Improved Catalytic Efficiency and Active Site Modification of 1,4-β-d-Glucan Glucohydrolase A from *Thermotoga neapolitana* by Directed Evolution*

James K. McCarthy‡, Aleksandra Uzelac, Diane F. Davis, and Douglas E. Eveleigh§

From the Department of Biochemistry and Microbiology, Cook College, Rutgers, The State University of New Jersey, New Brunswick, New Jersey 08901

*This work was supported by United States Department of Agriculture Competitive Grant 2001-35504-10158, the McNut-Chemie-Stennis Program, a National Institutes of Health and Rutgers University Biotechnology Training Fellowship (to J. K. M.), and the NJ Agricultural Experiment Station Grant K-NJ00133-01-04. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address: Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92039-0204.

§ To whom correspondence should be addressed. E-mail: eveleigh@aesop.rutgers.edu.

© 2004 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in U.S.A.

Published, JBC Papers in Press, December 4, 2003, DOI 10.1074/jbc.M305642200

**Thermotoga neapolitana** 1,4-β-d-glucan glucohydrolase A preferentially hydrolyzes cello-oligomers, such as cellotetraose, releasing single glucose moieties from the reducing end of the cello-oligosaccharide chain. Using directed evolution techniques of error-prone PCR and mutant library screening, a variant glucan glucohydrolase has been isolated that hydrolyzes the disaccharide, cellobiose, at a 31% greater rate than its wild type (WT) predecessor. The mutant library, expressed in *Escherichia coli*, was screened at 85 °C for increased hydrolysis of cellobiose, a native substrate rather than a chromogenic substrate. The most active clone caused an isoleucine to threonine amino acid substitution at position 170. Structural models for I170T and WT proteins were derived by sequence homology with Protein Data Bank code 1BGA from *Paenibacillus polymyxa*. Analysis of the WT and I170T model structures indicated that the substitution in the mutant enzyme repositioned the conserved catalytic residue Asn-163 and reconfigured entry to the active site.

The multistep transformation of cellulose biomass from the biopolymer to its glucose components is mediated in nature by the enzymatic hydrolysis of the polymer and its glucan intermediates. The potential of this microbial bioprocessing for production of oxymethyl ethers and transportation fuels has intensified the biotechnological focus on and basic research into the enzymology of cellulases (1–6). The large scale research initiatives underway to significantly improve biomass conversion yields (7) are supported by previous work that developed mutant organisms with enhanced activity against cellulosic substrates (8) and harnessed β-glucosidases, or β-glycosidases, for industrial application (9). *Thermotoga*, a Eubacterial genus diverging near the root of the Bacteria domain, evolved or received through lateral gene transfer (10, 11) a unique set of glycosyl hydrolases to degrade the variety of polymeric glucans found in its environment (12). In *Thermotoga neapolitana*, a Gram-negative, fermentative hyperthermophile (13), 1,4-β-d-glucan glucohydrolase (GghA)1 (EC 3.2.1.74) is one of three enzymes comprising the cello-oligosaccharide utilization pathway (14, 15). With optimal activity at 85 °C and low affinity for cellobiose (its preferred substrate is cellotetraose), GghA is an excellent target for a directed evolution effort. By "redesigning" GghA as a thermostable cellobiase with increased catalytic efficiency (*k*<sub>cat</sub>/*K*<sub>m</sub>) for cellobiose hydrolysis, the modified enzyme could play a significant role in a high temperature model cellulase system.

Directed molecular evolution, a widely used, dynamic tool for protein engineering (16, 17), draws its power from iterative cycles of random nucleotide-sequence mutagenesis, colony expression, and screening of mutants (18, 19). Its feasibility depends on a dynamic screening assay to identify enhanced mutant clones against a background of (tens of) thousands of normal and null clones. To screen for an evolved β-glucosidase (GghA) with increased activity toward a native substrate, cellobiose, we developed a thermostable (85 °C) coupled enzyme assay with glucokinase and glucose-6-phosphate dehydrogenase from *Thermotoga maritima* (20). Although β-glucosidase activity is rapidly and routinely assayed with chromogenic substrates such as various *p*-nitrophenyl-β-d-glycosides, using the *pNP*-β-d-glucoside (*pNPG*) analog in determining increased hydrolysis of cellobiose by GghA mutants might have resulted in evolving an “improved” enzyme that lacked a useful industrial application. As *pNPG* has both glycone and aglycone moieties, greater mutant enzyme activity might indicate that an amino acid substitution had occurred in a residue responsible for aglycone recognition or binding. Such substitutions might have a wholly different effect on cellobiose that lacks the steric and electrostatic characteristics of the aglycone moiety.

The first step in the directed evolution of GghA, expressed in *Escherichia coli*, has produced an enzyme with β-glucosidase-like character. The GghA mutant, based on a single amino acid substitution six positions from the catalytic residues conserved in all family 1 β-glucosidases, has a *V*<sub>max</sub> for cellobiose 31% higher than its wild type parent and shows a 31% increase in

---

1 The abbreviations used are: GghA, 1,4-β-d-glucan glucohydrolase; TEA, triethanolamine; WT, wild type; *pNPG*, *p*-nitrophenyl-β-d-glucoside.
catalytic efficiency when assayed at 85 °C, the temperature optimum of the wild type enzyme. In the absence of a published crystal structure for WT GghA, a three-dimensional molecular model, based on sequence homology to Pseudomonas polymyx (21), was derived for both wild type (WT) and mutant GghA proteins. Our analysis of the parent and mutant apoenzymes was facilitated by the breadth of structural data available for family 1 glycosyl hydrolases (22); the conservation of the eight-stranded TIM barrel-fold (23); the known residues involved in positioning and substrate recognition (24–26); the mechanism of hydrolysis; (27) the general acid catalysis and the conditions of specific residues in the active site microenvironment (27, 28).

Comparing the enhanced GghA apoenzyme to known family 1 structures enabled us to suggest how its single amino acid substitution produces the observed functional improvement.

**EXPERIMENTAL PROCEDURES**

**pTNGghA1 Expression System**—Isolation, purification, and kinetic characterization of WT T. neapolitana GghA from E. coli have been described previously (15, 20). Plasmid pTNGghA1, constructed from a pET-24d vector (Novagen, Madison, WI), contains the gghA gene under the control of an inducible T7lac promoter and includes a His6 sequence tag at the C terminus of the gene. Plasmids were harvested with the QIAfilter plasmid purification kit (Qiagen, Valencia, CA) and separated by agarose gel electrophoresis. DNA was run on the gel with 1 kb and 3 kb ladder (Invitrogen, Carlsbad, CA) to provide concentration. DNA was ligated into the linear vector (described above) using rapid DNA ligation protocols (32), into commercially available pET-24/28 vectors (Novagen, Madison, WI) and concentrated by SpeedVac vacuum filter plate (Millipore, Bedford, MA).

**Vector Preparation**—A pET-24d vector was constructed with a 900-bp spacer sequence flanked with HindIII restriction sites at the 5′ and 3′ ends inserted into the multiple cloning site of the vector. The construct was transformed into E. coli DH5α competent cells, plated on LB agar with 30 μg/ml kanamycin selection. A single colony from the plate, grown over night (37 °C), was used to inoculate 50 ml of LB broth (as above) (500 ml flask, shaking, 37 °C). One ml of the overnight broth culture was removed for glycerol stocks; from the remainder, the pET-plasmid plasmid was harvested using a QIAfilter plasmid purification kit. As needed, the pET-plasmid plasmid was double digested with Nhel and XhoI (New England Biolabs, Beverly, MA). The linear vector DNA was gel purified, cleaned up with the GeneClean spin kit (Qbiogene, Bio 101 Systems, Carlsbad, CA) and concentrated by SpeedVac vacuum centrifugation (Savant, Farmingdale, NY).

**Random Mutagenesis—Primers, 5′-GAAGAGGATATACATGG-CTAGCCTGAAAGG and 3′-GGTGTCGATGCTCGTTTGGT, were designed for error-prone PCR amplification (18, 19, 29, 30) of the gghA gene (1335 bp) in pTNGghA1 to include regions of the vector promoter and includes a His6 sequence tag in-frame. A 500-ng error-prone PCR mixture of mutagenic buffer (7 mM MgCl2, 50 mM KCl, 10 mM Tris-HCl (pH 8.5), 0.01% gelatin (w/v)), an asymmetrical mixture of dNTPs (0.2 mM DATP, 0.2 mM dGTP, 1.0 mM dCTP, and 1.0 mM dTTP), 250 pmol of each primer, 500 ng of pTNGghA1 template DNA, and a range of MnCl2 (0.075–0.15 mM final concentration) was aliquoted into 200-μl reagent mixture and amplified for 20 cycles: 60°C, 30 s; 94°C, 1 min; 72°C, 1 min. To start the reaction, 90°C, 30 s; 94°C, 30 s; 72°C, 1 min 30 s. After denaturation for 3 min at 96°C, all PCR mixtures were amplified for 20 cycles: 6 s at 94°C, 30 s at 61°C, and 1.45 min at 72°C with 0.7 units/50 μl of reaction of recombinant Taq polymerase (31).

To determine the appropriate mutation frequency, 3 parallel 500-μl reactions were set up with different MnCl2 concentrations, and amplified, producing 3 libraries of mutants.

**Ligase Transformation—**PCR products were cleaned up with the GeneClean spin kit, quantified by agarose gel electrophoresis, and double digested with Nhel and XhoI. Amplon digestions were also cleaned up, then concentrated (as above). Variant library sequences were ligated into the linear vector (described above) using rapid DNA ligation (Roche Applied Science) and transformed, using standard protocols (32), into commercially available E. coli non-expression host strain BL21 (DE3). The ligase transformation efficiency of a library was plated out as follows: 75 μl in one plate and the remaining volume (~225 μl) of the transformation reaction in a second plate. Transformed cells were incubated overnight at 37 °C. Transformation efficiencies were determined by comparing colony numbers on a 75-μl mutant library plate and on a 75-μl negative control plate (no insert).

To re-isolate the library, 1.5 ml of LB was added to each mutant plate. Colonies were gently removed with a cell scraper, collected, and combined. Plasmid preparations were made from each mutant library. Ten nanograms of mutant library plasmid DNA was used to transform chemically competent E. coli strain BL21(DE3) cells (Invitrogen) following standard protocols. A portion (140 μl) of each BL21(DE3) transformant was plated out on 22-μm bioassay trays (Genetic, New Milton, UK) with the addition of 250 μl of SOC (rich broth medium) to aid in spreading. Clones were incubated overnight at 37 °C.

**Harvesting Mutant Clones**—Microtitre plates were prepared with 200 μl of LB including kanamycin selection (30 μg/ml) per well. Transforms of the mutant libraries were picked from the bioassay trays and transferred to chemically competent DH5α by the addition of an appropriate polyethylene glycol solution (1.2 μM filter plate atop a vacuum manifold (Millipore, Bedford, MA). Filtrates of heat-treated, partially purified protein were collected in fresh plates, by applying a vacuum (18 psi) sustained by a pump, and were stored at 4 °C until assayed. The protein concentrations of the filtrates were estimated by the Bradford dye binding method (Bio-Rad) with bovine serum albumin as the standard. Control experiments using gel filtration (Bio-Rad) were done in the range of column/ buffer (50 ml/50 ml). Cells were heat killed empirically on a heatblock designed for microtitre plates and heated (85 °C, 3 min). Prior to use, cell broth was dissolved in phosphate citrate buffer (pH 5.2) (for the 1 ml cuvette assay for WT and mutant GghA, the cell broth was prepared in 50 mM Na2HPO4 buffer, pH 5.2), and incubated with glucose oxidase. Because of the sensitivity of the coupled assay, all cell broth had to be pretreated with glucose oxidase to remove glucose impurities (20). The pH of the cell broth preparation was then adjusted with TEA base to pH 7.3. To start the reaction, 90 μl of pre-heated (85 °C) substrate/cofactor mixture (cellobiose, 20 mM; ATP, 3.5 mM; and NADP, 1 mM) in TEA buffer was added with a multichannel pipettor to the hot assay plate on the heatblock. To prevent condensation on the lid of the plate during preheating and the assay, a heat block (85 °C) was inverted and placed on top of the assay plate lid. The reaction was stopped after 3 min by the addition of 25 μl of ice-cold stop buffer (50% assay buffer, 50% ethanol (95%)). In a control experiment, the stop buffer showed no distortion of absorbance at 340 nm. After the addition of the stop buffer, plates were allowed to cool (1 min). Changes in absorbance due to glucose oxidation (A595) and NADPH (A340) were measured against an enzyme-free reference well. To ensure that changes in absorbance were based on random mutagenesis effects, and not on media evaporation/condensation or well to well cross-talk, potential positive mutants were re-grown and re-screened in the range of column/row arrangements; un-inoculated columns separated the WT wells from control. Each strain, positive mutant or strain on LB kanamycin plates and single colonies were grown overnight (3 ml of LB, as above) for glycerol stocks and plasmid purification.

**Screening for Increased GghA Activity**—Wild type and mutant clones were screened for activity in 96-well plates at 85 °C using a coupled enzyme assay derived from T. maritima (20). The coupled assay, linking the formation of NADPH (measured at 340 nm) to the hydrolysis of cellobiose by GghA or GghA mutants (20) was scaled down to a 200-μl reaction. Thermostable glucokinase (1 unit ml−1) and glucose-6-phosphate dehydrogenase (0.4 units ml−1) in triethanolamine (TEA) buffer (80 mM TEA, 4 mM MgCl2, pH 7.8) were aliquotted into assay plate wells; the cell filtrate of individual clones (15 μl) was transferred column-wise to the assay plate bringing the volume in each well to 90 μl. Filtrate volumes were adjusted empirically on a heatblock designed for microtitre plates and heated (85 °C, 3 min). Prior to use, cellobiose was dissolved in phosphate citrate buffer (pH 5.2) (for the 1 ml cuvette assay for WT and mutant GghA, the cellobiose was prepared in 50 mM Na2O4 buffer, pH 5.2, and incubated with glucose oxidase. Because of the sensitivity of the coupled assay, all cell broth had to be pretreated with glucose oxidase to remove glucose impurities (20). The pH of the cellobiose preparation was then adjusted with TEA base to pH 7.3. To start the reaction, 90 μl of pre-heated (85 °C) substrate/cofactor mixture (cellobiose, 20 mM; ATP, 3.5 mM; and NADP, 1 mM) in TEA buffer was added with a multichannel pipettor to the hot assay plate on the heatblock. To prevent condensation on the lid of the plate during preheating and the assay, a heat block (85 °C) was inverted and placed on top of the assay plate lid. The reaction was stopped after 3 min by the addition of 25 μl of ice-cold stop buffer (50% assay buffer, 50% ethanol (95%)). In a control experiment, the stop buffer showed no distortion of absorbance at 340 nm. After the addition of the stop buffer, plates were allowed to cool (1 min). Changes in absorbance due to glucose oxidation (A595) and NADPH (A340) were measured against an enzyme-free reference well. To ensure that changes in absorbance were based on random mutagenesis effects, and not on media evaporation/condensation or well to well cross-talk, potential positive mutants were re-grown and re-screened in the range of column/row arrangements; un-inoculated columns separated the WT wells from control. Each strain, positive mutant or strain on LB kanamycin plates and single colonies were grown overnight (3 ml of LB, as above) for glycerol stocks and plasmid purification.

**Directed Evolution of T. neapolitana Glucan Glucohydrolase**
Purification of Wild Type and Mutant Glucan Glucohydrolase—The growth, expression, and purification of recombinant, wild type GghA from *Thermotoga* sp. has been described previously (15, 20, 35). The same protocols were followed for GghA mutants. Briefly, a loopful of glycerol stock was inoculated into 3 ml of LB containing 30 μg/ml kanamycin and grown to an A of 0.5 at 595 nm. The 3-ml culture was used to inoculate 500-ml of LB medium containing 30 μg/ml kanamycin and allowed to grow to an A of 0.9–1.0 (Fernbach flask, shaking, 37°C); protein expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (1 mM final concentration), and the culture was grown for an additional 4 h. Cells were pelleted by centrifugation (6,000 × g, 4°C, 15 min) and resuspended in 50 ml Tris-HCl (pH 7.5). Whole cells were lysed with lysozyme (4 mg ml⁻¹ of resuspension buffer) and shaken (room temperature, 15 min), cell lysates were heat treated (75°C, 15 min) to denature heat-labile proteins, then sonicated on ice (4 bursts, 45 s). Supernatants were recovered after centrifugation (15,000 × g, 30 min, 4°C), equilibrated with 300 mM NaCl, 50 mM NaH₂PO₄, and 10 mM imidazole, and loaded onto a Hi-Trap chelating column (Amersham Biosciences). Bound protein was eluted by a linear imidazole gradient (10–500 mM imidazole). Peak fractions were assayed for activity and analyzed by SDS-PAGE (36) stained with Coomassie Blue. Concentrations of pure protein were determined by measuring absorbance at 280 nm (ε GghA = 11.9 × 10⁴ m⁻¹ cm⁻¹) (37).

Enzyme Assays—Assays of purified, wild type and mutant GghA were performed at 85°C; the temperature optimum was previously established for the wild type (15), using a Shimadzu UV-265 spectrophotometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostatically controlled circulating water bath connected to the photometer. Assay temperature was maintained in the cuvette by a thermostat...
be estimated from the percent of null clones that arise in a mutant library (44, 46) with 30–40% null clones corresponding to a 0.2–0.3% mutation frequency. Fig. 1B shows the activity profile for three mutant clone libraries (192 clones each); the graph shows the data near zero activity to emphasize the differential percentages of each library. The libraries were designated IC, ID, and IE. With ~35% null mutants, library IE (MnCl₂, 0.075 M) provided the appropriate mutation frequency and was chosen for further screening.

In the first generation of directed evolution, 1500+ colonies were screened for increased activity toward cellobiose at 85 °C using the coupled enzyme assay. Activity corresponds to absorbance of NADPH at 340 nm. Seven clones showed increased absorbance as compared with wild type. Three potential positive mutants (Fig. 2) were re-grown in 96-well plates, and the cell filtrates were re-screened. A second cycle of growth and screening confirmed that all three were more active than the wild type. Mutant IE1E10 (IE-first generation, Library E; 1-plate #1; E10-well identifier) had 33% greater absorbance; IE1D5 showed a 41% increase, and IE8A9 (I170T) had an increase in absorbance of 45% as compared with the wild type. I170T, showing the highest absorbance, will be used for the next generation of mutagenesis.

Sequence Analysis—Complete DNA gene sequences were obtained for two positive (IE1D5 and I170T (IE8A9)) and two null mutants (IE1A5 and IE2H10). Eleven base shifts were identified over the 5,300+ bases sequenced (Table I). The actual mutation frequency of 0.21% was within the desired range of 0.2–0.3%. There were 8 transitions and 3 transversions. The T → C base shift in the most active mutant, I170T, caused an I170T change in the amino acid sequence. The wild type Ile-170 is 6 amino acids from Glu-164, the conserved catalytic acid-base residue. Along with the adjacent Asn-163, involved in hydrogen bonding with the substrate at the −1 position, the pair of residues are strictly conserved in family 1 and 5 glycosyl hydrolases (47). Comparative sequence analysis by WU Blast (33, 34) indicated that much of the sequence from residue 156 to 171, VKHWITLNEPVWVAVV, including the conserved catalytic glutamate and asparagine residues, is also highly conserved (bold face) in β-glucosidases and thioglucosidases in the Brassicales. Even isoleucine 170 (underlined) is conserved in a β-glucosidase from the fungus, Humicola grisea var. thermoidea and in one from the plant, Arabidopsis thaliana. As indicated in Table I, the substituted residues of both the positive and the null clones have been identified with elements of secondary structure in which they might participate. It is not surprising that the A → T transversion in IE1A5 produced a null mutant as the N163I substitution replaces the conserved active site asparagine involved in catalysis with a bulky hydrophobic residue.

Comparative Kinetic Characterization—Purified WT GghA and evolved I170T proteins each ran electrophoretically as a single band with an approximate molecular mass of 52–53 kDa, correlating with the DNA sequence data for gghA with the addition of the coding sequence for six histidines. Although WT GghA kinetics have been previously characterized (15, 20, 35), the current results are from identical and/or side by side experiments with both proteins performed under standard conditions. Fig. 3 shows the relative rates of the hydrolysis of cellobiose by wild type GghA and I170T proteins over a range of cellobiose concentrations. Kinetic parameters for both proteins were determined, and Table II summarizes the changes in kinetic behavior toward cellobiose arising from the I170T substitution. Although the Km values for both proteins were nearly the same, the maximum enzyme activity (Vₘₐₓ) of I170T, as compared with its parent, increased by 31% (108 ± 3 versus 75 ± 2 units mg⁻¹). The kₗₑₐ and the catalytic efficiency (kₗₑₐ/Kₘ) for I170T correspondingly increased 31%.

Protein Modeling—Crystal structures have been determined for a number of family 1 β-glucosidases (21, 24, 48). The enzyme from B. (now Paenibacillus) polymyxa, determined at 2.4 Å resolution (Protein Data Bank file 1BG), was used to construct models of both WT GghA and I170T apoenzymes. The structure models (Figs. 4, 6, and 7) suggested the relative locations of the substituted residues in the tertiary structure of the proteins (Table I). The models were refined by energy minimization using the Insight 2000 package, and their stereochemical qualities were assessed. Ramachandran plots obtained from PROCHECK showed that of the non-glycine, non-proline residues in wild type GghA, 98.2%, were within the allowed regions with 70.2% in the most favored regions. The values for the I170T mutant were similar, with 97.4% of the non-glycine, non-proline residues falling within the allowed regions of the plot and 73.8% found in the most favored regions. Solvent accessible surface area estimates (43) indicated that both the wild type isoleucine and the mutant threonine are buried at position 170. In the model structures of the apoenzymes, residue 170 is seen as a building block in the side wall of the active site cleft just beneath two partially accessible aromatic residues, Tyr-175 and Trp-66. The derived models of GghA and I170T (Fig. 4) fit well into the TIM barrel (ββα)₈ structures characteristic of the glycosidase 4/7 superfamily (23). The location of the conserved catalytic residues Asn-163, Glu-164, and Glu-349 at the C termini of the fourth and seventh β-strands, respectively, further classify GghA and I170T with other family 1, 2, 5, 10, and 17 glycosidases in the 4/7 superfamily of glycosyl hydrolases (49, 50).

**DISCUSSION**

Using protein engineering techniques of directed evolution, we have identified a first generation GghA mutant from T. neapolitana that, bearing a single amino acid substitution, I170T, significantly raises the Vₘₐₓ of the enzyme and increases its catalytic efficiency for cellobiose (Table II). The mutant, I170T, was picked from a library of 1,500 clones using a screening protocol based on the use of the recently reported thermostable, 96-well plate coupled enzyme assay. Developed with T. maritima glucokinase and glucose-6-phosphate dehy-
Directed Evolution of T. neapolitana Glucan Glucohydrolase

**Table I**

| Mutant clone | Activity | Base shift | Amino acid substitution | Location | Conserved residue |
|--------------|----------|------------|-------------------------|----------|------------------|
| IE1D5        | %        | A → C     | K90N                    | Surface  | No               |
| H11006       |          | T → C     | I170T                   | Buried   | No               |
| H11006       |          | A → T     | N163I                   | Active site | Yes          |
| IE1A5        | 17       | G → A     | V210M                   | Surface  | No               |
| IE2H10       | 30       | A → G     | K282R                   | Surface  | No               |
|              |          | A → T     | V308A                   | Surface  | No               |
|              |          | T → C     | F393S                   | Helix 5' | No               |
|              |          | A → T     | K317N                   | Loop C' | No               |
|              |          | C → T     | V380V                   | Sheet 5' | Yes              |
|              |          | A → G     | K410R                   | Surface  | No               |

*Activity as compared to wild type GghA.
*Secondary structure prediction (15).
*Secondary structure comparison to β-glucosidase from Sulfolobus solfataricus (24).

**Fig. 3.** Wild type versus I170T activity. Reaction rates of WT GghA (○) and mutant (I170T) (●) were measured over a range of cellubiose concentrations (1.5–200 mM), using the coupled enzyme assay to monitor formation of NADPH at 340 nm. One unit (U) of GghA activity corresponds to 2 μmol of glucose produced per min at 85 °C. Reaction mixtures contained TEA buffer (80 mM TEA, 4.4 mM MgCl₂, pH 7.8), cellubiose (pretreated with glucose oxidase), WT or I170T GghA (0.2 units/ml), glucokinase (1 unit/ml), glucose-6-phosphate dehydrogenase (0.4 units/ml), ATP (3.5 mM), and NADP (1 mM). All assays were performed at 85 °C. Vₘₐₓ values were determined by non-linear regression, and the curves were derived from best-fit values of the data. For the WT, r² was 0.996; for I170T, 0.997. The p values for both regressions were <0.0001.

**Table II**

| Enzyme  | Kₘ | Vₘₐₓ | kₘₐₓ | kₘₐₓ/Kₘ |
|---------|-----|------|-------|----------|
| GghA    | 43  | 75   | 64    | 1.6      |
| 1170T   | 43  | 75   | 92    | 2.2      |

**Fig. 4.** Secondary structure model of mutant GghA, I170T. The α-helices and β-strands are colored gray and yellow, respectively. Loops are colored blue; aperiodic secondary structure is in green. The ball and stick rendering of Glu-164 is in red, and Thr-170 is in green. The figure was produced using Insight 2000.
Directed Evolution of T. neapolitana Glucan Glucohydrolase

A well to well comparison of two microtiter plates further suggested there was little correlation between enzyme assays for cellobiose and those for pNPG (Fig. 5). Based on data collected in establishing the mutation frequency of given GghA libraries, where pNPG activity assays were used to rapidly determine the percentage of active clones on a 96-well plate, there were a number of instances where an identical plate, albeit grown at different times, was assayed for both for cellobiose and pNPG hydrolysis. The qualitative analysis shown in Fig. 5 indicated that there was only a 25% correlation between the cellobiose-coupled assay and the pNPG assay for any given mutant (well assignment) on the 96-well plate.

Family 1 1,4-β-D-glucan glucohydrolases exhibit a substrate preference for longer cello-oligomers, that is, as the degree of substrate polymerization (DP) increases the catalytic efficiency (kcat/Km) of the enzyme increases (15, 21). In the barley β-glucosidase βII, multiple surface residues, Trp, Tyr, Phe, and His, have been characterized as the binding subsites of the glucosyl units of the longer chain cello-oligomers (47), enabling the conserved catalytic mechanism of the enzyme to cleave single glucose units off the reducing end of the polymer in an energetically efficient manner. Breaking the scissile bond between the first and second glucoses eases the distorted conformation of the reducing glucosyl residue, as has been suggested for other glycosidases (55), and provides the energy to propel the substrate into position for the next cut. It is not surprising then that GghA shows little preference for cellobiose. Indeed, there is a 20-fold difference in the wild type Kcat values of the enzyme for cellotetraose and cellobiose, 2.15 versus 42.6 mm, respectively (15, 20, 35); furthermore, T. neapolitana expresses another catalytic enzyme, cellobiose phosphorylase, to utilize the disaccharide (15). The significant increase in the maximum rate of cellobiose hydrolysis and overall efficiency of the mutant enzyme, without a change in its Kcat value for cellobiose (Table II), suggests that the I170T substitution has affected the substrate-enzyme entry and/or exit interactions.

Structural changes in the mutant protein, I170T, as compared with WT GghA were investigated by molecular modeling. These models (Figs. 4, 6, 7), derived from β-glucosidase A from P. polymyxa (Protein Data Bank code 1BGA), enabled us to envision possible adjustments in three-dimensional structure and enzymatic mechanism that led to the 31% increase in the catalytic efficiency of the mutant. As the changes in I170T were relatively substantial, but the observed increases in catalytic activity, although significant, were not large, it is important to re-emphasize that the structural models were based on the apoenzyme conformations of the WT and I170T. The most prominent feature of the I170T apoenzyme model is the oval, crater-like shape of the active site. More characteristic of family 1 β-glycosidases (53, 56), the I170T active site opening is considerably different from that of the wild type model in which aromatic residues, the sugar binding subsites (see above), protrude into the narrow cleft. At the bottom of both are the highly conserved glutamates, Glu-164 and Glu-349, the catalytic acid and nucleophile, respectively, of family 1 glycosyl hydrolases (22, 23, 53, 57). The threonine for isoleucine substitution at residue 170, replacing a hydrophobic residue with a non-charged polar side chain, also changes the geometry of its immediate vicinity, presumably making new hydrogen bonds and “tightening” the turns of the short coil structure of which it is a part. This short helical motif, seen in the crystal structure representations of several β-glucosidases (21, 24, 56), runs parallel to the inside of the TIM barrel (Fig. 4). Starting just beyond the N terminus of the β3 strand, it includes not only the substituted I170T, but also conserved catalytic residues Asn-163 and Glu-164. The tightening of this coil (Fig. 6) results in a repositioning of several catalytic residues in the active site. Asn-163 is rotated completely away from its role in stabilizing one of the reducing-sugar hydroxyls. His-119, its side chain rotated 180° (Fig. 7), appears to provide both its nitrogen atoms for hydrogen bonding with that same sugar moiety. The two conserved glutamates, Glu-164 and Glu-349, the catalytic acid and nucleophile, respectively, of family 1 glycosyl hydrolases (22, 23, 53, 57), have been compressed. Other single amino acid substitutions in this conserved T/L/NEP region of the β3 strand have led to nearly complete loss of activity: the null mutant N163I reported here and in Verdoucq et al. (54). By superimposing large stretches of amino acids and looking for shifts in the areas not included in the superimpositions, qualitative comparisons can be made between the native and mutant model proteins. Two other tertiary structural changes appear to play a role in

![Fig. 5. Correlation of activity between substrates: pNPG and cellobiose.](Image)

![Fig. 6. Superimposition of WT and I170T GghA amino acid residues 140–175.](Image)
the geography of the active site of I170T. Upstream of Thr-170, in the loop between helix 3 and the b4 strand (Fig. 7), a vertical shift in the backbone angles repositions the side chain of Lys-157 at the back of the active site pocket. Its relocation suggests the formation of new hydrogen bonds with residues of the b6 strand. Another change in backbone angles, far downstream of Thr-170, causes the rings of Trp-396 to rotate up 45°, which in turn repositions Trp-322 up against the long wall of the crater. This shift appears to make the highly conserved catalytic core Glu-20, His-119, Asn-163, Glu-164, Asn-220, Glu-349, Glu-403, and Trp-404 more accessible to the substrate. Indeed, the increase in kcat/Km of I170T could be the result of the more open conformation of the mutant and be indicative of a higher on rate for cellobiose. The new conformation also suggests a faster dispersion of the product (glucose) away from the catalytic microenvironment after hydrolysis, a conformational change in the enzyme, as Rignall et al. (5) also noted, that, in effect, makes glucose a better leaving group.

By combining error-prone PCR mutagenesis with the high temperature, coupled assay, we have developed a robust, reliable method for the directed evolution of a thermostable glycosyl hydrodrolase. The single amino acid substitution I170T, so close to the catalytic center of the enzyme, provides an excellent opportunity to consider the conserved strength and evolutionary diversity of family 1 glycosyl hydrolases. Despite the "modeling" of the entry to the active site of the enzyme to better accommodate the smaller substrate, cellobiose, the conserved catalytic tools of the core b-glucosidase were retained. Within the framework of the national efforts to improve the efficiency of, and lower the costs associated with the transformation of biomass cellulosics to biofuels, the results of this first step in the evolution of a thermophilic b-glucosidase from a glaucan glucohydrolase hold great promise for the development of a key component of the model cellulase system of the future.

Acknowledgments—We thank Dr. Gerben Zylstra, Dr. Peter C. Kahn, Dr. Theodore Chase, Jr., Dr. Dinesh Yernool, and Dr. Gavin Swiatek for assistance and ongoing support of this work.

REFERENCES
1. Warren, R. A. (1996) Annu. Rev. Microbiol. 50, 183–212
2. Sekon, J., Adney, W. S., Himmel, M. E., Thomas, S. R., and Karpus, P. A. (1996) Biochemistry 35, 10648–10660
3. Sirois, M., Lehtin, J., Linden, M., Margollesclaire, E., Reinikainen, T., and Teeri, T. T. (1997) J. Bacteriol. 179, 49–57
4. Spiridonov, N. A., and Wilson, D. B. (2001) Curr. Microbiol. 42, 295–301
5. Rignall, T. R., Baker, J. O., McCarter, S. L., Adney, W. S., Vinzant, T. B., Becker, S. R., and Himmel, M. E. (2002) Appl. Biochem. Biotechnol. 96–100, 383–394
6. Andre, G., Kanchanaweng, P., Palma, R., Cho, H., Deng, X., Irwin, D., Himmel, M. E., Wilson, D. B., and Brady, J. W. (2003) Protein Eng. 16, 125–134
7. Sheehan, J. S., and Himmel, M. E. (2001) Agric. Food Industry Hi-Tech 12, 54–57
8. Montecourte, B. S., and Eveleigh, D. E. (1977) Appl. Environ. Microbiol. 34, 777–782
9. Neidleman, S. L. (2000) in Applied Biochemistry in Specialty Chemicals and Pharmaceuticals (Saha, B. C., and Demirjian, D. C., eds) Vol. 776, pp. 14–49, ACS Symposium Series, American Chemical Society, Washington, D. C.
10. Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G. (2002) Mol. Biol. Evol. 19, 2226–2238
11. Nesbo, C. L., Nelson, K. E., and Doolittle, W. F. (2002) J. Bacteriol. 184, 4475–4488
12. Sunna, A., Moracci, M., Rossi, M., and Antranikian, G. (1997) Extremophiles 1, 2–13
13. Huber, R., and Stetter, K. (1992) in The Prokaryotes (Balows, A., Truper, H. G., Dworkin, M., Harder, W., and Schleifer, K.-H., eds) Vol. IV, 2nd Ed., pp. 3809–3815, Springer-Verlag, New York
14. Bok, J. D., Yernool, D. A., and Eveleigh, D. E. (1998) Appl. Environ. Microbiol. 64, 4774–4781
15. Yernool, D. A., McCarthy, J. K., Eveleigh, D. E., and Bok, J. D. (2000) J. Bacteriol. 182, 5172–5179
16. Farinas, E. T., Bulter, T., and Arnold, F. H. (2001) Curr. Opin. Biotechnol. 12, 545–551
17. Zhao, H., Chockalingam, K., and Chen, Z. (2002) Curr. Opin. Biotechnol. 13, 104–110
18. Leung, D. C., and Goeddel, D. V. (1989) Technique J. Methods Cell Mol. Biol. 1, 11–15
19. Cadwell, R. C., and Joyce, G. F. (1992) PCR Methods Applications 2, 28–33
20. McCarthy, J. K., O’Brien, C. E., and Eveleigh, D. E. (2003) Anal. Biochem. 318, 196–203
21. Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Leducqua, J. L., and Polaina, J. (1998) J. Mol. Biol. 275, 491–502
22. Davies, G., and Henrissat, B. (1995) Structure 3, 853–859
23. Nagano, N., Porter, C. T., and Thornton, J. M. (2001) Protein Eng. 14, 845–855
24. Aguilar, C. F., Sanders, J., Moracci, M., Caramella, M., Nucci, R., Rossi, M., and Pearl, L. H. (1997) J. Mol. Biol. 271, 789–802
25. Hirmova, M., and Fischer, G. B. (1997) Carbohydr. Res. 305, 269–271
26. Ciceri, M., Blanchard, D., Bevan, D. R., and Essen, A. (2000) J. Biol. Chem. 275, 20002–20011
27. Tall, D., Withers, S. G., Gilkes, N. R., Kilburn, D. G., Warren, R. A., and

FIG. 7. Stereogram showing the GghA active site amino acid residues conserved in family 1 glycosyl hydrolases. The figure compares (A) wild type GghA to changes in (B) I170T. In the I170T model, His-119, Asn-163, and Trp-396 show significant shifts from their conserved locations. The figure was generated using Insight 2000.
Directed Evolution of *T. neapolitana* Glucan Glucohydrolase

Shafikhani, S., Siegel, R. A., Ferrari, E., and Schellenberger, V. (1997) *Bio-Techniques* **23**, 304–310

Hrmova, M., Macgregor, E. A., Biely, P., Stewart, R. J., and Fincher, G. B. (1998) *J. Biol. Chem.* **273**, 11134–11143

Tolley, S. P., Barrett, T. E., Suresh, C. G., and Hughes, M. A. (1993) *J. Mol. Biol.* **229**, 791–793

Jenkins, J., Lo Leggio, L., Harris, G., and Pickersgill, R. (1995) *FEBS Lett.* **362**, 281–285

Henrissat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.* **7**, 637–644

van Tilbeurgh, H., Loontiens, F. G., DeBruyne, C. K., and Classeyssens, M. (1988) *Methods Enzymol.* **160**, 45–59

Conn, E. E. (1992) in β-Glucosidases Biochemistry and Molecular Biology (Esen, A., ed) Vol. 533, pp. 15–41, ACS Symposium Series, American Chemical Society, Washington, D.C.

Czjzek, M., Cizek, M., Zamboni, V., Bevan, D. R., Henrissat, B., and Esen, A. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13555–13560

Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D. R., and Esen, A. (2003) *J. Biol. Chem.* **278**, 25055–25062

Zou, J., Klewegt, G. J., Stahlberg, J., Driguez, H., Nerinckx, W., Classeyssens, M., Kovula, A., Teeri, T. T., and Jones, T. A. (1999) *Struct. Fold. Des.* **7**, 1035–1045

Czjzek, M., Cizek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B., and Esen, A. (2001) *Biochem. J.* **354**, 37–46

Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S., and Henrissat, B. (1997) *Structure* **5**, 663–675
Improved Catalytic Efficiency and Active Site Modification of 1,4-β-d-Glucan Glucohydrolase A from *Thermotoga neapolitana* by Directed Evolution

James K. McCarthy, Aleksandra Uzelac, Diane F. Davis and Douglas E. Eveleigh

*J. Biol. Chem. 2004, 279:11495-11502.*

doi: 10.1074/jbc.M305642200 originally published online December 4, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M305642200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 53 references, 10 of which can be accessed free at http://www.jbc.org/content/279/12/11495.full.html#ref-list-1