerevisiae* exhibited higher apparent melting temperatures ( $T_m = 58\text{--}61^\circ\text{C}$ ; measurements were performed as described in Methods using differential scanning calorimetry) than Sc\_TrEGI ( $T_m = 57^\circ\text{C}$ ). Mutant G230A/D113S/D115T Sc\_TrEGI ( $T_m = 61^\circ\text{C}$ , Table 1) showed a  $\sim 4^\circ\text{C}$  increase in apparent melting temperature over

Sc\_TrEGI. No significant difference in half-life at 70°C was observed between Sc\_TrEGI and this mutant (Table 1). It is important to note that our initial efforts included expressing the TrEGI enzymes bearing a C-terminal his<sub>6</sub> tag, which lowered their activity on Avicel (up to 3-fold; compare Additional file 1: Figure S4A and Figure S5), but had little effect on their activity on IL-Avicel (data not shown). This result suggests abrogated binding to Avicel of TrEGI enzymes bearing a C-terminal his<sub>6</sub> tag.

**Table 1 Biochemical characterization of TrEGI expressed in different hosts**

Expression host	Enzyme	T <sub>m</sub> (°C)	t <sub>1/2</sub> (hr)			
			50°C	60°C	65°C	70°C
<i>S. cerevisiae</i>	WT	57	-	-	-	3 (0.4)
<i>S. cerevisiae</i>	G230A/D113S/D115T	61	-	-	-	3.7 (0.5)
<i>S. cerevisiae</i>	WT + Glutaminyl Cyclase	65	-	-	-	-
<i>N. crassa</i>	WT	65	>5 days	70.6 (3.5)	4.0 (0.2)	3.3 (0.1)
<i>N. crassa</i>	G230A/D113S/D115T	68	>5 days	176 (16)	3.3 (0.3)	3.5 (0.4)
<i>T. reesei</i>	WT	65	>5 days	74 (2.5)	5.8 (0.6)	6.2 (0.4)
<i>T. reesei</i>	G230A/D113S/D115T	68	>5 days	161 (5)	6.2 (0.7)	6.4 (0.3)

T<sub>m</sub> is the apparent melting temperature for the enzyme, and t<sub>1/2</sub> is half-life in hours. - = Not determined. Errors are reported in parentheses and represent one standard deviation for measurements. Apparent melting temperature data for Sc\_TrEGI (with and without glutaminyl cyclase treatment) and Tr\_TrEGI are shown in (Additional file 1: Figure S9). Apparent melting temperature data are not shown for other TrEGI enzymes. Half-life measurements for *T. reesei*- and *N. crassa*-expressed enzymes are shown in (Additional file 1: Figure S6 and Figure S7), respectively. Half-life measurements for *S. cerevisiae*-expressed TrEGI enzymes are not shown.

### Characterization of *Trichoderma reesei* endoglucanase I variants expressed in the filamentous fungi *N. crassa* and *T. reesei*

Sc\_TrEGI showed a much lower apparent melting temperature (T<sub>m</sub> = 57°C) and temperature optimum (50°C) compared to the native TrEGI (T<sub>m</sub> ~65°C, T<sub>opt</sub> 60°C) [18], which could be due to non-optimal folding and/or aberrant glycosylation of TrEGI when expressed in *S. cerevisiae*. In order to investigate these possibilities, and to compare the properties (activity and stability) of the engineered TrEGI enzymes to the native TrEGI, we pursued TrEGI expression in *T. reesei*. Thus the wild-type TrEGI and the best mutant (G230A/D113S/D115T TrEGI) were expressed in *T. reesei* QM9414. In addition, our interest in evaluating alternative filamentous fungal systems for cellulase expression prompted us to express TrEGI enzymes (wild-type, G230A/D113S/D115T, and G230E/D113S/D115T) in the filamentous fungus *N. crassa*, which is closely related to *T. reesei*. TrEGI enzymes (wild-type or G230A/D113S/D115T) expressed in *T. reesei* and *N. crassa* were found to have identical temperature optima on MUCellobiose (Figure 4C, 4F), and apparent melting temperatures (Table 1), which in the case of wild-type enzyme (Tr\_TrEGI and Nc\_TrEGI) is also identical to the values reported previously for the native TrEGI. Furthermore, little difference was observed between their activities on solid substrates Avicel (Figure 4A) and IL-Avicel (Figure 4B), thus validating *N. crassa* as a suitable surrogate of *T. reesei* for TrEGI production.

G230A/D113S/D115T Tr\_TrEGI showed similar or higher activity on solid substrates (Figure 4D and 4E) than Tr\_TrEGI when assayed at 50–70°C (Figure 4A and 4B); specifically, a ~2-fold increase in activity was observed at 65°C. In addition, G230A/D113S/D115T Tr\_TrEGI showed similar hydrolytic activity on solid substrates at 60°C as the Tr\_TrEGI at 50°C. As expected, similar improvements were also observed for *N. crassa*-expressed TrEGI variants.

Hence, B-factor guided engineering of TrEGI provided variants with hydrolytic activities on Avicel and IL-Avicel at 60°C similar to the activity of Tr\_TrEGI toward these same substrates at 50°C.

G230A/D113S/D115T Tr\_TrEGI (T<sub>m</sub> = 68°C), G230A/D113S/D115T Nc\_TrEGI (T<sub>m</sub> = 68°C), and G230E/D113S/D115T Nc\_TrEGI (T<sub>m</sub> = 68°C) showed a 3°C increase in apparent melting temperature (Table 1 and, Additional file 1: Table S1) compared to Tr\_TrEGI or Nc\_TrEGI (T<sub>m</sub> = 65°C). In order to assess whether the thermostable mutants also show improved half-lives over the wild-type TrEGI, half-lives of the thermostable mutant G230A/D113S/D115T TrEGI and the wild-type TrEGI (*T. reesei*- or *N. crassa*-expressed) were measured at 50°C, 60°C, 65°C and 70°C (Table 1 and Additional file 1: Figure S6–S7). Although the half-life for TrEGI enzymes at 50°C was too long (>5 days) to distinguish any gains made at this temperature (data not shown), the G230A/D113S/D115T TrEGI variant showed a >2-fold increase in half-life at 60°C compared to the corresponding recombinant (*T. reesei* or *N. crassa*) wild-type TrEGI (Table 1). The improved half-life (t<sub>1/2</sub> = 161 hr) of the G230A/D113S/D115T Tr\_TrEGI variant at 60°C results in improved or comparable performance of this mutant in biomass saccharification carried out at 60°C compared to the performance of Tr\_TrEGI at 50°C (Figure 4A, B, D and E). We also expect that the increased stability and half-life of the G230A/D113S/D115T Tr\_TrEGI at temperatures such as 60°C will correlate with improved half-life at lower temperatures (≤60°C) compared to Tr\_TrEGI. The half-lives of both these enzymes were significantly shorter (~3–6 hr) (Table 1) at higher temperatures (65°C and 70°C). It is surprising that little or no difference was found between the half-lives of G230A/D113S/D115T TrEGI and the wild-type enzyme (*T. reesei*- or *N. crassa*-expressed) at either 65°C or 70°C (Table 1), which may reflect a shift in the inactivation mechanism(s) with temperature

mediated by the different enzyme structures and/or glycosylation patterns as observed previously with other  $\beta$ -sandwich or multi-domain proteins [28,29].

#### Impact of N-terminal pyroglutamate on activity and stability of *T. reesei* endoglucanase I

TrEGI and G230A/D113S/D115T TrEGI expressed in both *T. reesei* and *N. crassa* were found to have much higher  $T_m$  values (7–8°C difference) than their *S. cerevisiae*-expressed counterparts (Table 1). In addition, they were up to 3.6-fold more active on Avicel (Figure 4A and D) at 50–65°C and up to 2.5-fold more active on IL-Avicel (Figure 4B and E) at 50–65°C compared to their counterparts expressed in *S. cerevisiae*. Given the similarity of molecular weights of the TrEGI glycoforms expressed in *S. cerevisiae* and *N. crassa* (Additional file 1: Figure S8), the lower stability and activity of *S. cerevisiae* TrEGI could potentially be due to non-optimal folding (due to protein modifications or lack thereof) and/or incorrect disulfide formation upon expression in *S. cerevisiae*. In this connection, we investigated the role of N-terminal pyroglutamate observed in the *T. reesei* EGI crystal structure.

The available crystal structure of *T. reesei* EGI (pdb code 1EGI) [7] reveals a buried N-terminal pyroglutamate in the natively-expressed enzyme resulting from the cyclization of the N-terminal glutamine. We hypothesized that the lack of this N-terminal glutamine cyclization (confirmed by N-terminal sequencing of *S. cerevisiae*-expressed TrEGI enzyme) in *S. cerevisiae*-expressed enzyme would most likely result in non-optimal protein structure and contribute to the observed decreased stability and activity compared to the *T. reesei*-expressed enzyme. Glutaminyl cyclases (E.C. 2.3.2.5) are known to catalyze the irreversible lactamization of a N-terminal glutamine to a hydrophobic pyroglutamate in many proteins and peptides [30]. Using a human glutaminyl cyclase enzyme [31] we cyclized the N-terminal glutamine residue of Sc\_TrEGI to a pyroglutamate, as evidenced by a blocked N-terminal sequence, typical of proteins/peptides with N-terminal pyroglutamate [32].

Treating Sc\_TrEGI with the human glutaminyl cyclase substantially increased its activity on solid substrates compared to the corresponding enzyme without glutaminyl-cyclase treatment (Figure 4A-B), and shifted its temperature optimum from 50°C to 60°C on soluble substrates (Figure 4C). The activity of the glutaminyl-cyclase treated Sc\_TrEGI on Avicel and IL-Avicel is similar if not identical to its counterpart expressed in *T. reesei* or *N. crassa* (Figure 4A-B).

The increased stability of Sc\_TrEGI from N-terminal glutamine cyclization was evaluated by measurements of apparent melting temperatures (Additional file 1: Figure S9). Glutaminyl-cyclase treated Sc\_TrEGI ( $T_m = 65^\circ\text{C}$ ) showed

an  $\sim 8^\circ\text{C}$  increase in apparent melting temperature compared to the corresponding uncyclized version (Table 1). The apparent melting temperature of Sc\_TrEGI upon cyclization (Table 1) was identical to that of TrEGI ( $T_m = 65^\circ\text{C}$ ). The large increase in  $T_m$  upon N-terminal glutamine cyclization indicates that the lower stability of the *S. cerevisiae* enzymes is due to the uncyclized N-terminal glutamine. In addition, it is worth noting that the differences in glycosylation between Sc\_TrEGI and Nc\_TrEGI (Additional file 1: Figure S8) are irrelevant to the activity and stability differences between these enzymes in comparison to the cyclization of the N-terminal glutamine. Recently our group also demonstrated a similar increase in stability and activity upon cyclization of N-terminal glutamine to pyroglutamate in *S. cerevisiae*-expressed *Talaromyces emersonii* Cel7A [33]. The ubiquitous presence of the N-terminal pyroglutamate in GH7 enzymes (Cel7A and Cel7B), along with its role in imparting high stability/activity for these enzymes, highlights the importance of N-terminal glutamine cyclization for achieving high-efficiency cellulases.

#### Impact of mutagenesis and expression host on activity and stability of *T. reesei* endoglucanase I

Combinations of TrEGI mutations at amino acids G230, and D113/D115 identified using the B-factor method resulted in improved TrEGI enzymes irrespective of the expression host for the enzyme. In particular, TrEGI mutant G230A/D113S/D115T showed improved activity on solid substrates as well as improved thermostability for each expression host when compared to wild-type TrEGI enzyme expressed in the same host. These mutations may rigidify the mobile portions of the TrEGI enzyme, and thus enhance its stability. In particular, it is likely that the mutation at G230, which is located on a solvent exposed loop, imparts improved stability by reducing the entropy (rigidifying) of the (partially) unfolded TrEGI, a well-known mechanism for improving protein stability by replacing conformationally flexible residues such as glycine (located in loops or secondary structural elements) with more rigid residues such as alanine [34,35]. Removal of unfavorable electrostatic interactions involving negative charges on D113/D115 in the wild-type enzyme could also lead to increased stability [36]. A small difference in mass was observed between Nc\_TrEGI and G230A/D113S/D115T Nc\_TrEGI (mass difference = 0–2 Da, Additional file 1: Figure S8 C/D) and between Sc\_TrEGI and G230A/D113S/D115T Sc\_TrEGI (mass difference = 2–59 Da, Additional file 1: Figure S8 A/B), after accounting for the mass difference resulting from the G230A/D113S/D115T mutation. This shows that there is little or no difference in the extent of glycosylation (total glycan content = 8–10 kDa) between the wild-type TrEGI and the mutant G230A/D113S/D115T TrEGI when they are expressed in the same fungal host. Thus, the improved stability of G230A/D113S/D115T

TrEGI compared to wild-type TrEGI when expressed in either *N. crassa* or *S. cerevisiae* is due to the G230A/D113S/D115T mutation and not due to changes in glycosylation.

The specific activity of TrEGI enzymes (wild-type and G230A/D113S/D115T) on solid substrates was dramatically altered upon expression in different hosts (cell-free, *S. cerevisiae*, *N. crassa*, and *T. reesei*) with *T. reesei* or *N. crassa* producing the most active enzymes on Avicel and IL-Avicel (Figure 4). Glycosylation of TrEGI enzymes renders them more thermostable compared to their cell-free expressed counterparts, which is consistent with previous reports showing the importance of glycosylation in protein thermostability [20-22,24]. The cyclization of N-terminal glutamine to a pyroglutamate in *S. cerevisiae* expressed TrEGI resulted in a TrEGI enzyme with near identical properties to the natively expressed enzyme.

## Conclusions

Using B-factor guided mutagenesis, *T. reesei* endoglucanase I variants were developed that were up to 2-fold more active on solid cellulosic substrates at 65°C than the recombinant wild-type enzyme expressed in the same hosts, and exhibited the same activity at 60°C as the recombinant wild-type enzyme at 50°C. The most stable TrEGI mutant G230A/D113S/D115T showed a ~3°C increase in  $T_m$ , as well as more than 2-fold improved half-life at 60°C over the native TrEGI. Amino acid changes at G230 coupled with the D113S/D115T double mutation in TrEGI gave rise to variants with improved thermostability and activity at 50–70°C, irrespective of their expression host. However, the stability and activity of the TrEGI enzymes varied significantly when expressed using *E. coli* cell extract, *S. cerevisiae*, *N. crassa*, or *T. reesei*. TrEGI variants expressed in *N. crassa* or *T. reesei* yielded the most active and stable TrEGI enzymes and *S. cerevisiae* yielded inferior TrEGI enzymes. Cyclization of N-terminal glutamine to a pyroglutamate in *S. cerevisiae*-expressed TrEGI improved its properties to closely match those of the native enzyme, highlighting the importance of this ubiquitous modification in Cel7B enzymes. In this study we show that a non-glycosylating host can be used to screen and discover thermostability imparting mutations in fungal enzymes that result in improved enzymes irrespective of the protein production host. The choice of host for protein engineering of cellulases may not be as critical as the protein production host, which plays a major role in determining the properties of the expressed enzymes.

## Methods

All experiments were carried out using either purified TrEGI enzymes (when expressed in *S. cerevisiae*, *N. crassa*, or *T. reesei*) or TrEGI enzymes expressed via cell-free protein synthesis without any purification. Purity of

purified enzymes (>95%) was determined using SDS-PAGE analysis with Coomassie blue dye staining.

**Construction of plasmid bearing *T. reesei* endoglucanase I Gene** for the *T. reesei* endoglucanase I (*cel7b*, UniProt No. P07981) synthesized with *E. coli* codon bias was purchased from GenScript (Genscript Corporation, Piscataway, NJ, USA). The structural gene of *cel7b* was PCR-amplified from plasmid pUC57-P07981 by using the forward primer (5'-AAAAAACATATG CAACAACCGGGCACCTCC-3') and reverse primer (5'-AAAAAAGTCGACTTACAGACATTG CGAGTAGTA-3'). PCR product was then cloned into pIVEX2.4d vector (Roche Applied Science) after a double digestion with NdeI and SalI to generate plasmid, pIVEX2.4d-TrEGI.

## Preparation of S30 cell-extract

The cell-extracts were prepared as reported previously [16]. To exhaust the reducing activity of cell-extract, the S30 extract was incubated with 10 mM oxidized glutathione for 2 hrs at 30°C. The residual glutathione molecules were removed by dialyzing the treated extract against 200 volumes of buffer C (buffer A without DTT and 2-ME) for 3 hrs at 4°C. The resulting extract was divided into small aliquots and stored at -80°C before using it for cell-free protein synthesis.

## Site directed mutagenesis of *T. reesei* endoglucanase I

Site directed mutagenesis at amino acids 238, 291, and 323 of *T. reesei* EGI (pIVEX2.4d-TrEGI as template plasmid) was performed using the site directed mutagenesis kit (Stratagene) by following the manufacturers protocol with primers shown in Additional file 1: Table S2. Mutagenesis at sites with amino acids 230, 113 and 115, 284–287, and 301–302 of *T. reesei* EGI (pIVEX2.4d-TrEGI as template plasmid) was performed using Overlap Extension PCR. All PCR reactions were carried out using Pfu DNA polymerase (Stratagene, La Jolla, CA). The plasmids were transformed into electrocompetent XL1 Blue cells and plated. Colonies were picked and transferred into 96-well plates using a colony picker (Qpix) and grown overnight at 37°C in 10% glycerol containing LB medium. For site A, ~3000 colonies were picked, corresponding to ~2% coverage; for all other sites, colonies corresponding to >95% coverage of sequence space were picked and grown overnight in 96-well plates (~3000 colonies for sites B and C, and ~500 colonies for sites D-G).

## Cell-free protein synthesis

Overnight culture bearing the mutant *T. reesei* EGI gene was used as a template for the PCR reaction. TrEGI DNA amplified by PCR using T7 fwd and T7 rev primers served as template for protein synthesis. The standard reaction mixture for cell-free protein synthesis consisted of the



following components in a total volume of 15  $\mu$ L: 57 mM of HEPES-KOH (pH 7.5); 1.2 mM of ATP; 0.85 mM each of CTP, GTP, and UTP; 0.17 mg/mL of *E. coli* total tRNA mixture (from strain MRE600); 90 mM of potassium glutamate; 80 mM of ammonium acetate; 12 mM of magnesium acetate; 34  $\mu$ g/mL of L-5-formyl-5, 6, 7, 8-tetrahydrofolic acid (folinic acid); 2.0 mM each of all 20 amino acids; 2% PEG (8000); 67 mM of creatine phosphate (CP); 3.2  $\mu$ g/mL of creatine kinase (CK); 2.5 mM oxidized glutathione; 2.5 mM reduced glutathione; 16.7  $\mu$ g/mL of PCR product; 27%(v/v) of S30-GroEL/ES cell-extract. The cell-free reaction was conducted at 30°C for 2 hours. For determining the amount of the protein synthesized using the cell-free synthesis, 10  $\mu$ M of L-[U-<sup>14</sup>C] Leucine (11.3 GBq/mmol, Amersham Biosciences) was also added to the cell-free reaction. To analyze the solubility of cell-free expressed protein, the cell-free mixture was centrifuged at 15,000 rpm for 10 min at 4°C. The soluble amount of synthesized protein was determined by analyzing the supernatant of the centrifugation. The amount of each cell-free synthesized cellulase was estimated from the TCA-insoluble radioactivity using a liquid scintillation counter (Tri-Carb 2810TR Liquid Scintillation Analyzer, Perkin-Elmer Inc.).

#### Screening *T. reesei* endoglucanase I mutants for improved thermostability/activity

Two  $\mu$ L of 1 M sodium acetate buffer pH 4.85 was added to 15  $\mu$ L of cell-free expressed TrEGI mutant to lower the pH to 5. TrEGI mutant was then heated to 50°C for 30 mins after which 1  $\mu$ L was spotted on CMC-agar plate (1% CMC, 1.5% agar, 50 mM pH 4.85 Acetate buffer) and incubated overnight (12–16 hr) at 50°C. The CMC-agar plates were stained with 15 mL of 1% congo red for 5 mins. The plates were then analyzed for presence of clearing zones, which represented active enzymes (hits). All experiments were carried out using TrEGI enzymes expressed using cell-free protein synthesis without any purification.

#### Measuring the thermostability of *Trichoderma reesei* endoglucanase I

To measure the thermostability of TrEGI, 10  $\mu$ L of the cell-free synthesized TrEGI (20 pmol) was mixed with 10  $\mu$ L of 200 mM of sodium acetate buffer (pH 4.85). After incubating the enzyme mixture at 50°C with different incubation time points (0, 15, 30, 45, 60 min), 80  $\mu$ L of 1.25% (w/v) CMC containing 100 mM sodium acetate buffer (pH 4.85) was added into the heat-pretreated enzyme mixture to start hydrolysis reaction for 1 hr at 50°C. The reducing sugar concentration was measured by the DNS assay as described previously [16]. All experiments were carried out in triplicate using TrEGI enzymes expressed via cell-free protein synthesis without any purification.

#### Endoglucanase activity measurement on solid cellulosic substrates

All cellulolytic assays for solid substrates were carried out in quadruplicate in 96-well plates in a final volume of 70  $\mu$ L containing 1% (w/v) substrate, 100 mM sodium acetate buffer pH 4.85, 0.2  $\mu$ M of the cellulase at 50–70°C. The 96-well plates were sealed with aluminum foil and incubated without shaking in a thermocycler with a heated top to minimize evaporation. Cellulase activities were measured for Avicel, and ionic-liquid pretreated Avicel (IL-Avicel). The mixtures were incubated at 50–70°C for 15 h after which they were cooled to 4°C prior to measuring the amount of soluble reducing sugar released using the glucose oxidase-peroxidase assay as described previously [16]. IL-avicel was prepared as described previously [16]. All experiments were carried out using either purified TrEGI enzymes (when expressed in *S. cerevisiae*, *N. crassa*, or *T. reesei*) or TrEGI enzymes expressed via cell-free protein synthesis without any purification. Enzyme purity (>95%) was determined using SDS-PAGE analysis with Coomassie blue dye staining.

#### Expression of *T. reesei* endoglucanase I in *S. cerevisiae*

For production of *T. reesei* EG1 and engineered mutants in *S. cerevisiae*, the corresponding genes (*E. coli* codon optimized as described above) were amplified from pIVEX 2.4d vector and were cloned with and without a his6 tag on the C-terminus in pCu424 vector [37]. An engineered  $\alpha$ -factor AppS4 [38] prepro leader sequence was appended onto the N-terminus of the genes to enable secretion of the enzyme. pCu424 containing *T. reesei* EG1 gene was transformed into *S. cerevisiae* strain YVH10 [39] with additional *PMR1* knockout using the LiAc method [40]. For expression, a saturated YPD medium preculture was used to inoculate 1 L selective medium (SC-Trp) and grown for three days at 30°C. The culture was spun down at 4000 g for 5 min and resuspended in YPD medium supplemented with 500  $\mu$ M CuSO<sub>4</sub> for a three-day induction at 25°C. Culture supernatant was concentrated to ~100 mL (10 kDa cut-off) prior to purification. TrEGI bearing the his6 tag was purified from culture supernatant using Ni-affinity chromatography. TrEGI without the his6 tag was purified from culture supernatant as described below for *N. crassa* TrEGI.

#### Deglycosylation of *T. reesei* endoglucanase I expressed in *S. cerevisiae*

Three to five mg of purified *T. reesei* EG1 expressed in *S. cerevisiae* was incubated in pH 7.0, 25 mM phosphate buffer along with 200 units of PNGase F (NEB) at 30°C for 18 h following which an additional 100 units of PNGase F were added and incubated for another 24 h. PNGase F treated *T. reesei* EG1 was purified using gel filtration chromatography.

#### Temperature optima of *T. reesei* endoglucanase I expressed in *S. cerevisiae*, *N. crassa*, and *T. reesei*

To measure the temperature optima, TrEGI and the mutants expressed in *S. cerevisiae* (glycosylated versions) and *T. reesei* were assayed for 10 min at different temperatures (40°C–65°C) using 1 mM MU-cellobiose in 100 mM sodium acetate buffer (pH 4.85) as the substrate in a 100 µL reaction volume in a PCR thermocycler. The reaction was quenched by adding 5 µL of the reaction mixture to 145 µL 110 mM CAPS buffer, pH 10.8. Activity was determined by measuring the fluorescence of methyl umbelliferone released. The TrEGI enzyme concentration ranged from 0.1–0.5 µM for initial-rate measurements. All experiments were carried out in quadruplicate. All experiments were carried out using purified enzymes.

#### Half-life measurements of *T. reesei* endoglucanase I expressed in *S. cerevisiae*, *N. crassa*, and *T. reesei*

TrEGI enzymes (1 µM) were incubated in 100 mM sodium acetate buffer pH 4.85 at 50°C, 60°C, 65°C and 70°C in a PCR thermocycler. Aliquots were withdrawn at different times and incubated at 4°C for 20 min. Endoglucanase activity was measured at 50°C using MU-cellobiose as described above. The enzyme half-life was calculated by fitting the activity data to exponential or linear functions. All experiments were carried out in duplicate. All experiments were carried out using purified enzymes.

#### Cloning, expression and purification of *T. reesei* endoglucanase I in *N. crassa*

*T. reesei* EGI and engineered mutants were expressed in *N. crassa* as described previously [33]. The wild-type TrEGI gene was amplified from *T. reesei* gDNA. The engineered TrEGI mutants were constructed using site-directed mutagenesis kit from Stratagene as described above. TrEGI was purified from culture supernatant in a series of steps. TrEGI was precipitated from the culture supernatant with 65% ammonium sulfate. The supernatant was spun down and the precipitated proteins were resuspended 20 mM Tris–HCl pH 7 buffer and desalted using HiPrep 26/10 desalting column with 20 mM Tris–HCl pH 7 as the running buffer. The protein was then concentrated and loaded on Mono Q column and TrEGI was eluted using 20 mM Tris–HCl pH 7 buffer between 25–75 mM NaCl. TrEGI containing fractions were pooled and further polished using size exclusion chromatography to give pure protein.

#### Cloning, expression and purification of *Trichoderma reesei* endoglucanase I in *T. reesei*

For production of *T. reesei* EGI and engineered mutants a EGI expression vector was constructed by introducing the EGI gene followed by the *N. crassa* cbhI terminator

behind the *cdna1* promoter in pPcdna1 expression plasmid (generous gift from Bernhard Seiboth) [41]. Transformation of *T. reesei* by electroporation was adapted as described previously [42]. Spores of *T. reesei* QM9414 were harvested from a 90 mm potato dextrose agar (PDA) plate and suspended in 1.1 M ice cold sorbitol. Spores were washed twice with 1.1 M ice cold sorbitol, pelleted at 800xg for 4 min, and then re-suspended in 100 µL of 1.1 M ice cold sorbitol. Two micrograms of purified PCR product of the expression cassette using the primers 5'- GGATCCGA GAGCTACCTTACATC -3' and 5'- CGAACTACCTCG CGAACTCG -3' was added to the spore suspension and incubated on ice for 30 min. Electroporation was conducted using a Gene Pulser Xcell (Bio-Rad Laboratories, Inc., Hercules, CA) in 1 mm gap cuvette at 1.6 kV, 600 Ω and 25 µF. Immediately following electroporation, 900 µL of 1.1 M ice cold sorbitol was added and gently mixed. The spore suspension was then added to 9 mL of YPD (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose) and incubated for 12 hr at room temperature. Spores were pelleted at 800xg for 4 min, re-suspended in 1 mL of 1.1 M sorbitol, and plated using the overlay method onto Mandels-Andreotti medium containing 200 µg/mL hygromycin B. Both top and bottom agars contained hygromycin B. Germinated spores were picked onto PDA slants containing hygromycin B. Positive transformants were selected by presence of EGI activity (AZO-CM-Cellulose Assay, Megazyme, Ireland) in the supernatant after growth in Mandels-Andreotti medium supplemented with 3% glucose, in which QM9414 does not secrete native cellulases [41,43] for 3–4 days. Culture broths were filtered through glass microfiber filters (934-AH, Whatman) followed by 0.22 µm PES filters (Corning). TrEGI was purified from culture supernatant as described above for *N. crassa* TrEGI.

#### Mass spectrometry of intact TrEGI protein

Mass Spectra were obtained via an Agilent 6510 Quadrupole Time-of-Flight (Q-TOF) mass spectrometer immediately following external mass calibration. TrEGI protein samples were diluted in water to a final concentration of 0.75 µM and 1 µL was injected using an autosampler. Samples were analyzed with an Agilent HPLC-chip cube using a custom HPLC-chip (SPQ381) containing Zorbax 300A SB-C3 (5 µm) separation (43 mm × 75 µm) and enrichment (4 mm, 40 nL) columns. Proteins were eluted under a gradient of 3% to 95% Solvent B over 9 min. Solvent A was water with 0.1% formic acid; Solvent B was acetonitrile with 0.1% formic acid. Solvent reagents were of LC/MS grade. The Q-TOF mass spectrometer was run under Extended Dynamic Range (2GHz). The HPLC-Chip voltage was 1.85 kV, gas temperature was 300°C, and drying gas flow was 4 L/min. One spectra was acquired per second with 1000 ms/spectrum using a mass range of 300–3,000 m/z. Data analysis was performed using Agilent Mass Hunter Qualitative

Analysis software. Mass Spectra extracted from chromatogram peaks were deconvoluted using the maximum entropy setting with a mass range of 20–80 kDa, mass step of 1 Da, S/N threshold of 30, minimum consecutive charge states of 5, minimum protein fit score of 8, and an average mass of 90% peak height.

### Glutaminyl Cyclase Treatment of *S. cerevisiae* Expressed TrEGI

Lyophilized human glutaminyl cyclase was purchased from Sino Biological (Daxing, China; 13752-H07B) and reconstituted to 0.2 mg/mL. Cyclization was carried out using 0.02 mg/mL cyclase and 50–100  $\mu$ M TrEGI in 50 mM sodium phosphate pH 7.4 for 18–48 h at 30°C. N-terminal edman sequencing was performed at University of California-Davis, proteomics core facility. All experiments were carried out using purified TrEGI.

### Differential scanning calorimetry

The apparent melting temperature of TrEGI (0.5–1.0 mg/mL) in 50 mM sodium acetate, pH 4.85, was measured using a TA Instruments NanoDSC. After equilibration, a temperature schedule from 25°C to 80°C was employed using a ramp rate of 1°C/min. Raw data are presented for some TrEGI variants (Additional file 1: Figure S9) and the apparent melting temperature corresponds to the apex of each peak. For all TrEGI enzymes mentioned herein, melting was irreversible. All experiments were carried out using purified enzymes.

### Additional file

**Additional file 1: Includes Figures S1-S9 and Tables S1-S2.**

### Abbreviations

TrEGI: *Trichoderma reesei* endoglucanase I; EGI: Endoglucanase I; Sc\_TrEGI: Wild-type *Trichoderma reesei* endoglucanase I expressed recombinantly in *Saccharomyces cerevisiae*; Nc\_TrEGI: Wild-type *Trichoderma reesei* endoglucanase I expressed recombinantly in *Neurospora crassa*; Tr\_TrEGI: Wild-type *Trichoderma reesei* endoglucanase I expressed recombinantly in *Trichoderma reesei* QM9414; CD: Catalytic Domain; CBM: Carbohydrate Binding Module; CMC: Carboxy Methyl Cellulose; DSC: Differential Scanning Calorimetry; MU-cellobiose: 4-Methyl Umbelliferone- $\beta$ -D-Cellobioside; PASC: Phosphoric Acid Swollen Celulose; pNPC: *p*-Nitrophenyl- $\beta$ -D-cellobioside; IL-Avicel: Ionic-liquid pretreated Avicel;  $\Delta T_{50}^{60}$ : Change in temperature at which the enzyme loses half its activity after an hour of incubation.

### Competing interests

A patent has been filed for the engineered *Trichoderma reesei* endoglucanase I variants reported herein.

### Authors' contributions

HAC designed and executed the experiments and analyzed the data. CMR transformed *N. crassa* and *T. reesei* with TrEGI genes and expressed the TrEGI proteins. TWK developed the cell-free expression platform used for expressing the TrEGI and performed all TCA-insoluble radioactivity experiments. NV performed cloning of the TrEGI genes in pCSR1:GPD plasmid. MEA performed the mass spectrometry experiments. CMD developed the yeast strain and vector

used for expressing the TrEGI. DSC, HWB, CMR, and HAC wrote the manuscript. All authors read and approved the manuscript.

### Acknowledgements

This work was supported by the Energy Biosciences Institute. We would like to thank Dr. Trevor Starr and Prof. N. Louise Glass for providing the pCSR1:GPD plasmid for recombinant TrEGI expression in *N. crassa*, Dr. Bernhard Seiboth for supplying us with the pPcdna1 vector, and Dr. Jerome Fox for providing the gDNA of *T. reesei*.

### Author details

<sup>1</sup>Energy Biosciences Institute, University of California, Berkeley, CA 94720, USA. <sup>2</sup>Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA. <sup>3</sup>Department of Biotechnology, Indian Institute of Technology, Kharagpur, West Bengal 721301, India. <sup>4</sup>Department of Chemistry, University of California, Berkeley, CA 94720, USA.

Received: 3 December 2014 Accepted: 28 January 2015

Published online: 21 February 2015

### References

- Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons BA, Blanch HW. The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnol Bioeng.* 2012;109:1083–7.
- Cherry JR, Fidantsef AL. Directed evolution of industrial enzymes: an update. *Curr Opin Biotechnol.* 2003;14(4):438–43.
- Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels.* 2012;5(1):45.
- Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J Mol Biol.* 2008;383(1):144–54.
- Viikari L, Alapuranen M, Puranen T, Vehmaanpera J, Siika-Aho M. Thermostable enzymes in lignocellulose hydrolysis. In: Olsson L, editor. *Biofuels*. vol. 108. Berlin: Springer-Verlag Berlin; 2007. p. 121–45.
- Dashtban M, Schraft H, Qin WS. Fungal Bioconversion of Lignocellulosic Residues; Opportunities & Perspectives. *Int J Biol Sci.* 2009;5(6):578–95.
- Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Aslam N, Walton JD. Synthetic Enzyme Mixtures for Biomass Deconstruction: Production and Optimization of a Core Set. *Biotechnol Bioeng.* 2010;106(5):707–20.
- Gao DH, Chundawat SPS, Krishnan C, Balan V, Dale BE. Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. *Bioresour Technol.* 2010;101(8):2770–81.
- Kleywegt GJ, Zou JY, Divne C, Davies GJ, Sinning I, Stahlberg J, et al. The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 angstrom resolution, and a comparison with related enzymes. *J Mol Biol.* 1997;272(3):383–97.
- Heinzelman P, Snow CD, Wu J, Nguyen C, Villalobos A, Govindarajan S, et al. A family of thermostable fungal cellulases created by structure-guided recombination. *Proc Natl Acad Sci U S A.* 2009;106(14):5610–5.
- Liang CN, Fioroni M, Rodriguez-Ropero F, Xue YF, Schwaneberg U, Ma YH. Directed evolution of a thermophilic endoglucanase (Cel5A) into highly active Cel5A variants with an expanded temperature profile. *J Biotechnol.* 2011;154(1):46–53.
- Liu WJ, Zhang XZ, Zhang ZM, Zhang YHP. Engineering of *Clostridium phytofermentans* Endoglucanase Cel5A for Improved Thermostability. *Appl Environ Microbiol.* 2010;76(14):4914–7.
- Voutilainen SP, Boer H, Alapuranen M, Janis J, Vehmaanpera J, Koivula A. Improving the thermostability and activity of *Melanocarpus albomyces* cellobiohydrolase Cel7B. *Appl Microbiol Biotechnol.* 2009;83(2):261–72.
- Reetz MT, Carballeira JD, Vogel A. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew Chem Int Ed.* 2006;45(46):7745–51.
- Jochens H, Aerts D, Bornscheuer UT. Thermostabilization of an esterase by alignment-guided focussed directed evolution. *Protein Engineering Design Select.* 2010;23(12):903–9.
- Kim TW, Chokhawala HA, Nadler D, Blanch HW, Clark DS. Binding Modules Alter the Activity of Chimeric Cellulases: Effects of Biomass Pretreatment and Enzyme Source. *Biotechnol Bioeng.* 2010;107(4):601–11.
- Nakazawa H, Okada K, Kobayashi R, Kubota T, Onodera T, Ochiai N, et al. Characterization of the catalytic domains of *Trichoderma reesei*

- endoglucanase I, II, and III, expressed in *Escherichia coli*. *Appl Microbiol Biotechnol*. 2008;81(4):681–9.
18. Baker JO, Tatsumoto K, Grohmann K, Woodward J, Wichert JM, Shoemaker SP, et al. Thermal-denaturation of *trichoderma-reesei* cellulases studied by differential scanning calorimetry and tryptophan fluorescence. *Appl Biochem Biotechnol*. 1992;34–5:217–31.
  19. Firth AE, Patrick WM. Statistics of protein library construction. *Bioinformatics*. 2005;21(15):3314–5.
  20. Imperiali B, O'Connor SE. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr Opin Chem Biol*. 1999;3(6):643–9.
  21. Beckham GT, Dai Z, Matthews JF, Momany M, Payne CM, Adney WS, et al. Harnessing glycosylation to improve cellulase activity. *Curr Opin Chem Biotechnol*. 2012;23(3):338–45.
  22. Kayser V, Chennamsetty N, Voynov V, Forrer K, Helk B, Trout BL. Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies. *Biotechnol J*. 2011;6(1):38–44.
  23. Jeoh T, Michener W, Himmel ME, Decker SR, Adney WS. Implications of cellobiohydrolase glycosylation for use in biomass conversion. *Biotechnol Biofuels*. 2008;1(1):10.
  24. Adney WS, Jeoh T, Beckham GT, Chou YC, Baker JO, Michener W, et al. Probing the role of N-linked glycans in the stability and activity of fungal cellobiohydrolases by mutational analysis. *Cellulose*. 2009;16(4):699–709.
  25. Dana CM, Saija P, Kal SM, Bryan MB, Blanch HW, Clark DS. Biased clique shuffling reveals stabilizing mutations in cellulase Cel7A. *Biotechnol Bioeng*. 2012;109(11):2710–9.
  26. Penttila ME, Andre L, Saloheimo M, Lehtovaara P, Knowles JKC. Expression of 2 *trichoderma-reesei* endoglucanases in the yeast *saccharomyces-cerevisiae*. *Yeast*. 1987;3(3):175–85.
  27. Vanarsdell JN, Kwok S, Schweickart VL, Ladner MB, Gelfand DH, Innis MA. Cloning, characterization, and expression in *saccharomyces-cerevisiae* of endoglucanase-i from *trichoderma-reesei*. *Bio/Technology*. 1987;5(1):60–4.
  28. Werbeck ND, Rowling PJ, Chellamuthu VR, Itzhaki LS. Shifting transition states in the unfolding of a large ankyrin repeat protein. *Proc Natl Acad Sci U S A*. 2008;105(29):9982–7.
  29. Wright CF, Lindorff-Larsen K, Randles LG, Clarke J. Parallel protein-unfolding pathways revealed and mapped. *Nat Struct Biol*. 2003;10(8):658–62.
  30. Fischer WH, Spiess J. Identification of a mammalian glutaminyl cyclase converting glutaminyl into pyroglutamyl peptides. *Proc Natl Acad Sci U S A*. 1987;84(11):3628–32.
  31. Huang KF, Liu YL, Wang AHJ. Cloning, expression, characterization, and crystallization of a glutaminyl cyclase from human bone marrow: A single zinc metalloenzyme. *Protein Expr Purif*. 2005;43(1):65–72.
  32. Mozdzanowski J. Deblocking of proteins containing N-Terminal Pyroglutamic Acid. In: Smith B, editor. *Protein sequencing protocols*. vol. 211. Totowa, NJ: Humana Press; 2003. p. 365–9.
  33. Dana CM, Dotson-Fagerstrom A, Roche CM, Kal SM, Chokhawala HA, Blanch HW, et al. The importance of pyroglutamate in cellulase Cel7A. *Biotechnol Bioeng*. 2014;111(4):842–7.
  34. 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(19):6663–7.
  35. van den Burg B, Eijsink VG. Selection of mutations for increased protein stability. *Curr Opin Biotechnol*. 2002;13(4):333–7.
  36. Perl D, Schmid FX. Electrostatic stabilization of a thermophilic cold shock protein. *J Mol Biol*. 2001;313(2):343–57.
  37. Labbe S, Thiele DJ. Copper ion inducible and repressible promoter systems in yeast. In: *Methods in Enzymology; Expression of recombinant genes in eukaryotic systems*. vol. 306. San Diego, CA, USA: Academic Press Inc; London, UK: Academic Press Ltd; 1999. p. 145–53.
  38. Rakestraw JA, Sazinsky SL, Piatasi A, Antipov E, Wittrup KD. Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *saccharomyces cerevisiae*. *Biotechnol Bioeng*. 2009;103(6):1192–201.
  39. Robinson AS, Hines V, Wittrup KD. Protein disulfide-isomerase overexpression increases secretion of foreign proteins in *saccharomyces-cerevisiae*. *Bio/Technology*. 1994;12(4):381–4.
  40. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc*. 2007;2(1):31–4.
  41. Uzbas F, Sezerman U, Hartl L, Kubicek CP, Seiboth B. A homologous production system for *Trichoderma reesei* secreted proteins in a cellulase-free background. *Appl Microbiol Biotechnol*. 2012;93(4):1601–8.
  42. Schuster A, Bruno KS, Collett JR, Baker SE, Seiboth B, Kubicek CP, et al. A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. *Biotechnol Biofuels*. 2012;5:1.
  43. Nakarisetala T, Penttila M. Production of *trichoderma-reesei* cellulases on glucose-containing media. *Appl Environ Microbiol*. 1995;61(10):3650–5.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

