CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Caltech Authors - Main

Accepted Preprint

Biocatalysis, Protein Engineering, and Nanobiotechnology **Biotechnology and Bioengineering**
DOI 10.1002/bit.25308

Engineered thermostable fungal cellulases exhibit efficient synergistic cellulose hydrolysis at elevated temperatures†

Devin L. Trudeau*[,a], Toni M. Lee*[,b], Frances H. Arnold[a,b]

*These authors contributed equally to this work.

Author Affiliation:
[a]California Institute of Technology, Division of Chemistry and Chemical Engineering
1200 E. California Blvd.
Pasadena, CA 91125

[b]California Institute of Technology, Division of Biology and Bioengineering
1200 E. California Blvd.
Pasadena, CA 91125

Corresponding Author:
Frances H. Arnold
1200 E. California Blvd. MC 210-41
Pasadena, CA 91125

†This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/bit.25308]

Additional Supporting Information may be found in the online version of this article.

© 2014 Wiley Periodicals, Inc.
Received March 06, 2014; Revision Received May 28, 2014; Accepted June 02, 2014

Accepted Preprint

Abstract

A major obstacle to using widely available and low-cost lignocellulosic feedstocksto produce renewable fuels and chemicals is the high cost and low efficiency of the enzyme mixtures used to hydrolyze cellulose to fermentable sugars. One possible solution entails engineering current cellulases to function efficiently at elevated temperatures in order to boost reaction rates and exploit several other advantages of a higher temperature process. Here we describe the creation of the most stable reported fungal endoglucanase, a derivative of *Hypocrea jecorina* (anamorph *Trichoderma reesei*) Cel5A, by combining stabilizing mutations identified using consensus design, chimera studies, and structure-based computational methods. The engineered endoglucanase has an optimal temperature that is 17 °C higher than wild type *H. jecorina* Cel5A, and hydrolyzes 1.5 times as much cellulose over 60 h at its optimum temperature compared to the wild type enzyme at its optimal temperature.This enzyme complements previously-engineeredhighly-active, thermostable variants of the fungal cellobiohydrolases Cel6A and Cel7A in a thermostable cellulase mixture that hydrolyzes cellulose synergistically at an optimum temperature of 70 °C over 60 h.The thermostable mixture produces three times as much total sugar as the bestmixture ofthewild type enzymes operating at its optimum temperature of 60 °C, clearly demonstrating the advantage of higher-temperature cellulose hydrolysis.

Keywords: cellulase, endoglucanase, cellobiohydrolase, thermostability, synergy, biofuels

Accepted Preprint

Introduction

Cellulases engineered for increased thermostability can reduce lignocellulosic biomass degradation times and costs, facilitating the use of these feedstocks for production of biofuels and chemicals(Turner et al. 2007). Thermostable cellulases can display increased cellulolytic activity at high temperatures and retain activity for longer periods of time than their less stable counterparts(Mingardon et al. 2011; Viikari et al. 2007). Moreover, biomass degradation at elevated temperatures reduces cooling costs following pre-treatment and reduces the risk of microbial contamination(Turner et al. 2007).

Cellulose degradation in nature is a complex process that can involve many different enzymes operating either independently or in large protein complexes known as cellulosomes (Brunecky, 2013). In aerobic fungi, there are three major classes of noncomplexed cellulolytic enzymes: cellobiohydrolases I and II processively hydrolyze from opposite ends (reducing and non-reducing, respectively) of the cellulose chain,while endoglucanases cleave intrachain bonds (Martinez et al. 2008; Rosgaard et al. 2007).This laboratory has engineered thermostable class I fungal cellobiohydrolases (Cel7A)(Heinzelman et al. 2010; Komor et al. 2012) and class II cellobiohydrolases (Cel6A)(Heinzelman et al. 2009a; Wu and Arnold 2013) using a combination of SCHEMA recombination, rational design, and directed evolution. Mixtures of thermostabilized Cel6A and Cel7A increase released cellobiose by up to 1.8 fold relative to a wild type Cel6A and Cel7A mixture when each mixture operates at its optimum temperature (70 °C for engineered and 60 °C for wild type)over a 60 h incubation(Wu and Arnold 2013).Engineering a thermostable fungal endoglucanase that retains high catalytic activity would constitute an important step in creating an effective thermostable fungal cellulolytic enzyme mixture.

Accepted Preprint

In the industrial fungal strain *Hypocrea jecorina* (anamorph *Trichoderma reesei*), the class II endoglucanaseCel5A (*Hj*Cel5A) accounts for up to 12% of the total secreted cellulase and 55% of the endoglucanase activity(Rosgaard et al. 2007; Suominen et al. 1993; Zhang and Lynd 2004). *Hj*Cel5A exhibits significantly decreased activity at 70 °C when expressed from either *Hypocrea jecorina* or *Saccharomyces cerevisiae* (Qin et al. 2008).Wild type *Hj*Cel5Aexpressed in *S. cerevisiae* was found to have optimum activity at 64 °C in a 2-hr cellulase hydrolysis assay (Lee 2014), making it incompatible with currently available thermostable fungal cellulases. We thereforestabilized this enzyme by protein engineering, which allowed us to assess the synergy of engineered thermostable cellobiohydrolases and endoglucanases and evaluate how the activity of thermostabilized cellulase mixtures increases at elevated temperatures.

Materials and methods

Plasmids and Strains

Genes encoding Cel6A, and Cel7A were cloned into the yeast secretion vector YEp352/PGK91-1-αss as described previously(Komor et al. 2012; Wu and Arnold 2013) ENREF 7. The gene encoding wild type Cel5A (including its cellulose binding module) was synthesized with *S. cerevisiae* codon optimization (DNA 2.0). Sequences of all genes are in Supplementary Information. Plasmids were transformed into the BY4742 Δkre2 strain of yeast

(BY4742; Mat a; his3D1; leu2D0; lys2D0; ura3D0; YDR483w::kanMX4) obtained from EUROSCARF.

Enzyme Purification

Yeast colonies expressing Cel5A and Cel6A with C-terminal His_6 tags and Cel7A with an N-terminal His_8 tag were grown at 30 °C: first overnight in 5 mL SD-Ura medium, expanded into 50 mL SD-Ura (+50 µg/mL kanamycin) medium for 24 h, and then expanded into 1 L YPD (+50 µg/mL kanamycin) medium for an additional 48 h. Cultures were centrifuged at 4500 *g* for 20 min, and the supernatant was filtered with 0.2 mm PES filter unit from Nalgene. Protein was purified using 5mL HisTrap columns (GE Healthcare). Purified cellulases were buffered-exchanged to 50mM sodium acetate buffer pH 5.0 using Vivaspin 20 ultrafiltration spin tubes (GE Healthcare). Protein concentrations were determined using A_{280}, with extinction coefficients calculated using ProtParam on the ExPASy server(Gasteiger et al. 2005).

Thermostability Measurements

100 µL samples in 50 mM sodium acetate buffer, pH 5.0 containing 0.2 µM Cel5A and 1% (w/v) Avicel were incubated at a range of temperatures for 2 h in an Eppendorf Mastercycler. A modified Park-Johnson reducing sugar assay was used to measure activity(Park and Johnson 1949). Briefly, reaction mixtures were spun at 1000 g for 5 min to remove Avicel. 50 µL of supernatant was removed and transferred to a mixture of 100 µL ferricyanide reagent (0.5 g/L $K_3Fe(CN)_6$, 34.84 g/L K_2HPO_4, pH 10.6) and 50 µL carbonate-cyanide reagent (5.3 g/L Na_2CO_3, 0.65 g/L KCN). The reaction was heated at 95 ºC for 15 min in an Eppendorf Mastercycler, and then cooled on ice for 5 min. 180 µL of the reaction was removed and mixed with 90 µL ferric iron

Accepted Preprint

solution (2.5 g/L $FeCl_3$, 10 g/L polyvinyl pyrrolidone, 2 N H_2SO_4). After 2 min, absorbance at 595 nm was taken, using solutions of 0 µM to 300 µM cellobiose as standards.

Cellulase Activity Measurements

All cellulase activity measurements were conducted in 50 mM sodium acetate buffer, pH 5.0. Constant temperature was maintained using an Eppendorf Mastercycler. To determine activity-temperature profiles of Cel5A, samples containing 0.2 µM of purified Cel5Aand 1% (w/v) Avicel were incubated at 60 and 70 °C for 60 h. To determine the activity of the Cel5A, Cel6A, and Cel7A mixtures, purified Cel5A, Cel6A, and Cel7A were combined at different ratios to a final concentration of 0.5 µMalong with 1% Avicel in 100 µL and incubated 60 °C and 70 °C for 60 h. After hydrolysis, reaction supernatants were sampled for reducing sugar concentrations via a modified Nelson–Somogyi assay(Green et al. 1989). Briefly, 50 µL of reaction solution was added to 40 µL carbonate-tartrate solution (144 g/L Na_2SO_4, 12 g/L potassium tartrate tetrahydrate, 24 g/L Na_2CO_3, 16 g/L $NaHCO_3$) and 10 µL copper solution (180 g/L Na_2SO_4, 20 g/L $CuSO_4.5H_2O$) and heated to 95 ºC for 15 min in an Eppendorf Mastercycler. The reaction was placed on ice for 5 min and then mixed with 50 µL arsenomolybdate solution (50 g/L $(NH_4)_2MoO_4$, 1.5 N H_2SO_4, 6 g/L NaH_2AsO_4). After mixing, absorbance at 520 nm was read, using 0 to 2mM cellobiose solutions as standards.

Cellulose hydrolysis to determine the activity of optimized engineered and wild type cellulase mixtures was carried out on 1%, 3%, and 5% Avicel at 60 °C and 70 °C. Samples were taken at 0 h, 4 h, 8 h, 15 h, 24 h, 36 h, 48 h, and 60 h, and reducing sugar concentration was quantified as above.

Accepted Preprint

Accepted Preprint

Ammonia fiber expansion (AFEX) pre-treated corn stover was obtained from the Great Lakes Bioenergy Research Center and milled to 250µm size as described by Banerjee and coworkers (Banerjee et al. 2010). Pre-treated lignocellulose from rice straw was obtained from Frank C. J. Chang and prepared as described by Hsu and coworkers (Hsu et al. 2010). Activity assays were carried out for optimized engineered and wild type cellulase mixtures on 3% substrate at 60 °C and 70 °C for 60 h. Reducing equivalent at the end of 60 h was quantified as described above.

Data analysis

Cellulase activity and thermostability data were plotted using Microsoft Excel. Synergy plots were made in Matlab (The Mathworks, Inc.), using the Ternplot package developed by Carl Sandrock(http://www.mathworks.com/matlabcentral/fileexchange/2299-ternplot).

Results

Engineering a thermostable fungal Cel5 endoglucanase

To create a thermostable *Hj*Cel5A, we combined stabilizing mutations identified from homologous recombination, consensus design, and various computational approaches (Lee 2014,Trudeau 2014).**Table 1** lists the mutations and the methods by which they were discovered.

Sixteen thermostabilizing mutations that did not compromise activity measured at 60 ºC were combined into a single variant. If two suitable mutations occupied the same site, the more thermostabilizing of the two was selected. The mutations in this combination mutant are T57N, E53D, S79P, T80E, V101I, S133R, N155E, G189S, F191V, T233V, G239E, V265T, D271Y, G293A, S309W, and S318P.

We transformed the resulting *Hj*Cel5A variant (OptCel5A) into the glycosylation-deficientBY4742 Δkre2 strain of *Saccharomyces cerevisiae*(Heinzelman et al. 2009b). When expressed and purified, OptCel5A has an optimal temperature of 81 °C when used to hydrolyze crystalline cellulose (Avicel) for 2 h (**Figure 1A**). OptCel5A therefore has an optimal temperature 17 °C higher than wild type *Hj*Cel5A. Based on optimal temperature on crystalline cellulose over 2h, OptCel5A is the most stable fungal endoglucanase reported.

We investigated the activity of OptCel5A and wild type*Hj*Cel5A over a 60 h hydrolysis at both 60 °C and 70 °C. OptCel5A had highest activity at 70 °C, hydrolyzing more than 1.5 timesas much cellulose as*Hj*Cel5A at its optimal temperature of 60 °C (**Figure 1B**). OptCel5A is compatible with the previously engineered thermostable Cel6A and Cel7A, which both operate optimally at 70 °C in 60 h hydrolysis experiments(Wu and Arnold 2013).

Synergy in cellulose hydrolysis

It has been known for 40 years that endoglucanases and cellobiohydrolases act synergistically to degrade cellulose(Baker et al. 1998; Tomme et al. 1995; Wood and McCrae 1978). We explored the synergy among cellobiohydrolases Cel6A and Cel7A and endoglucanase

Accepted Preprint

Cel5A by comparing mixtures of the wild type enzymes with engineered-thermostable cellulase mixtures. The engineered-thermostable mixture consists of OptCel5A as the Cel5A variant, 3C6P as the Cel6A(Wu and Arnold 2013), and TS8 as the Cel7A(Komor et al. 2012). Each of these enzymes has an optimal activity at or greater than 70 °C when measured over 60 h incubations on Avicel. In the wild type mixture (heterologously expressed in *S. cerevisiae*), we used Cel5A from *H. jecorina*, Cel6A from *Humicola insolens*, and Cel7A from *Talaromyces emersonii*; these are the most thermostable reported homologues of each enzyme. This wild type cellulase mixture exhibits optimal activity at 60 °C over 60 h(Wu and Arnold 2013).

In these experiments the total cellulase concentration remained fixed at 0.5 µM while the relative concentrations of each cellulase varied in steps of 0.1 µM. The resulting data wereused to produce a ternary synergy diagram(Baker et al. 1998). Reactions on Avicel were conducted over 60 h at 60 °C for wild type and 70 °C for the engineered enzymes, conditions consistent with previous synergy studies(Baker et al. 1998; Meyer et al. 2009; Wu and Arnold 2013) and realistic industrial conditions. As shown in **Figures 2A** and **B**, both enzyme mixtures exhibited substantial synergy, with the mixtures more active than any of the enzymes alone. The degree of synergy, obtained by dividing the activity of the mixture by the sum of the activities of the individual cellulases(Zhang and Lynd 2004), ranged from 1.0 to 1.6 for the wild type enzymes, and from 1.0 to 2.1 for the engineered enzymes.

In both wild type and engineered mixtures the highest cellulose hydrolysis activity occurred with relatively small amounts of endoglucanase (10-20% of total mixture), a finding observed in other synergy studies(Van Dyk and Pletschke 2012). Cellobiohydrolases processively hydrolyze along free cellulose ends, and are responsible for the bulk of hydrolysis. Endoglucanases non-processively produce free cellulose ends, increasing the effective concentration of

Accepted Preprint

cellobiohydrolase substrate. Less endoglucanase is required in an optimal mixture because the enzyme assists cellobiohydrolase function.

Wild type mixtures demonstrating the highest level of hydrolysis contained mostly Cel7A, while Cel6A was the primary component of the most effective engineered mixtures. As shown in **Figure 2C** and **D**, this change in optimal relative enzyme loadings reflects the relative activities of Cel6A and Cel7A in the wild type and engineered cases(Wu and Arnold 2013). **Figures 2C** and **D** also show the activities of the optimal cellulase mixtures for two and three enzymes. The best mixture of wild type enzymes in this experiment exhibitedmore than 1.5 foldthe activity of any of its constituent enzymes, while the optimal mixture of engineered thermostable enzymes was over 2.5 foldas active. The optimal engineered thermostable mixture also exhibited 1.2 times the activity of a mixture containing only engineered Cel6A and Cel7A with an equal total enzyme concentration.

An optimized mixture of engineered cellulases accelerates cellulose hydrolysis

We searched the region of maximum activity more closely in steps of 0.04 μM and found the optimal mixture for wild type to be 0.16: 0.28: 0.56 (Cel5A:Cel6A:Cel7A). The optimal engineered thermostable mixture is 0.08:0.56:0.36 (Cel5A:Cel6A:Cel7A). We call the optimized engineered thermostable mixture T-PRIMED. We evaluated the activity of T-PRIMED over 60 h at both 60 °C and 70 °C and compared it to the activity of the best wild type mixture. We ran this assay on 1%, 3%, and 5% Avicel to observe the effects of varying cellulose concentrations (**Figure 3A,B, C**). T-PRIMED exhibits the highest activity at 70 °C, where it is approximately three times as active as the best mixture of wild type enzymes at 60 °C. The activities of all cellulase mixtures

Accepted Preprint

increased at higher cellulose concentrations, with the activity ratios remaining approximately the same.

We also tested the activity of the mixtures on two industrially relevant lignocellulose substrates, milled corn stover and dilute acid-treated rice straw (**Figure 3D**). T-PRIMED was more active than the wild type enzymes on both substrates, with 1.8 foldthe activity of the wild type mixture on milled corn stoverand 2.5 foldhigher activity on treated rice straw.

Discussion

We report herethe most stable reported variant of*Hj*Cel5A, which we call OptCel5A, and the characterization of its synergy with other engineered thermostable cellulases in cellulose hydrolysis. This enzyme has an optimal temperature (measured in 2 h activity assays) of 81 ºC and releases more than 1.5 times as much soluble sugar over 60 h compared to wild type*Hj*Cel5A.

All cellulases investigated in this study (*Hj*Cel5A, OptCel5A, *Hi*Cel6A, 3C6P, *Te*Cel7A, and TS8) were expressed and purified from *S. cerevisiae*. Proteins expressed in *S. cerevisiae* may have different patterns of glycosylation and post-translational modifications compared to *H. jecorina*(Dana et al. 2014; Qin et al. 2008).*Hj*Cel5A expressed from *S. cerevisiae* strain H158 is reported to have two isoforms that differ in their glycosylation and have slight differences in thermostability (Qin et al. 2008). We used *S. cerevisiae* strain BY4742 Δkre2, a glycosylation-deficient strain, in order to minimize potential effects of glycosylation (Heinzelman et al. 2009b). It has also recently been reported that an N-terminal pyroglutamate modification in *Te*Cel7A can improve thermostability in the native host (Dana et al. 2014). Such effects were not investigated for the enzymes used here, and activity results for the engineered cellulases may change when they are expressed in other hosts.

Accepted Preprint

Accepted Preprint

OptCel5A is 17 ºC more stable than wild type *Hj*Cel5A and retains the same activity at 60ºC (Figure 1). Its higher stability enables the enzyme to remain active at elevated temperatures, where its specific activity increases. Leefound a large number of thermostabilizing mutations, many of which can have negative effects on endoglucanase activity (Lee 2014). Stabilizing mutations must therefore be chosen carefully to generate thermostable variants that are improved over the native enzyme and enable efficient operation at elevated temperature. In the OptCel5A combination mutant described here, we chose single mutations that were both thermostabilizing and did not decrease activity measured at 60 ºC in order to achieve maximal activity at high temperatures. These mutations were distributed across the protein, both in the solvent-inaccessible core (F191V, G293A, V101I), and on the surface (V265T, S318P, S79P, T57N, N155E, T80E, S133R) (Figure S1). These mutations were predicted to thermostabilize *Hj*Cel5A by a variety of mechanisms, including stabilizing the alpha helical dipole, stabilizing the protein backbone, and improving packing in the hydrophobic core (Table 1).

The higher stability of OptCel5A allows it to function at temperatures above 80 ºC, where activity increases approximately 1.5 fold in 2 h incubations on Avicel. This activity increase is comparable to those reported for the thermostabilized cellobiohydrolases Cel6A and Cel7A used in this study. The engineered thermostable Cel6A, 3C6P, has an optimum temperature 10 ºC higher than the most thermostable wild type, *Hi*Cel6A (75 ºC*vs*. 65 ºC), and an optimal activity 1.6 fold higher (Wu and Arnold 2013). The engineered thermostable Cel7A, TS8, has an optimum temperature 10 ºC higher than the most thermostable wild type, *Te*Cel7A (65 ºC vs. 55 ºC) and an optimal activity that is increased 1.5 fold(Komor et al. 2012).

Mingardon and coworkers explored the activity-temperature relationships among wild type mesophilic and thermophilic bacterial endoglucanases (Mingardon et al. 2011). They also found

that optimum activity increases when the bacterial endoglucanases are more thermostable and can be used at higher temperatures. *Thermobifida fusca* Cel9A (optimum temperature = 70 ºC) had approximately twice the activity of *Clostridium cellulolyticum* Cel9G (optimum temperature = 60 ºC), and *Clostridium thermocellum* Cel9I (optimum temperature = 80 ºC) had 1.3 fold higher activity than *Thermobifida fusca* Cel9A (optimum temperature = 70 ºC).

OptCel5A works synergistically with previously reported engineered thermostable cellobiohydrolases I and II(Wu and Arnold 2013), resulting in a mixture with more than twice the cellulase activity expected from the sum of the activities of individual enzymes. OptCel5A increases the activity of a previously reported mixture of engineered cellobiohydrolases(Wu and Arnold 2013) 1.2 fold. T-PRIMED, an optimized mixture of the three enzymes, releases over 3 fold more soluble sugar over 60 h from Avicel compared to a similarly optimized wild type mixture. T-PRIMED is also more active on cellulose substrates derived from corn stover and rice straw.

Notably, the relative amounts of constituent enzymes required for optimal cellulase activity switched from being dominated by Cel7A in the wild type mixture to having more Cel6A in T-PRIMED. This change reflects the change in relative activities of the wild type and engineered enzymes: wild type Cel7A is more active than wild type Cel6A, while engineered Cel6A (3C6P) is more active than the engineered Cel7A (TS8) (Figure 2C and D). The synergy between Cel6A and Cel7A is still poorly understood, but may arise from effects on enzyme mobility as the different enzymes interact on the cellulose surface (Igarashi et al. 2011) or low levels of endoglucanase activity in Cel6A (Boisset et al. 2001). These synergistic effects remain constant across a wide range of concentrations of each cellobiohydrolase (Igarashi et al. 2011).

Accepted Preprint

Accepted Preprint

An effectivemixture of wild type fungal cellulases is known to require at least three different activities: endoglucanase, cellobiohydrolase I, and cellobiohydrolase II; this holds true for engineered enzymes as well. The synergy values we observed are typical for reaction of fungal cellulase mixtures on Avicel, which range from 1.3 to 2.2 for *H. jecorina* cellulases(Zhang and Lynd 2004).The degree of synergism increased from a maximum of 1.6 for wild type cellulases to 2.1 for engineered thermostable cellulases. Although these datasuggest that synergy may be temperature dependent, wild type mixtures assessed for cellulase activity at 50 ºC, 60 ºC, and 70 ºC have similar synergy values (**Supplementary Figure 2**).

T-PRIMED displays an optimum temperature of 70 °C when digesting Avicel and corn stover, but demonstrates an optimal activity of 60 °C on treated rice straw. This change in optimum temperature may reflect subtle variations in cellulase thermostability related to the protective effects achieved when enzymes bind to cellulose(Mingardon et al. 2011). Substrate composition may affect the fraction of enzyme bound, the thermostability of the enzymes, and, consequently, the optimal temperature of the reaction(Zhang and Lynd 2004).

Synergy is also expected to decrease with hydrolysis time(Medve et al. 1994) and enzyme loading(Woodward et al. 1988), two properties that were not investigated in this study. Engineered cellulase mixtures will therefore require further optimization for particular applications. High-throughput approaches for optimizing cellulase mixtures, such as robotic platforms(Banerjee et al. 2010) and computationally guided approaches(Rivera et al. 2010) may assist in these efforts.

In summary, we have combined the results of multiple protein engineering efforts to: 1) create the most thermostable fungal endoglucanase reported, 2) create the most thermostable set of synergistically-acting cellulases reported to date, and 3) demonstrate an approximately three-fold enhancement in hydrolysis activity on crystalline cellulose for this set compared to a set of wild

type fungal enzymes. Our study demonstrates three important considerations for engineering systems of cellulolytic enzymes: 1) thermostabilization can enhance activity significantly by enabling hydrolysis at elevated temperatures, 2) when enzymes work cooperatively, it is necessary to engineer all key components of the system to attain the highest possible improvement, and 3)because the relative importance of enzymes in these systems can change, synergy experiments such as those carried out in this study should be used to find the optimum enzyme mixture.

Acknowledgments

The authors thank Indira Wu, Mathew Smith, Alex Nisthal, and two anonymous reviewers for helpful discussions and suggestions.This work was supported by the Institute for Collaborative Biotechnologies through grant [W911NF-09-D-0001] from the U.S. Army Research and The National Central University, Taiwan, though a Cooperative Agreement for Energy Research Collaboration. DLT was supported by a Canadian National Science and Engineering Research Council post-graduate fellowship [PGSD3-404332-2011]. TML was supported by a National Science Foundation graduate fellowship.A patent application has been submitted for the enzyme variants and mixtures described in this manuscript.

Accepted Preprint

Accepted Preprint

References

Baker JO, Ehrman CI, Adney WS, Thomas SR, Himmel ME. 1998. Hydrolysis of cellulose using ternary mixtures of purified celluloses. Biotechnology for Fuels and Chemicals: Springer. p 395-403.

Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Aslam N, Walton JD. 2010. Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. Biotechnology and bioengineering 106(5):707-720.

Billard H, Faraj A, Lopes Ferreira N, Menir S, Heiss-Blanquet S. 2012. Optimization of a synthetic mixture composed of major *Trichoderma reesei* enzymes for the hydrolysis of steam-exploded wheat straw. Biotechnol Biofuels 5(1):9.

Boisset C, Pétrequin C, Chanzy H, Henrissat B, Schülein M. 2001. Optimized mixtures of recombinant *Humicola insolens* cellulases for the biodegradation of crystalline cellulose. Biotechnology and bioengineering 72(3):339-345.

Brunecky R, Alahuhta M, Xu Q, Donohoe BS, Crowley MF, Kataeva IA, Yang S-J, Resch MG, Adams MWW, Lunin VV, Himmel ME, Bomble YJ. 2013. Revealing nature's cellulase diversity: the digestion mechanism of *Caldicellulosiruptor bescii* CelA. Science 342(6165):1513–1516.

Dana CM, Dotson-Fagerstrom A, Roche CM, Kal SM, Chokhawala HA, Blanch HW, Clark DS. 2014. The importance of pyroglutamate in cellulase Cel7A. Biotechnol Bioeng. 111:842-847.

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook: Springer. p 571-607.

Green F, 3rd, Clausen CA, Highley TL. 1989. Adaptation of the Nelson-Somogyi reducing-sugar assay to a microassay using microtiter plates. Anal Biochem 182(2):197-9.

Heinzelman P, Komor R, Kanaan A, Romero P, Yu X, Mohler S, Snow C, Arnold F. 2010. Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence blocks by SCHEMA structure-guided recombination. Protein Engineering Design and Selection 23(11):871-880.

Heinzelman P, Snow CD, Smith MA, Yu X, Kannan A, Boulware K, Villalobos A, Govindarajan S, Minshull J, Arnold FH. 2009a. SCHEMA recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability. Journal of Biological Chemistry 284(39):26229-26233.

Heinzelman P, Snow CD, Wu I, Nguyen C, Villalobos A, Govindarajan S, Minshull J, Arnold FH. 2009b. A family of thermostable fungal cellulases created by structure-guided recombination. Proceedings of the National Academy of Sciences 106(14):5610-5615.

Hsu T-C, Guo G-L, Chen W-H, Hwang W-S. 2010. Effect of dilute acid pretreatment of rice straw on structural properties and enzymatic hydrolysis. Bioresource technology 101(13):4907-4913.

Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttilä M, Ando T, Samejima M. 2011. Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. Science 333(6047):1279-1282.

Accepted Preprint

Komor RS, Romero PA, Xie CB, Arnold FH. 2012. Highly thermostable fungal cellobiohydrolase I (Cel7A) engineered using predictive methods. Protein Eng Des Sel 25(12):827-33.

Lee TM, Farrow MF, Arnold FH, Mayo SL. 2011. A structural study of Hypocrea jecorina Cel5A. Protein Science 20(11):1935-1940.

Lee TM. 2014. Computationally-guided thermostabilization of the primary endoglucanase from *Hypocrea jecorina* for cellulosic biofuel production.Thesis (Ph.D.), California Institute of Technology.

Martinez D, et al. 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nature Biotechnology. 26(5):553-60.

Medve J, Ståhlberg J, Tjerneld F. 1994. Adsorption and synergism of cellobiohydrolase I and II of *Trichoderma reesei* during hydrolysis of microcrystalline cellulose. Biotechnology and bioengineering 44(9):1064-1073.

Meyer AS, Rosgaard L, Sørensen HR. 2009. The minimal enzyme cocktail concept for biomass processing. Journal of Cereal Science 50(3):337-344.

Mingardon F, Bagert JD, Maisonnier C, Trudeau DL, Arnold FH. 2011. Comparison of family 9 cellulases from mesophilic and thermophilic bacteria. Applied and environmental microbiology 77(4):1436-1442.

Park JT, Johnson MJ. 1949. A submicrodetermination of glucose. J Biol Chem 181(1):149-51.

Qin Y, Wei X, Liu X, Wang T, Qu Y. 2008. Purification and characterization of recombinant endoglucanase of *Trichoderma reesei* expressed in Saccharomyces cerevisiae with higher glycosylation and stability. Protein Expr Purif 58(1):162-7.

Rivera EC, Rabelo SC, dos Reis Garcia D, da Costa AC. 2010. Enzymatic hydrolysis of sugarcane bagasse for bioethanol production: determining optimal enzyme loading using neural networks. Journal of Chemical Technology and Biotechnology 85(7):983-992.

Rosgaard L, Pedersen S, Langston J, Akerhielm D, Cherry JR, Meyer AS. 2007. Evaluation of minimal *Trichoderma reesei* cellulase mixtures on differently pretreated barley straw substrates. Biotechnology progress 23(6):1270-1276.

Suominen PL, Mantyla AL, Karhunen T, Hakola S, Nevalainen H. 1993. High frequency one-step gene replacement in *Trichoderma reesei*. II. Effects of deletions of individual cellulase genes. Molecular Genetics and Genomics 241(5-6):523-530.

Tomme P, Warren R, Gilkes N. 1995. Cellulose hydrolysis by bacteria and fungi. Advances in microbial physiology 37:1-81.

Trudeau DL.2014. Engineering enzyme systems by recombination. Thesis (Ph.D.), California Institute of Technology.

Turner P, Mamo G, Karlsson EN. 2007. Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb Cell Fact 6(9):1-23.

Van Dyk J, Pletschke B. 2012. A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes—Factors affecting enzymes, conversion and synergy. Biotechnology advances 30(6):1458-1480.

Viikari L, Alapuranen M, Puranen T, Vehmaanperä J, Siika-Aho M. 2007. Thermostable enzymes in lignocellulose hydrolysis. Biofuels: Springer. p 121-145.

Wood TM, McCrae SI. 1978. The cellulase of Trichoderma koningii. Purification and properties of some endoglucanase components with special reference to their action on cellulose when acting alone and in synergism with the cellobiohydrolase. Biochem J 171(1):61-72.

Accepted Preprint

Woodward J, Hayes MK, Lee NE. 1988. Hydrolysis of Cellulose by Saturating and Non–Saturating Concentrations of Cellulase: Implications for Synergism. Nature Biotechnology 6(3):301-304.

Wu I, Arnold FH. 2013. Engineered thermostable fungal Cel6A and Cel7A cellobiohydrolases hydrolyze cellulose efficiently at elevated temperatures. Biotechnology and Bioengineering 110(7):1874-1883.

Zhang YHP, Lynd LR. 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. Biotechnology and bioengineering 88(7):797-824.

Accepted Preprint

Tables

Table 1. Stabilizing mutations combined to create OptCel5A. Two related metrics were used to quantify thermostability change. Lee (2014) report change in T_{50}, the temperature at which an enzyme loses half of its activity after a 10 min incubation, as measured in a 1 h reaction at 60 ºC. Due to existence of refolding variants in their screen, Trudeau (2014)instead report change in T_{A50}, the temperature at which an enzyme loses half of its activity relative to activity at its optimum temperature, measured in a 2 h reaction. These metrics correlate closely when measured for the same variant.

Mutation	Thermostability increase (ºC)	Stabilization method	Source
F191V	0.9	Chimeragenesis	Trudeau 2014
T233V	0.9	Chimeragenesis	
V265T	2.0	Chimeragenesis	
S318P	3.4	FoldX/Chimeragenesis	Lee 2014
D271Y	2.7	FoldX/Chimeragenesis	Trudeau 2014
S79P	0.3	FoldX	Lee 2014
E53D	2.7	Consensus	Lee 2014
T57N	1.1	Consensus	
G293A	3.6	Consensus	
V101I	0.1	Core repacking	Lee 2014
N155E	0.5	Helix dipole stabilization	Lee 2014
T80E	0.5	Helix dipole stabilization	
S133R	0.4	Helix dipole stabilization	
G239E	0.2	Helix dipole stabilization	
S309W	0.4	Triad ΔΔG	Lee 2014
G189S	0.9	Backbone stabilization	Lee 2014

Accepted Preprint

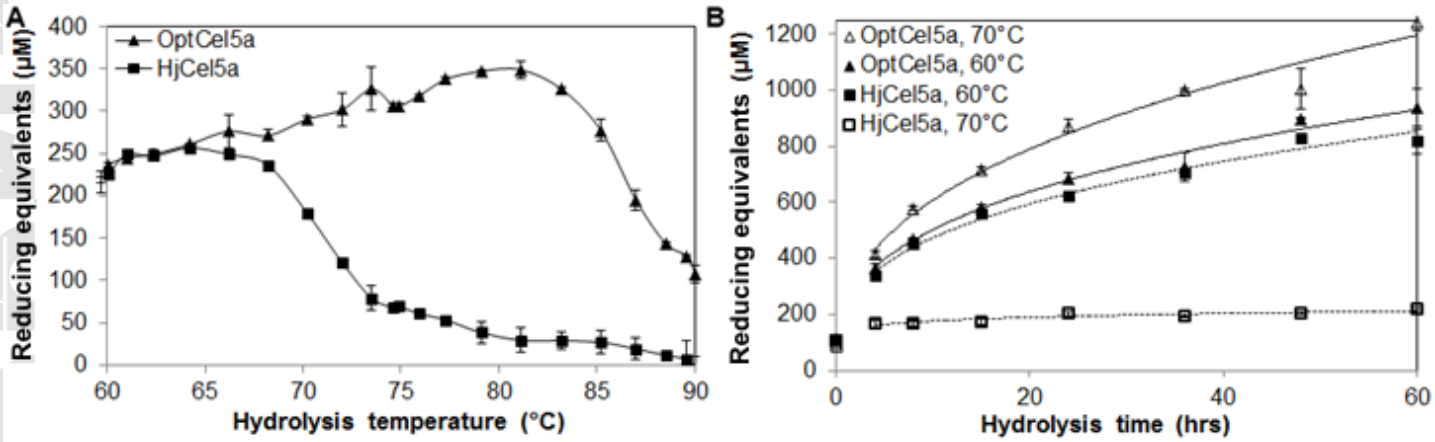
List of Figures

Figure 1. A highly stable engineered Cel5A endoglucanase. A) Total reducingequivalents released after 2 h Avicel hydrolysis with *Hj*Cel5A and OptCel5A. B) Total reducingequivalents released over 60 h Avicel hydrolysis at 60 °C and 70 °C with *Hj*Cel5A and OptCel5A. Reactions contained 0.2 μM enzyme and 1% Avicel.

Figure 2. Synergistic cellulose hydrolysis by wildtype (A,C) and engineered-thermostable (B,D) Cel5A, Cel6A, and Cel7A. Reactions contained a total concentration of 0.5 μM of cellulase and 1% w/v Avicel. Edges are labeled with the concentration of the noted cellulase, which ranges from 0% to 100% of the total. Each vertex represents 100% of an individual cellulase, each edge represents a mixture of two cellulases, and the interior of the triangle represents a mixture of all three cellulases. Black dots are individual measurements (in duplicate), and colors are arithmetic averages between each point, with red representing maximum activity and blue representing minimum activity. Colors are normalized for each synergism test. The absolute activities of the individual enzymes as well as the best mixtures for double and triple enzyme combinations are shown for wild type (**C**) and engineered thermostable enzymes (**D**).

Figure 3. Total reducingequivalents released during 60 h hydrolysis with wild type and engineered-thermostable cellulase mixtures at 60 °C and 70 °C, on both 1% (**A**), 3% (**B**), and 5% (**C**) w/v Avicel; and after 60 h hydrolysis on milled corn stover and dilute-acid treated rice straw (**D**). The wild type mixture is 0.16:0.28:0.56 (Cel5A:Cel6A:Cel7A) and the engineered thermostable mixture is 0.08:0.56:0.36 (Cel5A:Cel6A:Cel7A), with a total concentration of 0.5 μM, as described in the text.

Figure 1



Accepted Preprint

Figure 2

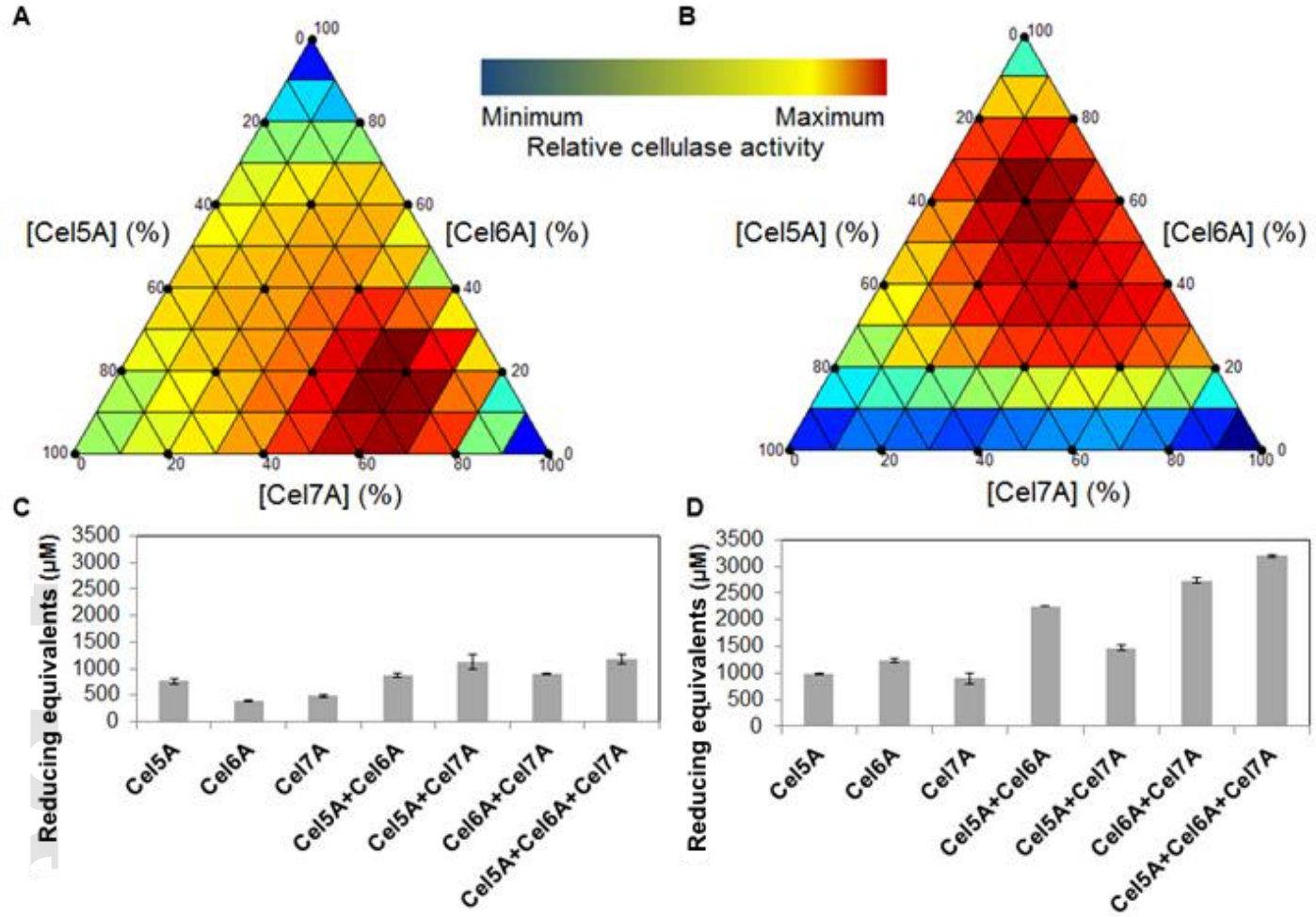
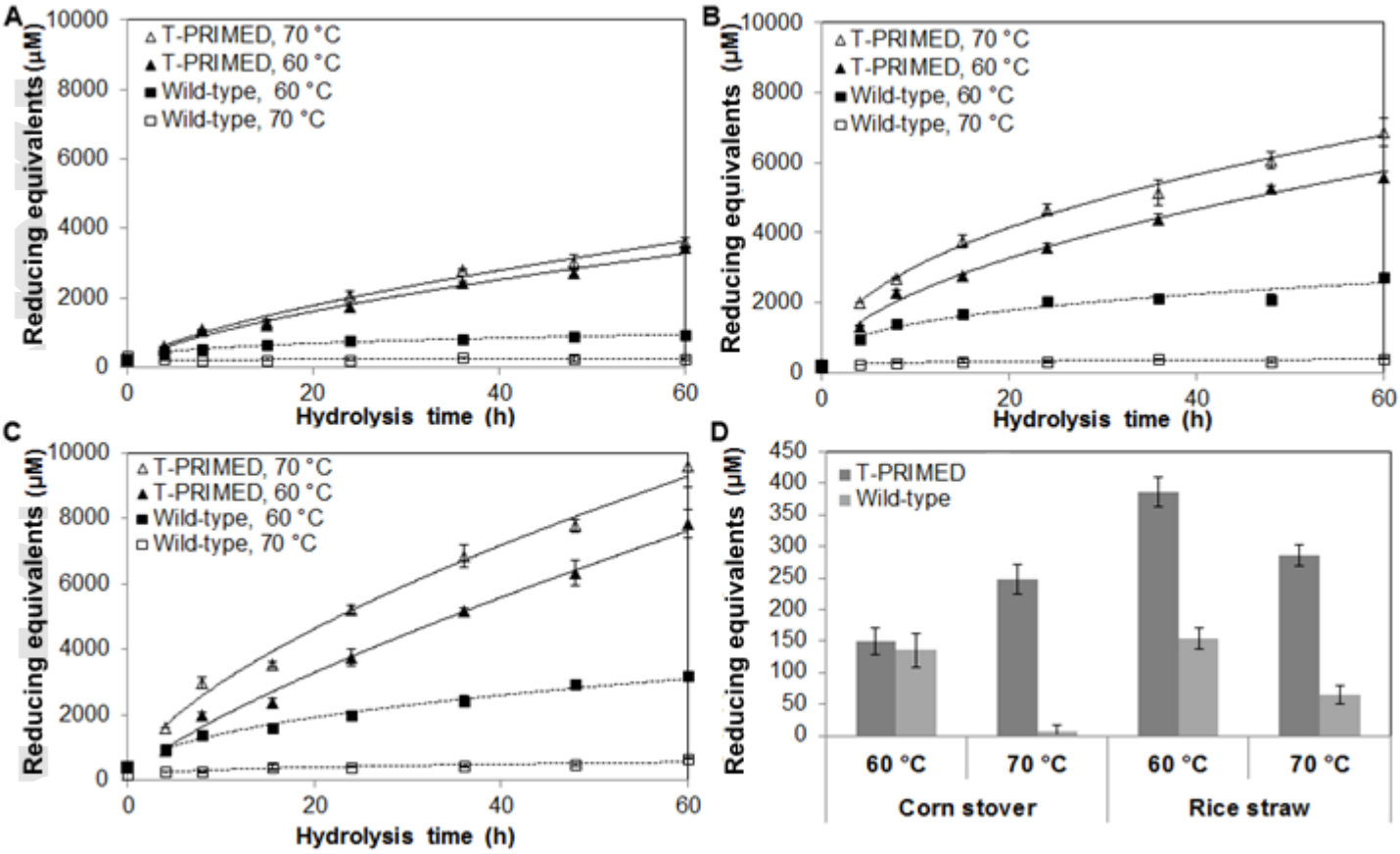

A
B
Minimum
Maximum
Relative cellulase activity
[Cel5A] (%)
[Cel6A] (%)
[Cel7A] (%)
C
D
Reducing equivalents (µM)
Cel5A
Cel6A
Cel7A
Cel5A+Cel6A
Cel5A+Cel7A
Cel6A+Cel7A
Cel5A+Cel6A+Cel7A

Figure 3



Accepted