THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 29, Issue of July 19, pp. 26501–26507, 2002

Printed in U.S.A.

A Novel, High Performance Enzyme for Starch Liquefaction

DISCOVERY AND OPTIMIZATION OF A LOW pH, THERMOSTABLE α-AMYLASE*

Received for publication, April 3, 2002, and in revised form, May 3, 2002
Published, JBC Papers in Press, May 6, 2002, DOI 10.1074/jbc.M203183200

Toby H. Richardson‡, Xuqiu Tan, Gerhard Frey, Walter Callen, Mark Cabell, David Lam, John Macomber, Jay M. Short, Dan E. Robertson, and Carl Miller

From the Diversa Corporation, San Diego, California 92121

High throughput screening of microbial DNA libraries was used to identify α-amylases with phenotypic characteristics compatible with large scale corn wet milling process conditions. Single and multiorganism DNA libraries originating from various environments were targeted for activity and sequence-based screening approaches. After initial screening, 15 clones were designated as primary hits based upon activity at pH 4.5 or 95 °C without addition of endogenous Ca^{2+}. After further characterization, three enzyme candidates were chosen each with an exceptional expression of one or more aspects of the necessary phenotype: temperature stability, pH optimum, lowered reliance on Ca^{2+} and/or enzyme rate. To combine the best aspects of the three phenotypes to optimize process compatibility, the natural gene homologues were used as a parental sequence set for gene reassembly. Approximately 21,000 chimeric daughter sequences were generated and subsets screened using a process-specific, high throughput activity assay. Gene reassembly resulted in numerous improved mutants with combined optimal phenotypes of expression, temperature stability, and pH optimum. After biochemical and process-specific characterization of these gene products, one α-amylase with exceptional process compatibility and economics was identified. This paper describes the synergistic approach of combining environmental discovery and laboratory evolution for identification and optimization of industrially important biocatalysts.

Endo-1,4-α-D-glucan glucohydrolase (α-amylase,[1] EC 3.2.1.1) is currently used in a broad array of industrial applications. These include starch hydrolysis for the production of ethanol and high fructose corn syrup, starch soil removal in laundry washing powders and dish-washing detergents, textile de-sizing, the production of modified starches, baking, hydrolysis of oil-field drilling fluids, and paper recycling. Since 1980, the most widely used enzyme for these applications has been the α-amylase isolated from the ubiquitous mesophilic soil bacterium *Bacillus licheniformis* (1–3). This enzyme operates optimally at 90 °C and pH 6, and it requires addition of calcium (Ca^{2+}) for its thermostability (4), conditions that are substantially different from those encountered in the various industrial processes where the enzyme is utilized. The disparity between these industrial requirements and the native environment for the α-amylase results in sub-optimal enzymatic performance in many applications.

Corn wet milling is an example of a multistep industrial process where there is considerable scope for enzyme performance improvement. Initially, whole corn kernels are fractionated into semi-purified streams of protein, fiber, oil, and starch. The resulting starch fraction has a pH of 4.5. The next process step involves liquefaction of the semi-purified starch to glucose oligomers by the *B. licheniformis* α-amylase, ideally at a pH of ~4.5 and a temperature of 105 °C. However, because the enzyme is unstable under these conditions (5), the pH must be increased to 5.7–6.0 and calcium added (5). The second step in the process involves saccharification of the liquefied product using a glucoamylase enzyme isolated from an *Aspergillus* sp. Because the glucoamylase activity optimum is pH 4.2–4.5 (5), the pH must be returned to pH 4.5 for this step to proceed efficiently. The final step in the process involves conversion of the saccharified glucose to high fructose corn syrup using glucose isomerase. Prior to this final step, the calcium added in the first step and the salts generated from the second step must be removed. Removal of these compounds is an expensive part of the overall process and could be avoided if the liquefaction step utilized an α-amylase enzyme capable of liquefying starch at pH 4.5 and 105 °C without the addition of calcium.

The study of hyperthermophilic bacteria and archaea, which grow optimally at temperatures >80 °C, has resulted in the discovery of many thermostable enzymes with industrial applications. These include a number of extracellular amylolytic enzymes from cultured microbes (6–15). A recent review of the sequence similarity of these enzymes (16) revealed that they are all members of the glycosyl hydrolase family 13 (17) and that they share a high degree of amino acid sequence identity (>76%), yet they display biochemical phenotypes that vary significantly in parameters such as T_{opt}, pH_{opt}, and metal ion requirements (*e.g.* Ca^{2+}).

This work describes the discovery of a group of thermostable α-amylase genes from nature and the subsequent laboratory evolution of a novel and improved chimeric α-amylase with performance characteristics ideal for the corn wet milling process. The approach taken utilizes access to biodiversity, ultra-high throughput screening, and GeneReassembly™, a proprietary DNA recombination technology.

EXPERIMENTAL PROCEDURES

Genomic DNA Libraries—More than 2000 genomic DNA libraries from DNA purified directly from environmental samples, isolates, and

* 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF504062, AF504063, AF504064, and AF504065.

‡ To whom correspondence should be addressed: Diversa Corporation, 4955 Directors Place, San Diego, CA 92121. Tel.: 858-526-5211; Fax: 858-526-5711; E-mail: trichardson@diversa.com.

[1] The abbreviations used are: α-amylase, endo-1,4-α-D-glucan glucohydrolase; HPLC, high pressure liquid chromatography; DE, dextrose equivalents; MOPS, 4-morpholinepropanesulfonic acid; RBB, Remazol Brilliant Blue; DGA, Diversa unit of glucoamylase.

primary enrichments have been constructed. Samples were collected under formal agreement with all legal parties. The method of generating these libraries has been described previously in United States patent number 5,958,672 (18). These libraries represent a vast array of ecological niches and biotopes that span the globe. A targeted subset of these libraries was chosen for α-amylase discovery based on the temperature and pH of the environmental sample used to generate the library. Examples of environments used to generate libraries in this study include a variety of different geographical deep-sea enrichments and acid soil environments.

α-Amylase Discovery, Expression Screening—Approximately 50,000 clones from selected λ DNA libraries were plated to produce plaques on semi-solid medium according to standard procedures (19). An isopropyl-β-D-thiogalactopyranoside-soaked (1 mM) nitrocellulose membrane was overlaid to both induce and replicate the clones (minimum of 5 min). After duplication, a substrate solution containing 0.1–1.0% azo-dyed starch (Sigma) in 50 mM Tris-HCl, pH 7.4, and 1% w/v agarose was heated until completely liquefied. The solution was cooled to around 50–70 °C, applied over the surface of the original plates, and allowed to cool to room temperature. Finally, the plates were heated to ~75 °C for 15–60 min in a sealed bag. Positive plaques, identified by the appearance of a clearing zone or "halo," were then purified.

Sequence-based Screening—A 500-bp PCR product corresponding to amino acid residues 30–197 derived from the α-amylase BD5031 was radioactively labeled and used to probe appropriate DNA libraries using standard methods (19).

Subcloning and Expression of Amylase Genes—All α-amylase genes were PCR-amplified using the α-amylase degenerate primers, without their native signal sequence, and cloned into a *Pseudomonas-Escherichia coli* shuttle vector using appropriate restriction enzymes and standard techniques (19). All molecular biology manipulations were performed in *E. coli*, and the resulting constructs were transformed into a proprietary *Pseudomonas fluorescens* strain. To express the various α-amylase constructs, they were grown to mid-log phase and induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside for 19 h at 30 °C. Cell growth rate was measured by absorbance at 600 nm using a Beckman DU-7400 spectrophotometer. The amount of protein expression was examined using SDS-PAGE with 10% acrylamide and a discontinuous buffer system. Gels were stained with Coomassie Blue, and the appropriate molecular weight region was examined for determination of α-amylase protein compared with the total protein.

Reassembled Library Construction—Nine fragments (*FI—FIX* in Fig. 1; each about 150 bp in length) were prepared from each of the three wild-type genes covering the whole open reading frame. Related DNA fragments were pooled (1:1:1), cut with the appropriate restriction enzymes, and gel-purified. Equal amounts of fragment pools were combined and ligated. The resulting ligation products were gel-purified and ligated to yield full-length reassembled α-amylase genes. Purified product was cut with the appropriate restriction enzymes, gel-purified, and ligated into an *E. coli/Pseudomonas* shuttle vector. The ligation mixture was transformed into *E. coli*. Plasmid DNA was isolated from pooled colonies (~21,000) and transformed into *P. fluorescens*.

Expression and α-Amylase Activity Assay—Reassembled libraries were transformed into *P. fluorescens*, and the transformants were sorted into 384-well plates at 1 cell per well. The plates were incubated for 24 h at 30 °C. Clones were replicated into a new 384-well plate for storage at 4 °C. An equal volume of 12 M urea as cell volume was added to each well, and the plates were incubated for 1 h at room temperature. For the assay, 5 μl of lysate was added to 75 μl of 1% RBB-insoluble cornstarch in 50 mM NaOAc buffer, pH 4.5, in a new 96-well plate. The plates were sealed and incubated at 90 °C for 20 min. After cooling to room temperature, 100 μl of ethanol was added to each well. The plates were then spun at 4000 rpm for 20 min using a tabletop centrifuge. The supernatant was transferred to a new plate, and the absorbance was measured at 595 nm using a Beckman DU-7400 spectrophotometer.

Recovery of Amylases—The α-amylases were recovered from 10-liter shake flask fermentations. Five grams of fermentation broth was mixed with 25 ml of deionized water in a 50-ml Falcon tube. The solution was vortexed and centrifuged at 13,000 rpm for 30 min (Sorvall, RC5C). The supernatant was removed, and the pellet was resuspended in 25 ml of fermentation broth with 50 mM HEPES, pH 7.0, containing 0.2 mM dithiothreitol. The samples were heated to 80 °C for 30 min and cooled to room temperature. The samples were re-centrifuged at 13,000 rpm for 30 min (Sorvall, RC5C). The clarified supernatant was collected and concentrated by Centriprep-30.

α-Amylase Secondary Characterization Assays—Amylase activity was determined in a continuous spectrophotometric assay by measuring the release of *p*-nitrophenol from the substrate, 5 mM *p*-nitrophenyl-α-D-hexa-(1,4)-glucopyranoside in 50 mM MOPS buffer, pH 7, at 80 °C. To calibrate the enzyme assay, 50 μl of enzyme solution (diluted such that the $\Delta A_{405/min}$ was <1) was added to 950 μl of the substrate solution in a pre-heated cuvette at 80 °C, and the increase in absorbance at 405 nm was measured over 1 min. One standard unit of thermostable α-amylase was defined as being equal to the amount of enzyme that will catalyze the release of 1 mmol/ml/min of *p*-nitrophenol under the defined conditions of the assay.

For test samples, 10 μl of amylase solution from a shake-flask culture broth was added to 300 μl of 2% RBB corn starch (Sigma catalog number S-7776) in 50 mM NaOAc buffer pH 4.5 in a 1.5-ml microcentrifuge tube. The mixture was incubated at 95 °C for 15 min and then cooled to room temperature. Ethanol (750 μl of 100%) was added to the mixture, and the tube was spun at 16,000 × *g* for 5 min. The supernatant was transferred to a 1-ml glass cuvette and read in a spectrophotometer at 595 nm.

Glucoamylase Activity Assay—Glucoamylase activity was determined in a continuous spectrophotometric assay by measuring the release of *p*-nitrophenol from the 5 mM substrate *p*-nitrophenyl-α-D-glucopyranoside in 50 mM MOPS buffer, pH 7, at 80 °C. To calibrate the enzyme assay, 50 ml of enzyme solution (diluted such that the $\Delta A_{405/min}$ was <1) was added to 950 ml of substrate solution in a pre-heated cuvette at 80 °C, and the increase in absorbance at 405 nm was measured over 1 min. One standard "Diversa unit" of glucoamylase (DGA) is equal to the amount of enzyme that will catalyze the release of 1 mmol/ml/min of *p*-nitrophenol under the defined conditions of the assay.

Laboratory Scale Starch Liquefaction—Laboratory scale liquefactions were performed in an incubator (Werner Mathis, Labomat). Cornstarch (Sigma) was thoroughly mixed with 100 mM NaOAc buffer in a ratio of 3:7 (w/w). α-Amylase was then added, mixed, and the pH adjusted to the desired value. The resulting mixture was heated for 5 min at 105 °C, followed by 90 min at 95 °C, and then cooled to 60 °C. An aliquot of this solution was diluted (1:200, w/v) in distilled water, and the pH was adjusted to ~2.0 by the addition of HCl. Liquefied starch solutions were sampled for dextrose equivalents determination.

Dextrose Equivalents (DE) Determination—The DE of liquefied syrups was determined by measuring the quantity of reducing sugars (as glucose) by the neocuproine procedure (20). The amount of glucose in the sample was determined by comparison to a known glucose standard (0.2 mg/ml). The percentage of glucose to the total carbohydrate (w/w) in the sample also known as DE is a measure of the starch hydrolysis.

Thermostability in the Presence and Absence of Calcium—The reassembled and parental α-amylases were analyzed for stability at pH 4.5 and 90 °C in 66 mM sodium acetate buffer. The samples were heated at 90 °C for 10 min and cooled on ice. Residual activity was assayed using of *p*-nitrophenol hexa-glucopyranoside (2.5 mM) in 200 mM MOPS buffer, pH 7. Activity pre- and postincubation was measured to calculate residual activity. An analogous experiment was conducted at pH 4.5 and a temperature of 100 °C in 66 mM NaOAc buffer in the presence of 40 ppm calcium. The samples were boiled for 10 min, cooled on ice, and assayed for residual activity as described above.

Pilot Plant Starch Liquefaction—Line starch directly from an industrial starch process was pumped into a 60-liter feed tank where pH, dry solids, and calcium level were adjusted prior to liquefaction. The α-amylase was added to the 32% dry solid slurry and pumped at 0.7 liters/min to a pressurized mixing chamber, where the starch slurry was instantaneously heated to 105 °C. The gelatinized, partially liquefied starch was pumped under pressure to give the desired dwell time (5 min) at 105 °C. The pressure was released into a flash tank, and samples were taken. The liquefied starch was collected in 1-liter glass bottles and held in a water bath at 95 °C for 90 min.

Saccharification—Liquefied starch was treated typically with 10 DGA units per kg of liquefied starch of glucoamylase at pH 4.5 and a temperature of 60 °C. Samples were recovered at various time points, and the extent of saccharification was monitored over time (2–88 h) by HPLC.

Oligosaccharide Profile—Liquefaction and saccharification carbohydrate profiles were measured by high pressure liquid chromatography (HPLC) using a Bio-Rad Aminex HPX-87C column in calcium form at 80 °C using refractive index detection.

RESULTS

Discovery of Amylases—Discovery and optimization of robust enzymes to catalyze specific chemistries in an industrial setting rely on the implementation of a number of technologies. Under agreements with countries around the world, samples

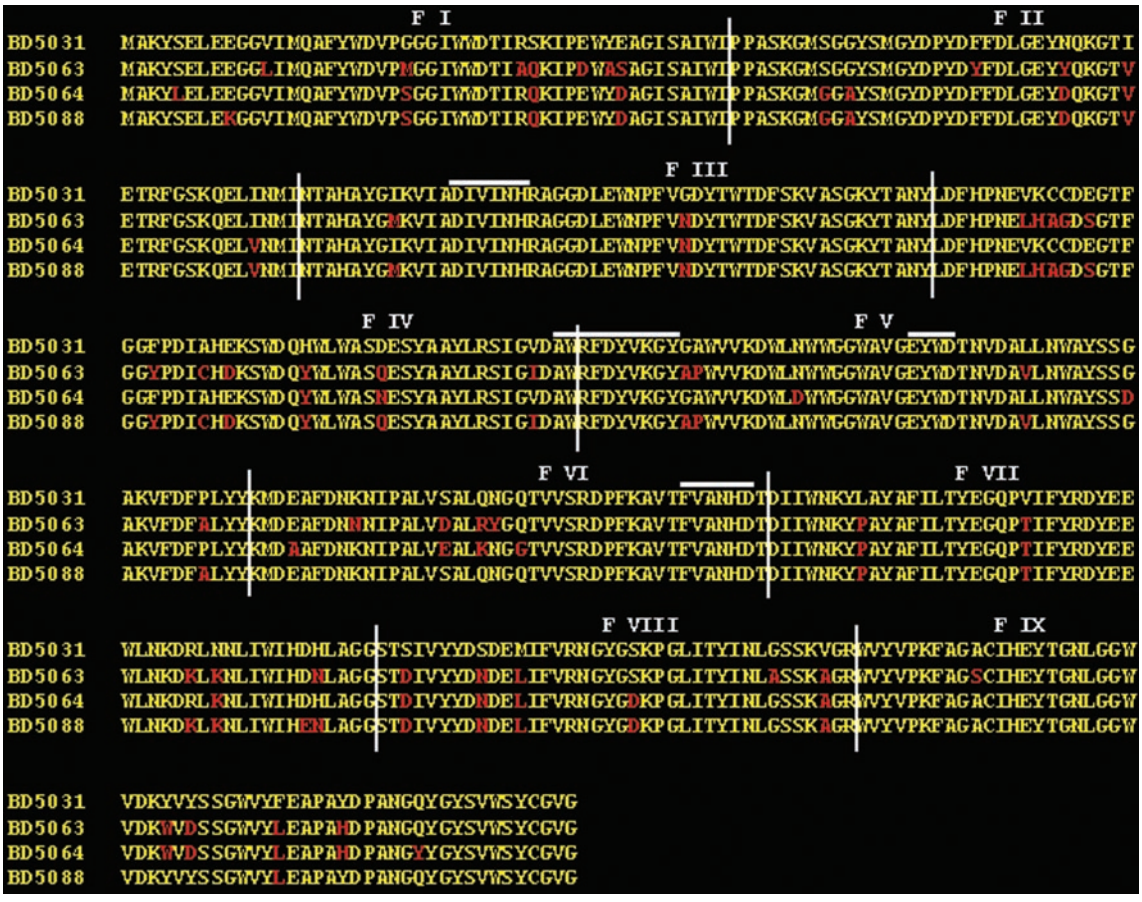

FIG. 1. **Gene reassembly.** Alignment of the three parents used for the reassembly and comparison to the best reassembled daughter, BD5088. Different amino acids (compared with BD5031) are shown in *red*. Reassembly was performed using 9 fragments (*FI–FIX*). *White vertical bars* indicate the junctions of these fragments. *White horizontal bars* indicate four highly conserved regions in α-amylases.

have been collected from a myriad of natural environments, many of which display environmental variables of extreme pH, temperature, pressure, organic contaminants, salt concentration, etc. The DNA representing the microbial constituents of the sample is processed into complex gene libraries. These libraries, made from DNA extracted from soils or aqueous environments, are screened for genes coding for enzymes of desired phenotypes using targeted expression assays and sequence homology-based methods (21).

To target discovery of new α-amylases applicable to the corn wet milling process, libraries containing the genomes of between 1 and 15,000 organisms were screened either by enzymatic activity or sequence homology to known α-amylases. These technologies enable screening of 10^5–10^9 clones/day from gene libraries using pH and temperature conditions approximating those of the corn wet milling process. α-Amylase expression screening was performed in semi-solid agar plates or in liquid-based format in microtiter plates using azo-dyed starch as the substrate (see "Experimental Procedures"). Clones with the highest activity on the azo-dyed starch were sequenced. Open reading frames were subjected to homology searches (BLAST) using the non-redundant gene data base at the National Center for Biotechnology Information.

Libraries also were screened using degenerate PCR primers; in brief these primers were designed to incorporate sequences found at the N and C termini of known hyperthermophilic α-amylases. By using these probes, single organism or environmental DNA libraries were screened for sequence homologues. Homologous DNA inserts were sequenced and re-cloned for expression, and their α-amylase activity was verified.

Following this concerted discovery effort, three enzymes, designated BD5031, BD5064, and BD5063 (internal classification labels) were selected. These enzymes each had near-optimal characteristics for the corn wet milling process. The sequences of these three proteins are shown in Fig. 1. BD5031 and BD5064 (both 461 amino acids in length) showed high levels of sequence identity to other hyperthermophilic α-amylases and were most similar (95 and 85% amino acid identity, respectively) to the α-amylase of *Pyrococcus* sp. KOD1 (12). BD5063 shared significant identity to both BD5031 and BD5063 (88 and 90% amino acid identity, respectively). Phylogenetic analysis revealed that all three enzymes were members of the glycosyl hydrolase family 13 (17) and are likely to be from organisms closely related to the order of Thermococcales.

Characterization of Wild-type α-Amylases—α-Amylases were expressed in a *P. fluorescens* host using a proprietary cloning vector. BD5031 was produced with and without its signal peptide sequence.

In order to simulate the first step of the corn wet milling process, heat-treated cell lysates were tested for their ability to liquefy starch under pH conditions ranging from 4.25 to 6.25 (Fig. 2). The pH optima of the three wild-type α-amylases, BD5031, BD5063, and BD5064, were compared with the commonly used *B. licheniformis* α-amylase. Each of these three newly discovered α-amylases had superior characteristics compared with the *Bacillus* enzyme, but BD5063 had the greatest activity at pH 4.5 and a temperature of 105 °C.

The relative level of α-amylase expression was found to be BD5031 > BD5064 > BD5063 (Table I). As with other heterologously expressed archaeal amylases reported previously (22), the majority of the α-amylase protein expressed in the *Pseudomonas* host was found as insoluble aggregates.

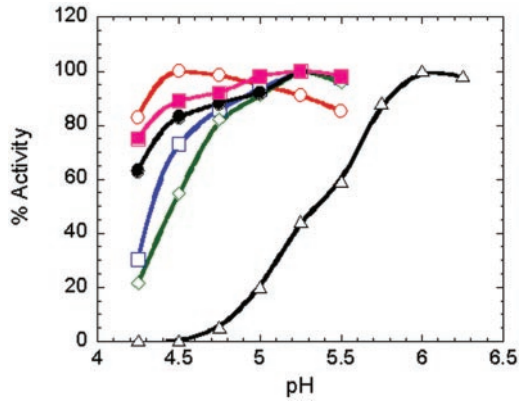

FIG. 2. **Determination of α-amylase pH optima.** The pH optima of wild-type and reassembled α-amylases were measured under typical industrial liquefaction conditions: 32% w/w starch slurry, 5 min treatment at 105 °C followed by 90 min at 95 °C. The dosage of amylase needed to achieve a target DE of 12 at the optimum for each α-amylase was determined. The same dosage was then used in liquefaction experiments at the remaining pH values, and the response was measured. The percentage of the maximal response at each pH is given as follows: △, *B. licheniformis* α-amylase; ◇, BD5031; □, BD5064; ●, BD5088; ○, BD5063; ■, BD5096.

TABLE I
Summary of the biochemical and growth characteristics of clones containing the wild-type α-amylases

Name	BD5031	BD5064	BD5063
Expression[a]	>30%	~20%	~5–10%
Relative activity[b]	15%	21%	100%
pH optimum	5.25	5.25	4.5
Growth in fermenter[c]	>200	~30	~80

[a] The levels of expression are the % of total protein expressed in a *Pseudomonas* host.
[b] The relative activity represents whole cell activity at a per cell level at pH 4.5 relative to BD5063.
[c] Final A_{575}.

After scaling to 10-liter fermenters, the pH and protein expression characteristics of all three enzymes were similar compared with those measured previously in smaller shake flasks (25 ml of media in a 250-ml flask). In the larger fermenter, BD5063 continued to maintain the best pH optimum (measured as liquefaction activity at the target pH) but was produced at much lower levels than the other two enzymes.

Reassembly of Amylases—In order to optimize the α-amylase productivity and expression phenotypes to better suit the commercial process, gene reassembly was performed using the three wild-type genes (Fig. 1) as parental sequence. Pooled fragments were ligated to produce chimeric α-amylase genes. Ligation of fragments from different parents with identical overhangs was predicted to produce all possible combinations with equal probability; the possible number of combinations of nine fragments from three different parents is 3^9 = 19,683, with only 3 of the 19,683 combinations regenerating the wild-type sequences.

The reassembled genes were ligated into an *E. coli*/*Pseudomonas* shuttle vector and introduced into *E. coli* by transformation. Plasmid DNA from ~21,000 pooled colonies was isolated and introduced into *P. fluorescens* also by transformation. Twenty five random clones from the primary reassembled library were sequenced to determine the distribution of the 27 parental fragments among the daughter clones and to verify the efficacy of the reassembly protocol (Table II). Each clone had a unique sequence and exhibited an apparently random distribution of all parental sequence fragments with the exception of fragment IV.

TABLE II
Analysis of fragments of reassembled α-amylase clones

25 random clones and 20 up-mutants were sequenced. The composition of the clones from the three different wild-type enzymes is shown (in percent). In fragments I–VII, one parent is preferred over the other two in the improved clones (≥50%; marked in gray). The derived consensus composition as well as the composition of two improved clones discussed in the text is shown at the bottom.

	Fragments								
Random Clones	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**
BD5063	36	48	32	44	28	56	40	16	28
BD5064	56	28	28	0	24	28	12	60	20
BD5031	8	24	40	56	48	16	48	24	52
Upmutants	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**
BD5063	50	25	30	80	50	20	70	45	35
BD5064	25	75	5	0	35	20	15	35	20
BD5031	25	0	65	20	15	60	15	20	45
Consensus	BD5063	BD5064	BD5031	BD5063	BD5063	BD5031	BD5063	BD5063	BD5063
Best upmutants									
BD5088	BD5064	BD5064	BD5063	BD5063	BD5063	BD5031	BD5063	BD5064	BD5031
BD5096	BD5063	BD5064	BD5063	BD5063	BD5063	BD5063	BD5063	BD5031	BD5064

Primary, Secondary, and Tertiary Screening of Reassembled α-Amylases—To identify enhanced α-amylases from the pool of reassembled clones, a screening strategy was designed to identify clones that exhibited high activity at pH 4.5 based upon normalized cell density. The protocol for the primary high throughput assay is shown schematically in Fig. 3. In brief, the reassembled clones were individually distributed into 384-well plates at an average inoculum of 1 cell per well and incubated for 24 h. The cells were then lysed with 6 M urea for 1 h to solubilize all α-amylase protein. An aliquot of cell lysate was transferred to a well in the screening plate containing Remazol Brilliant Blue (RBB)-labeled cornstarch in 50 mM NaOAc buffer at pH 4.5. The assay plates were incubated at 90 °C for 20 min. Following an ethanol precipitation step, enzymatic activity (release of RBB) was assessed by measuring absorbance at 595 nm.

Detectable amylase activity was found in ~40% of the reassembled clones. 145 daughter clones showed an increased activity on a per cell basis at pH 4.5, when compared with the three parental α-amylases in the primary screen. Seventy one of the best up-mutants were grown in shake flasks and subjected to a secondary screen for starch liquefaction. From the liquefaction screen a total of 34 clones was confirmed to have higher liquefaction activity at pH 4.5 when compared with the three wild-type enzymes on a per cell basis (data not shown). The α-amylase activity in each of these 34 clones was also tested for improved thermostability at pH 4.5.

Twenty two of the 34 reassembled enzymes exhibited improved thermostability at 90 °C in the absence of added calcium when compared with BD5063, the most stable of the wild-type parental α-amylases. Similarly, 20 reassembled clones exhibited greater thermostability at 100 °C in the presence of 40 mg/liter calcium when compared with the wild-type α-amylase BD5063 (data not shown).

Twenty seven of the 34 clones described above were further evaluated for their capacity to express α-amylase activity during large scale fermentation. Two of the reassembled clones, BD5088 and BD5096, were robust and grew to high cell density (>200 A_{575}). By using these two clones, the α-am-

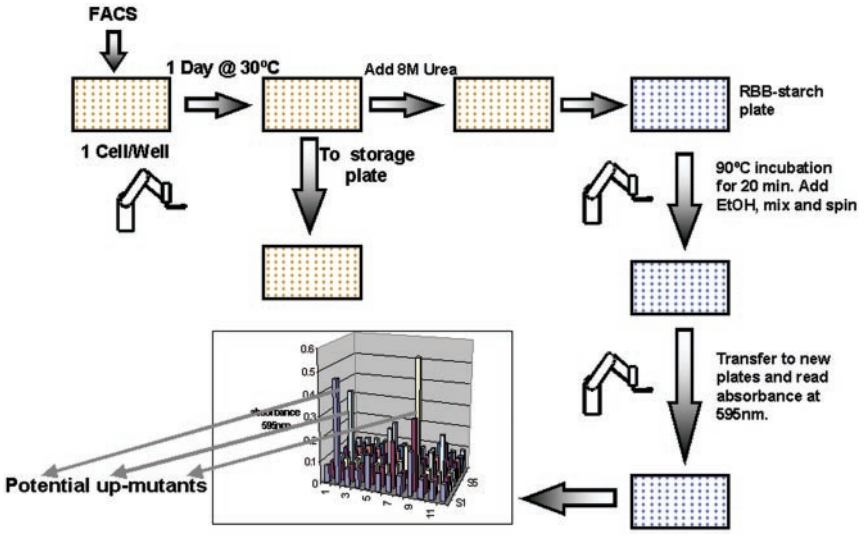

FIG. 3. **A schematic representation of a high throughput α-amylase screen.** The library of reassembled clones was sorted into 96- or 384-well plates by fluorescence-activated cell sorting (*FACS*). Recombinant proteins were solubilized with 6 M urea and assayed on RBB starch plates (see "Experimental Procedures"). Putative up-mutants were identified by an increase in A_{595}.

ylase activity per liter of fermentation broth, as judged by the liquefaction assay, was increased 2–4-fold over the best wild-type α-amylase BD5031. The relative thermostability of the α-amylases BD5088 and BD5096 produced by reassembly, compared with the three wild-type α-amylases (BD5031, BD5064, and BD5063) and the *B. licheniformis* α-amylase, is shown in Fig. 4. The half-lives of the chimeric α-amylases, at 90 °C and pH 4.5, and in the absence of calcium ions, were improved 40-fold when compared with the most stable wild-type α-amylase, BD5031. Calculation of improvement over *B. licheniformis* α-amylase is complicated by its immediate denaturation under the assay conditions. The pH profiles of α-amylases produced by clones BD5088 and BD5096 were similar to the wild-type α-amylase produced by BD5063, in the pH range 4.25–4.75 (Fig. 2).

The fragment composition of the identified up-mutants showed some remarkable differences compared with the clones that had been randomly sequenced (Table II). For fragments I, IV, V, and VII, BD5063 is the most frequent parental sequence contributor, whereas BD5064 dominates for fragment II. Fragment II of BD5031 is not contained in any of the sequenced up-mutants, whereas BD5031 is the most frequent contributor in positions III and VI.

Confirmation of Improved Performance in an Industrial Corn Wet Milling Facility—Because of its optimal combination of temperature stability, pH profile, expression characteristics, and lack of stringent calcium ion requirement, the reassembled daughter gene product BD5088 was chosen as the primary candidate enzyme for testing in an industrial corn wet milling facility. In order to confirm that the enhanced thermostability along with good growth characteristics would lead to improved performance in the actual corn wet milling liquefaction process, the evolved α-amylase BD5088 was compared on a gram loading basis with the parental α-amylase BD5063 under laboratory scale conditions (data not shown). Based upon equivalent liquefaction performance at pH 4.5 and 105 °C, the enhanced thermostability of BD5088 was reflected by a greater degree of starch hydrolysis at 115 °C. To validate the liquefaction performance under industrial conditions, the reassembled BD5088 α-amylase was employed in a continuous 2 gallon per min pilot reactor to liquefy cornstarch at pH 4.5.

In the industrial scale starch liquefaction process, a 32% w/w starch slurry was pumped through a narrow orifice, with steam injected to immediately raise the temperature to 105–107 °C. The α-amylase was continuously added into the starch slurry prior to contact with the steam. The gelatinized, partially hydrolyzed starch was kept under pressure for 5 min, released into a flash tank, and kept at 95 °C for an additional 90 min. Before proceeding with the next step in the process (glucoamy-

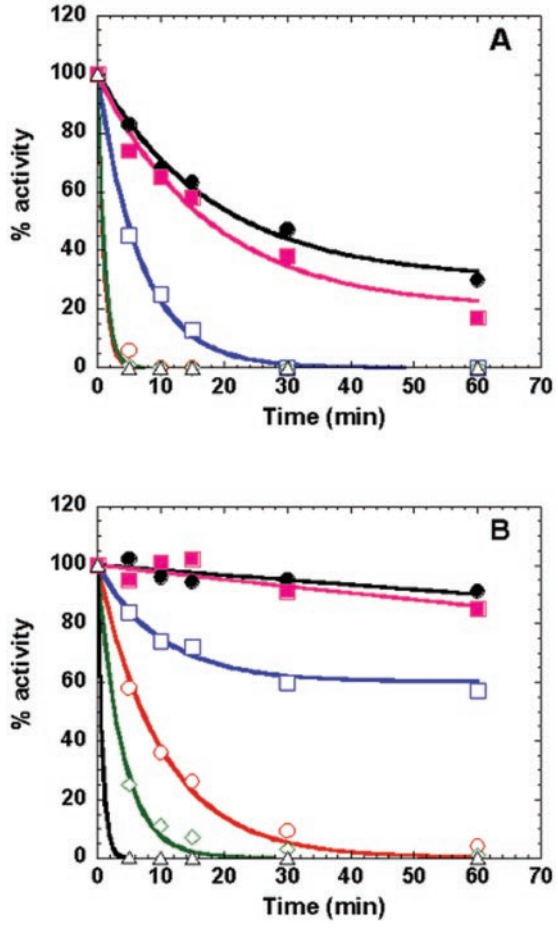

FIG. 4. **A comparison of wild-type and selected reassembled α-amylases with the *B. licheniformis* α-amylase.** The amylases were incubated in acetate buffer at 100 °C at a protein concentration between 0.3 and 0.4 mg/ml. Time points were taken at 0, 5, 10, 15, 30, and 60 min. The activity was measured according to the of *p*-nitrophenol (*pNP*) hexa-glucopyranoside assay (see "Experimental Procedures"). The residual activity was plotted as the percentage of initial activity. *A*, pH 4.5 in the absence of calcium; *B*, pH 4.5 with 40 mg/liter calcium. △, *B. licheniformis* α-amylase; ◇, BD5063; □, BD5031; ●, BD5088; ○, BD5064; ■, BD5096.

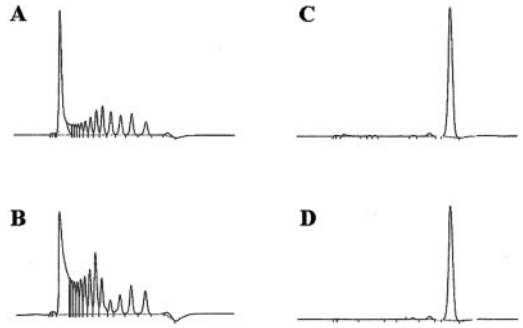

FIG. 5. **Oligosaccharide profile of BD5088 and *B. lichenformis* α-amylases, following liquefaction and saccharification of starch.** Liquefied starch syrups (18 DE) were prepared from BD5088 at pH 4.5 and from the *B. licheniformis* α-amylase at pH 6.0 using the 2 gpm pilot liquefaction system unit. *A*, *B. licheniformis*; *B*, BD5088. These liquefied syrups were then saccharified at pH 4.5 and 60 °C with the same dosage of a commercial glucoamylase; *C*, *B. licheniformis*; *D*, BD5088. The saccharide profile was measured by HPLC (see "Experimental Procedures").

lase-catalyzed saccharification), the degree of hydrolysis and concentration of non-hydrolyzed starch was measured.

The performance of BD5088 in the pilot scale process, using starch from a commercial corn wet milling facility, was similar to the performance under laboratory conditions. The oligosaccharide HPLC profile of the BD5088-liquefied starch was compared with the profile generated by the α-amylase from *B. licheniformis* (Fig. 5, *A* and *B*) and showed a very similar degradation pattern with slightly higher molecular weight oligomers formed by BD5088. The BD5088-liquefied starch was saccharified with a commercial *Aspergillus* glucoamylase to produce a glucose syrup containing the same level of glucose as a syrup produced with a current commercial enzyme process (Fig. 5, *C* and *D*).

DISCUSSION

The ability to access and rapidly screen discrete biotopes to discover novel α-amylases has proven to be extremely valuable in this study. Although the naturally occurring α-amylases (BD5031, BD5063, and BD5064) described in this work were similar to each other at the amino acid level, the environmental conditions, *i.e.* high temperature and low pH, prevailing at the sites of sample collection apparently has had a strong influence on their biochemical properties. Differences in biochemical characteristics between hyperthermophilic α-amylases have been observed previously (16) and have been re-iterated by Holden *et al.* (23). In this paper eight *Thermococcus* strains, from three geographically distinct environments, were compared. Although there was little difference in their 16 S rRNA signatures, large phenotypic differences were observed with respect to their α-amylase phenotypes.

In the present study, the results confirm the importance of biodiversity and reinforce the value of searching different environments to capture enzymatic diversity. Although significantly different enzyme properties were displayed by a collection of wild-type α-amylases from different environments, no single enzyme possessed all the desired properties for the targeted industrial application. α-Amylase BD5031 was expressed at high levels in shake flasks and in a fermenter, but its performance was not optimal at the target liquefaction pH of 4.5; α-amylase BD5063 was optimally active at pH 4.5 but was only expressed at low levels in a fermenter; α-amylase BD5064 hydrolyzed starch rapidly but exhibited lower thermostability.

We and others (24–28) have demonstrated that directed evolution is an efficient tool for the improvement of enzymes. A number of methods for generating chimeric genes from natural gene families have been reported, each utilizing a unique process and each resulting in novel daughter genes. Most of the methods used for DNA shuffling rely on the extension of cross-annealed random DNA fragments from different parents (29–32). However, because the number and location of crossover events is dependent on the relative parental sequence identity, the sites of crossover are not easily controlled. These methodologies used for gene reassembly result in the regeneration of a large proportion of full-length parental sequences, thereby diminishing the structural diversity of the chimeric library. The presence of these parental genes can dramatically increase the number of clones that need to be screened to obtain coverage of the library and identify useful variants. Libraries of single crossover chimeras, independent of DNA sequence identity, can be constructed using procedures described recently (33, 34), but the effectiveness of these methods remains to be shown.

Ligation methods are also effective for generating gene chimeras. Building reassembled genes from defined DNA fragments allows for precise control over the number and locations of the crossover sites. If necessary, genes can be recombined at non-conserved sites. DNA fragments with compatible overhangs are joined with the same efficiency, independent of their origin and sequence identity. Each possible combination of fragments (including wild-type parent sequences) is formed with the same probability. The complexity of the reassembled library can be tuned to suit the specific screening strategy. A higher complexity can be achieved by using more parent genes or by increasing the number of fragments. The fragments can be generated by synthesis, enzymatic cleavage, or PCR. The use of synthetically prepared oligonucleotides affords the opportunity to optimize codon usage and to add/eliminate specific restriction sites during the reassembly. Because the reassembly sites are known precisely, it is also possible to include knowledge of structural elements in the design of a GeneReassembly™ procedure.

A multistep, high throughput screen was developed to identify the amylase that best incorporated the targeted phenotypes from over 19,000 possible candidates after reassembly of the three wild-type α-amylases. The initial selection was based upon the highest response in a microtiter plate high throughput activity screen using RBB-insoluble cornstarch as substrate. Although the RBB assay was not the screen with the highest throughput, it was selected because the results obtained correlated well with liquefaction assay data. This primary screen identified clones with enhanced specific activity at pH 4.5 and/or active clones exhibiting a higher expression level. The best candidates derived from reassembly were grown and expressed in shake flasks and were subjected to a laboratory scale starch liquefaction assay based upon an equal number of cells from a shake flask culture. As in the primary screen, the best clones were those with either improved specific activity or showing a higher expression level. α-Amylases from the best performing clones found during the secondary screen were dosed into a laboratory scale starch liquefaction test based upon equal protein level. These reassembled α-amylases were also evaluated for thermostability at pH 4.5 and 100 °C and were compared with the wild-type enzymes.

The final selection of the best reassembled α-amylase was based upon application-specific activity at pH 4.5, shake flask expression level, and fermentation titer. The selection process was validated by a pilot scale liquefaction trial at an industrial corn wet miller. The identified up-mutants showed a trend with respect to their parent composition in 7/9 fragments used for the gene reassembly procedure. Fragments VIII and IX showed an approximately equal distribution of sequence from all three parents. The derived consensus as well as the composition of

the best up-mutants showed that all three parents contributed to generating enzymes with the desired phenotype.

Results reported here provide further support for the combined application of discovery and laboratory evolution tools for fully exploring biodiversity and enzyme sequence space. These modern molecular biology techniques have been demonstrated to yield a multitude of robust, natural enzymes well suited to industrial applications or to laboratory strategies of evolutionary improvement.

Acknowledgments—We thank Ronald Swanson, Eric Mathur, and Kevin Gray for contributions to this work and Keith Kretz, Mervyn Bibb, and Geoff Hazlewood for critical comments regarding the manuscript.

REFERENCES

1. Saito, N. (1973) *Arch. Biochem. Biophys.* **155,** 290–298
2. Yuuki, T., Nomura, T., Tezuka, H., Tsuboi, A., Yamagata, H., Tsukagoshi, N., and Udaka, S. (1985) *J. Biochem.* (*Tokyo*) **98,** 1147–1156
3. Matsuzaki, H., Yamane, K., Yamaguchi, K., Nagata, Y., and Maruo, B. (1974) *Biochim. Biophys. Acta* **365,** 235–247
4. Violet, M., and Meunier, J. C. (1989) *Biochem. J.* **263,** 665–670
5. Hashida, M., and Bisgaard-Frantzen, H., (2000) *Trends Glycosci. Glycotechnol.* **12,** 389–401
6. Frillingos, S., Linden, A., Niehaus, F., Vargas, C., Nieto, J. J., Ventosa, A., Antranikian, G., and Drainas, C. (2000) *J. Appl. Microbiol.* **88,** 495–503
7. Dong, G., Vieille, C., Savchenko, A., and Zeikus, J. G. (1997) *Appl. Environ. Microbiol.* **63,** 3569–3576
8. Jorgensen, S., Vorgias, C. E., and Antranikian, G. (1997) *J. Biol. Chem.* **272,** 16335–16342
9. Leveque, E., Haye, B., and Belarbi, A. (2000) *FEMS Microbiol. Lett.* **186,** 67–71
10. Imanaka, T., Tachibana, Y., Suzuki, Y., Kojima, I., and Utsura, K. (1996) JP Patent 96191138(09173077), 12
11. Lee, Jong T., Kanai, H., Kobayashi, T., Akiba, T., and Kudo, T. (1996) *J. Ferment. Bioeng.* **82,** 432–438
12. Tachibana, Y., Leclere, M., Fujiwara, S., Takagi, M., and Imanaka, T. (1996) *J. Ferment. Bioeng.* **82,** 224–232
13. Koch, R., Spreinat, A., Lemke, K., and Antranikian, G. (1991) *Arch. Microbiol.* **155,** 572–578
14. Brown, S. H., Costantino, H. R., and Kelly, R. M. (1990) *Appl. Environ. Microbiol.* **56**, 1985–1991
15. Chung, Y. C., Kobayashi, T., Kanai, H., Akiba, T., and Kudo, T. (1995) *Appl. Environ. Microbiol.* **61,** 1502–1506
16. Leveque, E., Janecek, S., Haye, B., and Belarbi, A. (2000) *Enzyme Microb. Technol.* **26,** 3–14
17. Henrissat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.* **7,** 637–644
18. Short, J. M. (June 3, 1996) U. S. Patent 96-657409(5958672), 22
19. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning*: *A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
20. Dygert, S., Li, L. H., Florida, D., and Thoma, J. A. (1965) *Anal. Biochem.* **13,** 367–374
21. Robertson, D., Mathur, E., Swanson, R. V., Marrs, B. L., and Short, J. (1996) *SIM News* **46,** 3–8
22. Linden, A., Niehaus, F., and Antranikian, G. (2000) *J. Chromatogr. B Biomed. Appl.* **737,** 253–259
23. Holden, J. F., Takai, K., Summit, M., Bolton, S., Zyskowski, J., and Baross, J. A. (2001) *FEMS Microbiol. Ecol.* **36,** 51–60
24. Arnold, F. H., and Volkov, A. A. (1999) *Curr. Opin. Chem. Biol.* **3,** 54–59
25. Giver, L., Gershenson, A., Freskgard, P. O., and Arnold, F. H. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 12809–12813
26. Ness, J. E., Welch, M., Giver, L., Bueno, M., Cherry, J. R., Borchert, T. V., Stemmer, W. P., and Minshull, J. (1999) *Nat. Biotechnol.* **17,** 893–896
27. Schmidt-Dannert, C., and Arnold, F. H. (1999) *Trends Biotechnol.* **17,** 135–136
28. Coco, W. M., Levinson, W. E., Crist, M. J., Hektor, H. J., Darzins, A., Pienkos, P. T., Squires, C. H., and Monticello, D. J. (2001) *Nat. Biotechnol.* **19,** 354–359
29. Stemmer, W. P. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 10747–10751
30. Stemmer, W. P. (1994) *Nature* **370,** 389–391
31. Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. (1998) *Nat. Biotechnol.* **16,** 258–261
32. Crameri, A., Raillard, S. A., Bermudez, E., and Stemmer, W. P. C. (1998) *Nature* **391,** 288–291
33. Ostermeier, M., Jae Hoon, S., and Benkovic, S. J. (1999) *Nat. Biotechnol.* **17,** 1205–1209
34. Sieber, V., Martinez, C. A., and Arnold, F. H. (2001) *Nat. Biotechnol.* **19,** 456–460