Organic Process Research & Development **2011,** *15,* 275–278

Improved Saccharification of Wheat Straw for Biofuel Production Using an Engineered Secretome of *Trichoderma reesei*

Céline Ayrinhac,[†] Antoine Margeot,[‡] Nicolas Lopes Ferreira,[‡] Fadhel Ben Chaabane,[‡] Frédéric Monot,[‡] Gilles Ravot,[†] Jean-Marie Sonet,[†] and Laurent Fourage*[,†]

Protéus SA, 70, allée Graham Bell, Parc scientifique Georges Besse, 30035 Nimes Cedex 1, France, and IFP Energies Nouvelles, 1 et 4, avenue de Bois-Préau, 92852 Rueil Malmaison Cedex, France

Abstract:

For the purpose of a industrial process of cellulosic ethanol production, an efficient *β*-glucosidase was evolved by L-Shuffling starting from three parental genes (*i.e.*, *Chaetomium globosum* glucosidase putative gene, *Trichoderma reesei bgl1* gene, and *Neurospora crassa* glucosidase putative gene, named genes A, B, and C, respectively) originating from microbial biodiversity and showing 70% of identity at the amino acid level. Enzyme B (encoded by *bgl1* gene) was chosen as a reference so that the backbone of the evolved enzymes would be based on this enzyme. Two rounds of L-Shuffling and colonies screening (20,000 colonies per round) on chromogenic glucose substrate were performed. Compared with native *β*-glucosidase, the most evolved enzyme has a 242-fold increased k_{cat} for the *p*NPGlc substrate. After expression of this improved *β*-glucosidase in *T. reesei*, a new efficient enzymatic cocktail was secreted by the strain allowing for a 4-fold decrease in cellulase loading without any loss in hydrolysis performance of degradation of a steam-exploded wheat straw compared to the untransformed parental strain.

Introduction

The current worldwide *bioethanol* production reached ∼74 billions liters in 2009[1] in more than 40 countries. This represents nearly 400% growth since 2000. The global *bioethanol* production will reach 85.9 billion liters in 2010, and it is expected to keep on growing in the coming years.[2] First-generation *bioethanol* (1G *bioethanol*) production uses easily processed sugar cane, sugar beets, and starch from grain-producing crops as feedstock, which raised public concerns over competition between the food industry and the energy industry, and over the large amount of arable land required for crops. Hence, the attractiveness of second-generation technologies that enables the production of *bioethanol* (2G *bioethanol*) or other energetic molecules from cheaper feedstock such as ligno-cellulosic material from agricultural wastes. However, in order to enable the economically viable production of *bioethanol* from such ligno-cellulosic feedstock instead of those from starch or sugar, it is crucial to obtain (hemi)cellulose-degrading enzymes with improved efficiency, as these enzymes are major cost contributors in the *bioethanol* production process. Biotechnology has therefore become one of the key growth drivers for so-called "cellulosic" 2G *bioethanol*.[3] Due to its capacity to produce large amounts of cellulose-degrading enzymes, *Trichoderma reesei* has being extensively studied in various fields of white biotechnology, especially in biofuel production from lignocellulosic biomass.

Molecular directed evolution using either mutagenesis/screening methods[4,5] or molecular gene shuffling methods[6−8] or a combination of both procedures provides a fast track to obtain enzymes that better fit with the requirements of targeted industrial processes.[9,10] Gene shuffling methods allow for poolwise recombination of parental genes. Such poolwise recombination enables mutations from many parental genes to recombine in a single progeny, thus increasing the number of positive mutations that can be accumulated between two selection events and enabling a broader sequence space to be explored. L-Shuffling[11] is a ligation-based gene shuffling method, which permits the recombination of gene variants without the use of any polymerase, therefore reducing the risk of adding unwanted mutations. In this technology, the random recombination of parental gene fragments is assured by the ligation by means of a suitable DNA ligase of the ends of fragments of the parental gene variants hybridized onto an assembling template. Ligated and nonligated fragments are then denatured again, and cycles of hybridization−denaturation−ligation are repeated until full-length genes are rebuilt. This ligation-based recombination process enables randomized recombination of the parental gene fragments, hence maintaining

* Author for correspondence. Telephone: +33 (0)4 66 70 64 64. Fax: +33 (0)4 66 70 64 60. E-mail: lfourage@proteus.fr.
[†] Protéus SA.
[‡] IFP Energies Nouvelles.

(1) *2010 Ethanol Industry Outlook*; Renewable Fuels Association: Washington, DC, 2010 (http://www.ethanolrfa.org/page/-/objects/pdf/outlook/RFAoutlook2010_fin.pdf).
(2) *2010 Biofuels Production Forecast*; Global Renewable Fuels Alliance, **2010** (http://www.globalrfa.org/pr_032110.php).
(3) Margeot, A.; Hahn-Hagerdal, B.; Edlund, M.; Slade, R.; Monot, F. *Curr. Opin. Biotechnol.* **2009**, *20*, 372–380.
(4) Chodorge, M.; Fourage, L.; Ullmann, C.; Duvivier, V.; Masson, J. M.; Lefevre, F. *Adv. Synth. Catal.* **2005**, *347*, 1022–1026.
(5) Chodorge, M.; Lefevre, F.; Masson, J. M.; Fourage, L. Method of determining the mutational load of a gene library obtained by random mutagenesis of a particular gene and means for implementing same. WO/2006/003298, 2006.
(6) Stemmer, W. P. *Nature* **1994**, *370*, 389–391.
(7) Stemmer, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10747–10751.
(8) Zhao, H.; Giver, L.; Shao, Z.; Affholter, J. A.; Arnold, F. H. *Nat. Biotechnol.* **1998**, *16*, 258–261.
(9) Fourage, L.; Lefevre, F.; Wieland, S.; Weber, A.; Beckers, A.; Maurer, K. H.; Kottwitz; B. Alkaline protease variants having improved performance and washing and cleaning agents containing said alkaline protease variants having improved performance. WO/2005/118793, 2005.
(10) Bruehlmann, F.; Wahler, D.; Fourage, L. Modified 13-hydroperoxide lyases and uses thereof. WO/2009/001304, 2009.
(11) Dupret, D.; Masson, J. M.; Lefevre, F. Method for obtaining in vitro recombined polynucleotide sequences, sequence banks and resulting sequences. WO/2000/09679, 2000.

10.1021/op100218a
Published on Web 11/19/2010

and cumulating the DNA information of the parental genes and generating a high proportion of functional variants. Results of the molecular evolution of a key component of the *Trichoderma* cellulose-degrading enzymatic pool using L-Shuffling demonstrate how the resulting evolved secretome enables lowering the cost-contribution of the saccharification step.

Experimental Section

Materials. *p*-Nitrophenyl β-D-glucopyranoside (*p*NPGlc) was purchased from Sigma. Plasmid pET26b+ used for creating L-Shuffling libraries was supplied by Novagen.

Construction of Plasmid pET26cay. The cDNA encoding glucosidases was amplified by PCR with the specific primers SgeneA (5′-GGAATTCCATATGCTGGAGGCCGCCGACTGG-3′) and ASgeneA (5′-CCCAAGCTTCTAGGCGGTCAGGCT-GCC-3′); SgeneB (5′-GGAATTCCATATGGTTGTACCTC-CTGCAGGGAC-3′) and ASgeneB (5′-CCCAAGCTTCTA-CGCTACCGACAGAGTG-3′) and SgeneC (5′-GGAATTC-CATATGGAGACAAGCGAGAAGCAGG-3′) and ASgeneC (5′- CCCAAGCTTCTAGTATACGTCGAACTTGCC-3′). Amplification was performed with 30 cycles of 94 °C for 1 min, 60 °C for 45 s, and 72 °C for 1.5 min, with *Pfu* DNA polymerase (Promega) using the following PCR mix: 2 μL of DNA template at 20 ng/μL; 2 μL of primers at 10 pmol/μL; 4 μL of dNTPs at 5 mM each; 10 μL of 10x buffer; 1 μL of *Pfu* at 2.5 U/μL and 79 μL of *d*-H_2O. After digestion using *Nde*I and *Hind*III restriction nucleases, the 2.1-kb PCR products were cloned into pET26b+ vector using the corresponding restriction sites leading to pETcay construct. After validation by automated DNA sequencing with the T7 promoter and terminator primers, the resulting plasmids, named pETcayB, A, and C, respectively, were used to prepare the two L-Shuffling libraries as described previously.[12]

Screening of L-Shuffling Library. For the first round of evolution, *Escherichia coli* MC1061(DE3) colonies expressing β-glucosidase variants were grown in 96-well microtiterplates at 37 °C during 4 h and 20 °C during 20 h after induction with 100 μM IPTG in 150 μL of Luria−Bertani (LB) medium complemented with 60 mg/mL of kanamycin. For the second round, cells were grown during 20 h at 20 °C without IPTG induction. After centrifugation (4 min at 4000 g) and resuspension in 100 μL of 100 mM succinate buffer at pH 5.0 comprising 2.2 mM of *p*NPGlc, cells were incubated for 3 h at 23 °C. After this incubation, activities were estimated spectrophotometrically by measuring the release of *p*-nitrophenolate ion at 414 nm after adding one volume of Na_2CO_3. In each plate, the values were compared to the value obtained with cells expressing the WT-glucosidases tested in the same conditions.

Characterization of Glucosidase Variants. Growth conditions used for the characterization of the β-glucosidase variants were the same as the ones used during the screening of the second round of L-Shuffling (without IPTG induction). After centrifugation (5 min at 8000 g), *E. coli* MC1061(DE3) clones expressing the improved glucosidase variants were resuspended in 0.8 mL of 100 mM succinate buffer at pH 5.0 and different amounts of resuspended cell pellets were incubated 1.5 h using saturating concentration of *p*NPGlc at 50 °C.

Engineered *Trichoderma* Strains. The WT *bgl1* gene (Bgl1 + strain) or the 100B11 synthetic gene (TR3002 strain) was cloned in a vector allowing expression in the CL847 *T. reesei* industrial strain under control of the strong inducible *cbh*1 promoter. The transformation procedure was carried out as previously described.[13] The strain TR3002 was selected as the clone with the best specific β-glucosidase activity among 10 clones assayed for shake-flask cellulase production as previously described.[14]

Fermentation Conditions. The previously described fermentation conditions[15] were used in this study. The enzymes secreted in the supernatants were recovered by filtration as previsously described. Protein concentrations were assayed using the Lowry method.[16] β-glucosidase and Filter-Paper (FPase) activity were performed as previously described[17] using IUPAC standard procedures.

Wheat Straw Saccharification. Enzymatic cocktails secreted by *T. reesei* CL847 (expressing native *bgl1* gene), *T. reesei* CL847-bgl1+ (overexpressing native *bgl1* gene) and TR3002 (expressing 100B11 variant) strains, wild-type strain and engineered strain containing improved β-glucosidase respectively, were used for saccharification of steam exploded wheat straw under acidic conditions. Kinetics were determined in bioreactor (Bio-Laffite), under stirring conditions (500 tr/min). The wheat straw was pretreated by steam explosion (19 bar, 3 min) previously soaked in 0.05 M H_2SO_4 for 15 h. After filtration, the solid phase mainly containing cellulose and lignin was solubilized at 10% dry matter (DM) in a total volume of 2 L of 1 M acetate buffer at pH 4.8 and 48 °C. Protein concentration was determined using the Folin method.[18] Enzymatic cocktails were used at 20 mg and 5 mg per g of DM, and kinetics were carried out for 72 h. Enzymatic activities were inactivated by thermodenaturation, and after centrifugation and filtration, the released glucose was titrated with a glucose analyser using the glucose oxidase method (Analox, UK).

Results

Molecular Evolution of *Trichoderma reesei* β-Glucosidase Activity. In order to evolve β-glucosidase activity, a first round of L-Shuffling was carried out using two parental sequences sharing 70% of amino acid identity. For microplate expression, induced cultures using 100 μM of IPTG were needed in order to detect the glucosidase activity of the reference parental protein B (shown in red in Figure 1). It is important to note that activity of the second parental protein (protein A) was not detectable under these screening conditions. Around 20,000 clones were analyzed as described in the experimental design using 2.2 mM of *p*-nitrophenyl β-D-glucopyranoside (*p*NPGlc) as substrate.

(12) Ravot, R.; Masson, J. M.; Lefevre, F. *Methods in Microbiology: Application of Extremophiles*; Academic Press: London, 2005.

(13) Penttilä, M.; Nevalainen, H.; Rättö, M.; Salminen, E.; Knowles, J. *Gene* **1987**, *61*, 155–164.

(14) Lopes-Ferreira, N.; Margeot, A.; Mathis, H.; Fourage, L. Beta-glucosidase variants having improved activity, and uses thereof. WO/2010/029259, 2010.

(15) *Ibid*.

(16) Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. *J. Biol. Chem*. **1951**, *193*, 265–275.

(17) Herpoël-Gimbert, I.; Margeot, A.; Dolla, A.; Jan, G.; Mollé, D.; Lignon, S.; Mathis, H.; Sigoillot, J.-C.; Monot, F.; Asther, M. *Biotechnol. Biofuels* **2008**, *1*, 18–30.

(18) Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. *J. Biol. Chem*. **1951**, *193*, 265–275.

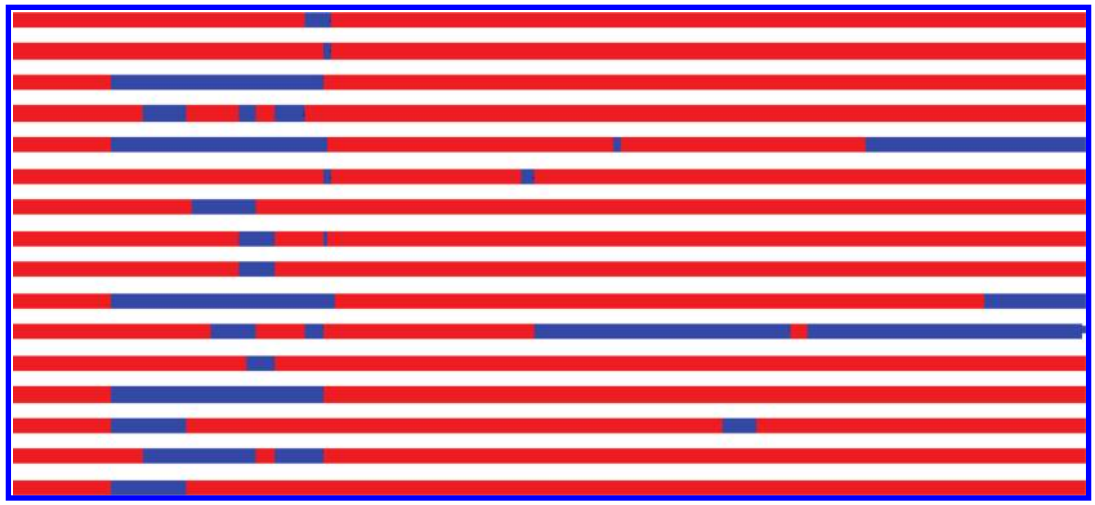

Figure 1. **Sequence alignment of improved variants resulting from L-Shuffling experiment. Red: fragments originating from the reference parental gene encoding protein B; blue: fragments originating from gene encoding protein A.**

Table 1. **Measurement of the β-glucosidase activity of variants and the reference protein (protein B) from the first round of L-Shuffling[a]**

clone numbers	k_{cat} (min^{-1})	factors of improvement
SH1-1	11.7	17
SH1-2	7.2	10
SH1-3	16.0	23
B	0.7	1

[a] k_{cat} values were estimated using saturating concentration of the substrate (10 mM of pNPGlc) and are compared to the k_{cat} of reference protein B for quantifying the improvement.

Table 2. **Measurement of the β-glucosidase activity of variants resulting from the second round of L-Shuffling[a]**

clone numbers	k_{cat} (min^{-1})	factors of improvement
SH2-1	169.7	242
SH2-2	159.1	227
SH2-3	100.9	144
B	0.7	1

[a] k_{cat} values were estimated using saturating concentration of the substrate (10 mM of pNPGlc) and are compared to the k_{cat} of reference protein B for quantifying the improvement.

Among the selected improved clones, 16 were sequenced, showing large sequence diversity even if a hot-spot was suspected in the N-terminal part of the protein. As expected, due to the chosen L-Shuffling strategy, the skeleton of the improved variants was based on the reference protein. In addition, among the best performers, up to seven recombination events per gene were observed (Figure 1).

These best performers were analyzed for their β-glucosidase activity improvement on *p*NPGlc. k_{cat} values were estimated using different amounts of enzymes at saturating concentration of the substrate (10 mM of *p*NPGlc). The k_{cat} value of the wild-type β-glucosidase A could not be determined because of its low activity in the tested conditions. The expression levels of the variants and wild-type β-glucosidases were the same when no IPTG induction was used for the validation of the activity improvement (data not shown). As shown in Table 1, after one round of molecular evolution, a strong improvement of the catalytic constant was observed for the best improved clone 164A2 (SH1-3: 23-fold improvement using *p*NPGlc substrate). Despite no observable activity with β-glucosidase A under screening and validation conditions, the introduction of DNA fragments from this sequence into the reference β-glucosidase B resulted in improved variants of the enzyme.

Using the 16 best performers from the first round of L-Shuffling and introducing the WT β-glucosidase C as an additional parental gene, a second round of molecular evolution was launched. The same screening strategy was applied except that IPTG induction was no longer necessary due to the strong improvement achieved during the first round. Around 20,000 clones were analyzed as described in the experimental design using 2.2 mM of *p*NPGlc as substrate. Among the selected improved clones, we determined the DNA sequence of 14 clones. The sequences confirmed the presence of a hot-spot in the N-terminal part of the protein (Figure 1). k_{cat} values of the three best improved clones were assessed (Table 2). Surprisingly, despite the high β-glucosidase activity of protein C in the tested conditions, no DNA fragment of that gene C was found in the sequence of the three best performers, although other less improved variants resulting from the same round of shuffling contained fragments originating from gene C.

Saccharification Experiment Using the Engineered *Trichoderma reesei* Secretome. β-Glucosidase activity of the best improved variant (SH2-1: 100B11) was also checked after cloning and production in *T. reesei* leading to TR3002 strain.

Table 3. **Comparison of the β-glucosidase activity of engineered *T. reesei* CL847-*bgl1*+ and TR3002 strains vs that of the native CL847 strain[a]**

	Trichoderma reesei strains		
	CL847	CL847-bgl1+	TR3002
factor of improvement of the β-glucosidase activity (compared to CL847 strain)	–	4.5	11.9
FPase specific activity (IU/mg) of cellulolytic cocktail	0.5	0.4	0.65

[a] CL847-bglI+ is a *T. reseei* strain CL847 overexpressing native β-glucosidase gene; TR3002 is a *T. reseei* strain CL847 expressing the best gene resulting from the second round of L-Shuffling (i.e. gene 100B11).

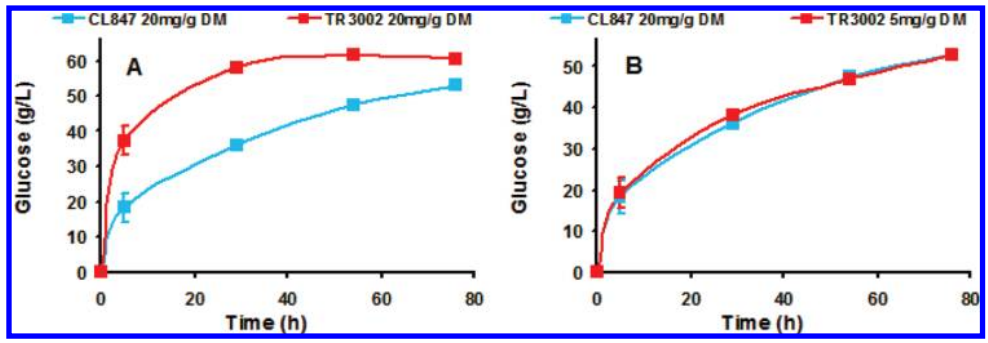

Figure 2. **Kinetics of wheat straw saccharification using either improved (red) TR3002 or wild-type (blue) CL847 *T. reesei* secretomes. (A - left side) equal amounts of secretome, *i.e.* 20 milligrams of secretome per gram of dry matter; (B - right side) 20 mg/g of wild-type secretome vs 5 mg/g of improved secretome. The release of glucose is quantified using HPLC.**

Table 3 shows a factor of improvement for the variant compared to both wild-type *T. reesei* CL847 strain[19] and a recombinant *T. reesei* strain (CL847-*bgl1*+) overexpressing a single additional copy of native β-glucosidase *bgl1* gene of *T. reesei*. The TR3002 strain shows a factor of improvement of ~12. The concentration of excreted β-glucosidase protein in this strain remains approximately 1.5−2-fold greater than the one of the WT strain, as estimated by two dimensional-electrophoresis.[17] Consistently, filter paper activity was slightly increased in the TR3002 strain, probably owing to cellobiose end-product inhibition release of exoglucanase enzymes (Table 3).

Then, cellulolytic cocktails produced by CL847 and TR3002 strains[15] were used for catalyzing steam-exploded wheat straw hydrolysis under the conditions described in the Experimental Section. The results obtained are shown in Figure 2. After 24 h, the maximum yield (60 g/L of glucose) is obtained when TR3002 secretome was used as biocatalyst, whereas after 72 h of incubation this maximum yield was not held using the same amount of CL847 secretome (20 mg of secretome/g of DM, Figure 2A). This major improvement enabled a 4-fold decrease in secretome loading (5 mg of secretome/g of DM) without any loss in activity performance of hydrolysis of a steam exploded wheat straw in bioreactor (figure 2B).

Conclusion

The present study demonstrates that random recombination approach using different sequences sharing a high identity (70% at the amino acid level), enables identifying strongly improved variants in order to design a bespoke enzyme. Although no structural information was available for the targeted reference protein and no β-glucosidase activity was demonstrated for the two parental proteins A and C before this study, using an appropriate screening assay and an efficient L-Shuffling technology, enzyme variants showing up to 242-fold improvement were identified after only two shuffling rounds. Moreover, after replacement of the *bgl1* gene by this improved β-glucosidase in *T. reesei*, a good productivity of the new gene in an industrial *Trichoderma* strain was observed without significant changes of ratio between the different cellulosic activities (cellobiohydrolases (~60−80% w/w); endoglucanases (~10−20% w/w) and β-glucosidase (~1% w)). The production by *T. reesei* of a new efficient enzymatic cocktail (in which β-glucosidase activity is 11 times higher, data not shown) allows a 4-fold decrease in cellulase loading without any loss in hydrolysis performance for degradation of a steam explosed wheat straw in bioreactors. This example demonstrates first how an evolved secretome through molecular evolution using *in vitro* recombination processes enables to lower the cost-contribution of the saccharification step. Second, this example shows that the optimal conditions defined by process developers should be used as a starting point for the selection and/or design of the "ideal" biocatalyst[20] and such approach requires multidisciplinary skills to be successfully addressed.

Acknowledgment

The authors thank the teams at the Biotechnology Department of IFP and Protéus for their contributions. This work was supported by grants from French PNRB (Programme National de Recherche sur les Bioénergies).

Received for review August 6, 2010.

OP100218A

(19) Durand, D.; Clanet, M.; Tiraby, G. *Enzyme Microb. Technol.* **1988**, *10*, 341–346.

(20) Tramper, J. *Biotechnol. Bioeng.* **1996**, *52*, 290–295.