

Video Article

# GENPLAT: an Automated Platform for Biomass Enzyme Discovery and Cocktail Optimization

Jonathan Walton<sup>1,2</sup>, Goutami Banerjee<sup>2</sup>, Suzana Car<sup>2</sup>

<sup>1</sup>DOE Plant Research Laboratory, Michigan State University

<sup>2</sup>DOE Great Lakes Bioenergy Research Center, Michigan State University

Correspondence to: Jonathan Walton at [walton@msu.edu](mailto:walton@msu.edu)

URL: <https://www.jove.com/video/3314>

DOI: [doi:10.3791/3314](https://doi.org/10.3791/3314)

Keywords: Bioengineering, Issue 56, cellulase, cellobiohydrolase, glucanase, xylanase, hemicellulase, experimental design, biomass, bioenergy, corn stover, glycosyl hydrolase

Date Published: 10/24/2011

Citation: Walton, J., Banerjee, G., Car, S. GENPLAT: an Automated Platform for Biomass Enzyme Discovery and Cocktail Optimization. *J. Vis. Exp.* (56), e3314, doi:10.3791/3314 (2011).

## Abstract

The high cost of enzymes for biomass deconstruction is a major impediment to the economic conversion of lignocellulosic feedstocks to liquid transportation fuels such as ethanol. We have developed an integrated high throughput platform, called GENPLAT, for the discovery and development of novel enzymes and enzyme cocktails for the release of sugars from diverse pretreatment/biomass combinations. GENPLAT comprises four elements: individual pure enzymes, statistical design of experiments, robotic pipeting of biomass slurries and enzymes, and automated colorimetric determination of released Glc and Xyl. Individual enzymes are produced by expression in *Pichia pastoris* or *Trichoderma reesei*, or by chromatographic purification from commercial cocktails or from extracts of novel microorganisms. Simplex lattice (fractional factorial) mixture models are designed using commercial Design of Experiment statistical software. Enzyme mixtures of high complexity are constructed using robotic pipeting into a 96-well format. The measurement of released Glc and Xyl is automated using enzyme-linked colorimetric assays. Optimized enzyme mixtures containing as many as 16 components have been tested on a variety of feedstock and pretreatment combinations.

GENPLAT is adaptable to mixtures of pure enzymes, mixtures of commercial products (e.g., Accellerase 1000 and Novozyme 188), extracts of novel microbes, or combinations thereof. To make and test mixtures of ~10 pure enzymes requires less than 100 µg of each protein and fewer than 100 total reactions, when operated at a final total loading of 15 mg protein/g glucan. We use enzymes from several sources. Enzymes can be purified from natural sources such as fungal cultures (e.g., *Aspergillus niger*, *Cochliobolus carbonum*, and *Galerina marginata*), or they can be made by expression of the encoding genes (obtained from the increasing number of microbial genome sequences) in hosts such as *E. coli*, *Pichia pastoris*, or a filamentous fungus such as *T. reesei*. Proteins can also be purified from commercial enzyme cocktails (e.g., Multifect Xylanase, Novozyme 188). An increasing number of pure enzymes, including glycosyl hydrolases, cell wall-active esterases, proteases, and lyases, are available from commercial sources, e.g., Megazyme, Inc. ([www.megazyme.com](http://www.megazyme.com)), NZYTech ([www.nzytech.com](http://www.nzytech.com)), and PROZOMIX ([www.prozomix.com](http://www.prozomix.com)).

Design-Expert software (Stat-Ease, Inc.) is used to create simplex-lattice designs and to analyze responses (in this case, Glc and Xyl release). Mixtures contain 4-20 components, which can vary in proportion between 0 and 100%. Assay points typically include the extreme vertices with a sufficient number of intervening points to generate a valid model. In the terminology of experimental design, most of our studies are "mixture" experiments, meaning that the sum of all components adds to a total fixed protein loading (expressed as mg/g glucan). The number of mixtures in the simplex-lattice depends on both the number of components in the mixture and the degree of polynomial (quadratic or cubic). For example, a 6-component experiment will entail 63 separate reactions with an augmented special cubic model, which can detect three-way interactions, whereas only 23 individual reactions are necessary with an augmented quadratic model. For mixtures containing more than eight components, a quadratic experimental design is more practical, and in our experience such models are usually statistically valid.

All enzyme loadings are expressed as a percentage of the final total loading (which for our experiments is typically 15 mg protein/g glucan). For "core" enzymes, the lower percentage limit is set to 5%. This limit was derived from our experience in which yields of Glc and/or Xyl were very low if any core enzyme was present at 0%. Poor models result from too many samples showing very low Glc or Xyl yields. Setting a lower limit in turn determines an upper limit. That is, for a six-component experiment, if the lower limit for each single component is set to 5%, then the upper limit of each single component will be 75%. The lower limits of all other enzymes considered as "accessory" are set to 0%. "Core" and "accessory" are somewhat arbitrary designations and will differ depending on the substrate, but in our studies the core enzymes for release of Glc from corn stover comprise the following enzymes from *T. reesei*: CBH1 (also known as Cel7A), CBH2 (Cel6A), EG1(Cel7B), BG (β-glucosidase), EX3 (endo-β1,4-xylanase, GH10), and BX (β-xylosidase).

## Video Link

The video component of this article can be found at <https://www.jove.com/video/3314/>

## Protocol

### 1. Enzyme Production

1. Produce proteins in *Pichia pastoris* by cloning of the corresponding genes into vectors pPICZ or pPICZ $\alpha$  and transformation into *P. pastoris*. Grow cells in 300-ml batches in baffled 500-ml flasks with methanol induction every 24 hr (Banerjee et al., 2010a).
2. Concentrate and desalt the culture filtrates by tangential flow filtration.
3. Store the enzymes in small aliquots in 20% glycerol at -80°C. The concentration range is 1-10 mg/ml.

### 2. Design of Experiment

1. Input the number of enzymes to be tested, their upper and lower proportions (e.g. a lower proportion of 5% for core enzymes), and the type of experimental design (e.g., quadratic or cubic) into the Design-Expert software.
2. Calculate the amount in  $\mu$ g of each enzyme to be added to make a total loading of 15 mg/g glucan, and then calculate the amount of each enzyme in  $\mu$ l to add to each reaction.
3. Generate Excel worksheets specifying source labware, source wells, destination labware, destination wells, and transfer volumes for dispensing of water and enzymes according to the experimental design and import it into the Biomek FXP software using the transfer-from-file function.

### 3. Enzymatic Hydrolysis

1. Suspend the biomass slurry in 50 mM sodium citrate, pH 4.8, containing 5  $\mu$ g/ml tetracycline and 5  $\mu$ g/ml cycloheximide, in the paddle reservoir.
2. Robotically pipet 200  $\mu$ l of the slurry into each well of a 96-well deep-well reaction plates using Span-8 1000- $\mu$ l pipet tips with the last 0.5 cm cut off. Mix the slurry once at 100  $\mu$ l/sec and then aspirate the same volume from the paddle reservoir and dispense into the plates.
3. Dispense the enzymes into the same plate using the program entered into the Beckman FX.
4. Seal the plates with pierceable capmats.
5. Invert and tap the plates several times to overcome surface tension. Place the plates at 50°C for 48 hr in the hybridization incubator rotating at 10 rpm.

### 4. Determination of Glc and Xyl

1. Centrifuge the reaction plates for 3 min at 1250 x g in a swinging bucket centrifuge.
2. Transfer 100  $\mu$ l of the supernatant into regular 96-well plates using the AP96 pod of the Biomek FX.
3. Incubate the plates at 90-95°C for 10 min in order to inactivate the enzymes, and then centrifuge at 1250 x g for 30 sec.
4. For Glc measurements, transfer 12  $\mu$ l of the supernatants into regular 96-well plates. Add 192  $\mu$ l of the glucose oxidase/peroxidase (GPOD) reagent. Incubate the plates at 50°C for 20 min and read the absorbances at 510 nm in a microplate reader. Blank is reaction mixture with no enzyme.
5. For Xyl measurements, transfer 4  $\mu$ l into 384-well plates. Add the assay reagents and read the plates at 340 nm. Blank is the reaction mixture with no enzyme.

### 5. Data Analysis

1. Convert the absorbance values to their % Glc and Xyl yields based on the known Glc and Xyl content of the original biomass. Import these percentages as the responses into the Design-Expert software. Check the statistical parameters to be sure that the criteria for a robust model are fulfilled.
2. Confirm the best model prediction experimentally.

### 6. Representative Results with GENPLAT:

**(1) Creation of optimized mixtures of individual pure proteins.** The results indicate which enzymes are important, and in what proportions, and also which enzymes are *not* important. Some proteins that are abundant in the *T. reesei* secretome (Nagendran et al., 2009), such as Cip1 and Cip2, appear to play no role in Glc release from corn stover pretreated by ammonia fiber expansion (AFEX) or alkaline hydrogen peroxide (AHP) (Banerjee et al., 2010b). On the other hand, some proteins are unexpectedly important. For example, endo-xylanases of Glycosyl Hydrolase families 10 and 11 are both important for Glc release; one xylanase cannot substitute for the other (Banerjee et al., 2010b,c). Cel61A, a minor protein in the *T. reesei* secretome, is very important for Glc but not Xyl release. Some enzymes are important for release of both Glc and Xyl, whereas others are necessary only for one or the other (Banerjee et al., 2010c).

A synthetic mixture containing 16 components, each at a concentration of 0.94 mg/ml (i.e., a non-optimized mixture), released 38% of the available Glc from AFEX-pretreated DDG. Under identical hydrolysis conditions, an optimized mixture, in which the concentrations of the 16 components ranged from 0% to 32%, released 52% of the Glc (Fig. 2). This experiment illustrates the utility of constructing defined mixtures.

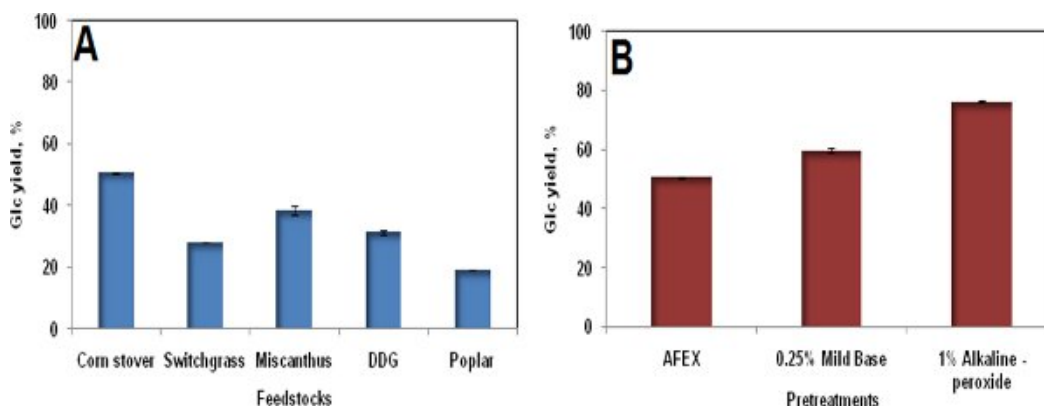
**(2) Creation of optimized cocktails for different pretreatment/substrate combinations.** Current commercial cellulase preparations are "one-size-fits-all". In reality, the best cocktail depends on both the substrate and the pretreatment. A single product such as Accellerase 1000 gives

different yields from different substrates pretreated in the same way, and from the same substrate exposed to different pretreatments (Fig. 1). We have used GENPLAT to optimize mixtures of pure enzymes for multiple pretreatments and multiple feedstocks (Banerjee et al., 2010c). Figure 2 shows how the optimal enzyme proportions of 16 pure enzymes differs for AFEX-pretreated corn stover, AHP-pretreated corn stover, and AFEX-pretreated DDG. Only 11 enzymes are shown in Fig. 2 because the optimal proportions of the other five (Cel61B, AbfB, Cip1, Cip2, and Cel12A) were found to be 0% (Banerjee et al., 2010c).

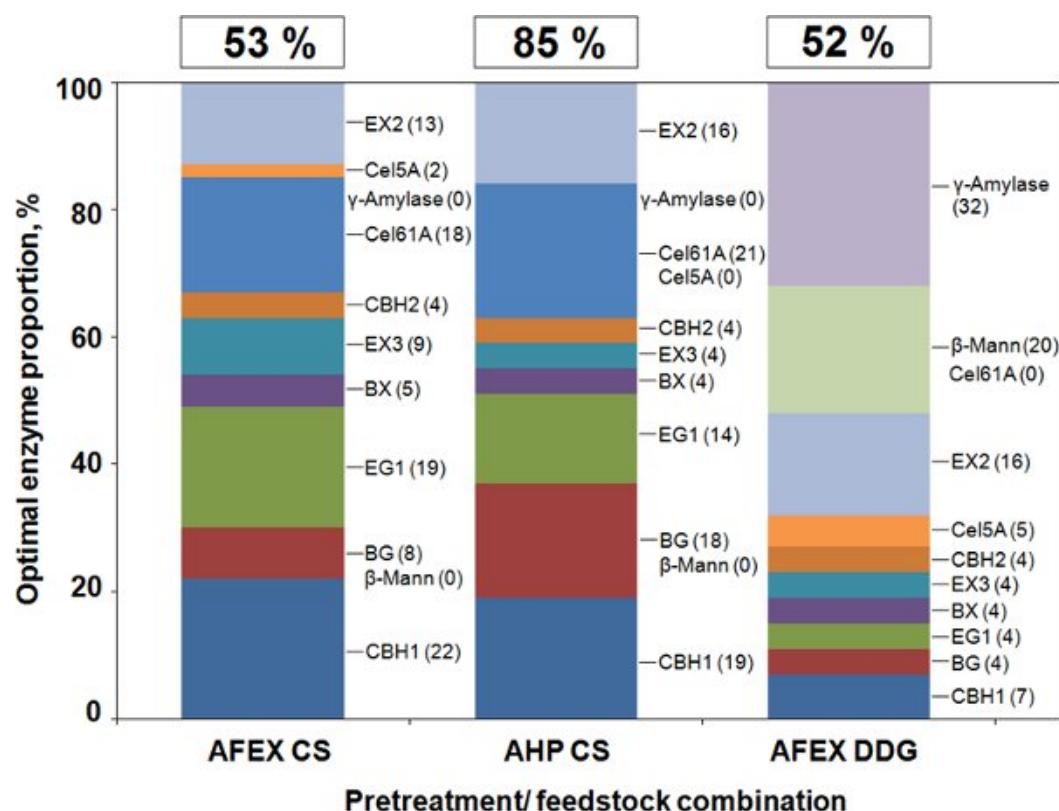
**(3) Novel enzyme discovery.** The construction of optimized mixtures of pure proteins is an ongoing process. As new proteins become available in pure form (from commercial sources, by heterologous expression, or by purification), they can be systematically tested by adding them to the current optimized set. Novel proteins that contribute to Glc or Xyl release then become part of a new, optimized set. In this way, GENPLAT becomes a platform for the identification of new enzymes for biomass deconstruction. We have made extracts of fungi other than *T. reesei* and *A. niger* grown on different substrates. One particular extract boosted Glc yield when added to our core set. The responsible protein was purified and shown to be a novel glycosyl hydrolase not present in any commercial enzyme preparation (Banerjee and Walton, unpublished results).

**(4) Discovery of better enzymes.** Having shown with GENPLAT which enzymes are most important for degradation of a particular biomass, we have a better understanding of which enzymes should be the focus for improvement. Better examples of the critical enzymes could be found by searching in nature or made by protein engineering. (An enzyme could be "better" than its *T. reesei* homolog in several ways, depending on the desired property, e.g., higher specific activity, greater synergy, decreased proteolytic sensitivity, or enhanced thermal stability). GENPLAT provides a meaningful platform for assaying potentially better enzymes, because it utilizes a realistically complex cocktail (important because no biomass enzyme works in isolation) and a realistic substrate (important because results with pure cellulose or other artificial substrates may not be relevant to natural lignocellulosic materials).

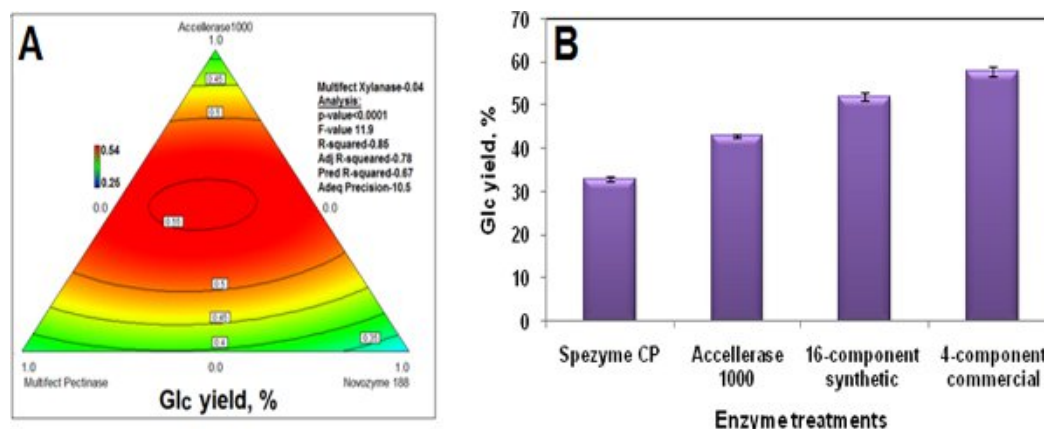
**(5) Commercial enzyme optimization.** Sufficiently large quantities of pure enzymes do not yet exist in the real world. Until they do, ethanol producers must rely on commercial products such as Accellerase 1000 and MultifectXylanase. However, mixtures of commercial enzymes generally perform better than any individual one, and GENPLAT can be used to design optimized cocktails of multiple commercial enzymes. Figure 3 shows the use of GENPLAT to optimize a mixture of four commercial enzyme preparations. The optimal proportions for release of Glc from AFEX-pretreated DDG were found to be 47% Accellerase1000, 27% Multifect Pectinase, 22% Novozyme 188, and 4% Multifect Xylanase (Banerjee et al., 2010c). Figure 3B shows the differences in Glc yield from AFEX-DDG with four enzyme preparations; the optimized commercial mixture gives higher Glc yield than Accellerase1000 alone, SpezymeCP alone, or a 16-component synthetic mixture. This experiment also indicates that the 16-component synthetic mixture is missing one or more enzymes needed for optimal Glc release, which are present in one or more of the commercial enzymes. GENPLAT can be used to identify such components.



**Figure 1. No single commercial enzyme preparation is optimal for all substrates and all pretreatments.** Five substrates and three pretreatments were compared under similar conditions of grinding (0.5 mm particle size), enzyme loading (15 mg/g glucan), and hydrolysis conditions (48 hr, 50°C). **A.** Digestion of different AFEX-pretreated substrates with Accellerase 1000. **B.** Digestion of corn stover pretreated by three different methods with Accellerase 1000 (see Banerjee et al., 2010c).



**Figure 2.** Optimal proportions of 11 enzymes for release of Glc from AFEX-pretreated corn stover (AFEX-CS), alkaline hydrogen peroxide-pretreated corn stover (AHP-CS), or AFEX-pretreated dried distillers' grains (AFEX-DDG). Data are from Banerjee et al. (2010c). Numbers along the top of the bars are the Glc yields as a percent of the total Glc in the biomass sample.



**Figure 3. A.** The use of GENPLAT to produce an optimized mixture of four commercial enzyme preparations for release of Glc from AFEX-pretreated DDG. Multifect Xylanase made the smallest contribution (4%) and was therefore omitted from this ternary diagram. **B.** Glc yields from AFEX-DDG with different enzyme treatments under identical hydrolysis conditions (15 mg/g glucan enzyme, 48 h, 50°C). "Four-component commercial" has the proportions determined from the experiment shown in Fig. 3A.

## Discussion

It is widely recognized that reducing the cost of enzymes is important to the development of an economical lignocellulosic ethanol industry. Currently available commercial enzyme cocktails are complex and poorly defined mixtures of many proteins (Nagendran et al., 2009), and they are adapted mainly for use on acid-pretreated corn stover. In order to accelerate the development of better enzyme cocktails, several laboratories have developed high-throughput platforms for enzyme discovery and characterization. Efforts in this area have incorporated one or more of the following properties also found in GENPLAT: robotic dispensing of enzymes and biomass slurries, statistical design of experiment, and/or automated determination of Glc and Xyl (Berlin et al., 2007; Decker et al., 2009; Kim et al., 1998; King et al., 2009). GENPLAT extends these earlier efforts, most significantly in the complexity of the enzyme mixtures that can be analyzed from at most 6 components in the earlier studies to more than 16 in our latest work (Banerjee et al., 2010c). Additional key features of GENPLAT are the use of a bead mixing chamber

(paddle reservoir) that can keep stover slurries suspended during dispensing; gentle mixing during digestion by end-over-end rotation; and automated colorimetric determination of Glu and Xyl.

## Disclosures

No conflicts of interest declared.

## Acknowledgements

This work was funded in part by the U.S. Department of Energy Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494) and grant DE-FG02-91ER200021 from the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences. We thank John Scott-Craig and Melissa Borrusch for their material and conceptual contributions.

## References

1. Anderson, M.J., & Whitcomb, P.J. *DOE Simplified: Practical Tools for Effective Experimentation*, 2<sup>nd</sup> Edition. Productivity Press, New York (2007).
2. Banerjee, G., Car, S., Scott-Craig, J.S., Borrusch, M.S., Aslam, N., & Walton, J.D. Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. *Biotechnol. Bioengineer.* **106**, 707-720 (2010a).
3. Banerjee, G., Car, S., Scott-Craig, J.S., Borrusch, M.S., Bongers, M., & Walton, J.D. Synthetic multi-component enzyme mixtures for deconstruction of lignocellulosic biomass. *Bioresour. Technol.* **101**, 9097-9105 (2010b).
4. Banerjee, G., Car, S., Scott-Craig, J.S., Borrusch, M.S., & Walton, J.D. Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnol. Biofuels.* **3**, 22 (2010c).
5. Berlin, A., Maximenko, V., Gilkes, N., & Saddler, J. Optimization of enzyme complexes for lignocellulose hydrolysis. *Biotechnol. Bioeng.* **97**, 287-296 (2007).
6. Decker, S.R., Brunecky, R., Tucker, M.P., Himmel, M.E., & Selig, M.J. High throughput screening techniques for biomass conversion. *Bioenerg. Res.* **2**, 179-192 (2009).
7. Gao, D., Chundawat, S.P., Krishnan, C., Balan, V., & Dale, B.E. Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. *Bioresour. Technol.* **101**, 2770-2781 (2010).
8. Harris, P.V., Welner, D., McFarland, K.C., Re, E., Poulsen, J.C.N., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S.Y., & Lo Leggio, L. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. *Biochemistry.* **49**, 3305-3316 (2010).
9. Kim, E., Irwin, D.C., Walker, L.P., & Wilson, D.B. Factorial optimization of a six-cellulase mixture. *Biotechnol. Bioeng.* **58**, 494-501 (1998).
10. King, B.C., Donnelly, M.K., Bergstrom, G.C., Walker, L.P., & Gibson, D.M. An optimized microplate assay system for quantitative evaluation of plant cell wall-degrading enzyme activity of fungal culture extracts. *Biotechnol. Bioeng.* **102**, 1033-1044 (2009).
11. Nagendran, S., Hallen-Adams, H.E., Paper, J.M., Aslam, N., & Walton, J.D. Reduced genomic potential for secreted plant cell-wall-degrading enzymes in the ectomycorrhizal fungus *Amanita bisporigera*, based on the secretome of *Trichoderma reesei*. *Fung. Genet. Biol.* **46**, 427-435 (2009).
12. Rosgaard, L., Pedersen, S., Langston, J., Akerhielm, D., Cherry, J.R., & Meyer, A.S. Evaluation of minimal *Trichoderma reesei* cellulase mixtures on differently pretreated barley straw substrates. *Biotechnol. Prog.* **23**, 1270-1276 (2007).