SynTec
Final Technical Report

Synthetic biology for Tailored Enzyme Cocktails

Janine Lin & Sarah Teter, Novozymes, Inc.
June 30, 2016

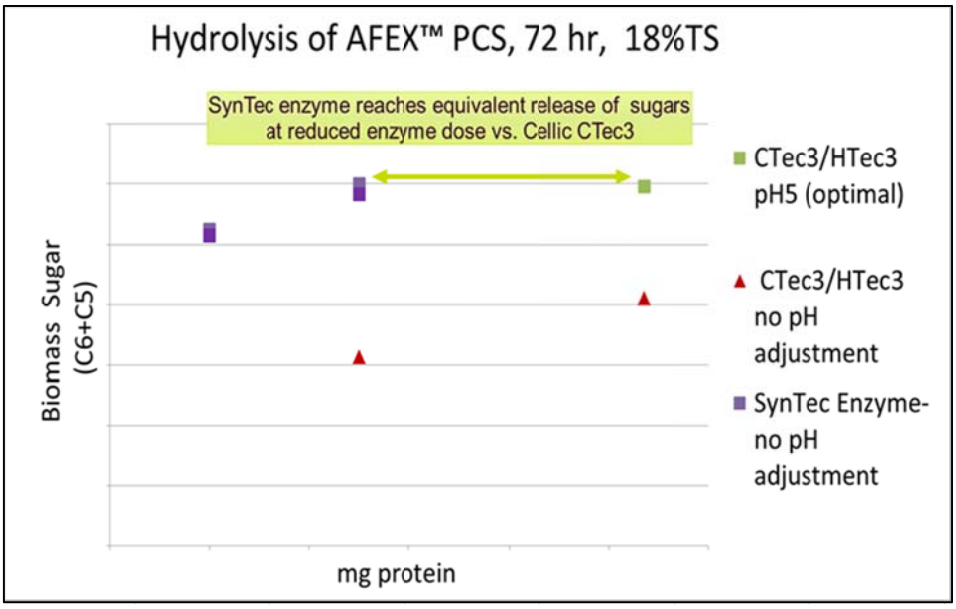
Prepared for the U.S. Department of Energy under
Award Number: DE-EE0006110

Project Team: Novozymes, Inc (Davis, CA);
Novozymes North America (Franklinton, NC);
MBI (Lansing, MI)

Executive Summary

Using a novel enzyme screening method inspired by synthetic biology, Novozymes developed new technology under SynTec which allows for more rapidly tailoring of enzyme cocktails. The methodology can be applied to specific feedstocks, and or coupled to address a specific hydrolytic conversion process context. Using combinatorial high throughput screening of libraries of enzyme domains, we can quickly assess which combination of catalytic modules delivers the best performance for a specific condition.

To demonstrate the effectiveness of the screening process, we measured performance of the output catalytic cocktail compared to CTec3/HTec3. SynTec benchmark cocktail - blend of Cellic® CTec3 and HTec3. The test substrate was - ammonia fiber expansion pretreated corn stover (AFEX™ PCS).CTec3/HTec3 was assayed at the optimal pH and temperature (pH 5.0 and 50°C- green squares in figure) and also in the absence of any pH adjustment (50°C- red triangles in figure). The new enzyme cocktail discovered under SynTec was assayed in the absence of any pH adjustment and at the optimal temperature (53°C- purple squares in figure). Conversion is delivered by SynTec enzyme at significant dose reduction relative to CTec3/HTec3 at the controlled pH optimum, and without titrant required to maintain pH, which delivers additional cost savings relative to current state of the art process.



In this 2.5 year $4M project, the team delivered an experimental cocktail that significantly outperformed CTec3/HTec3 for a specific substrate, and for specific hydrolysis conditions. As a means of comparing performance improvement delivered per research dollar spent, we note that SynTec delivered a similar performance improvement to the previous award, in a shorter time and with fewer resources than for the previously successful DOE project DECREASE, a 3.5 year, $25M project, though this project focused on a different substrate and used different hydrolysis conditions. The newly implemented technology for rapid sourcing of new cellulases and hemicellulases from nature is an example of Novozymes' continued innovation that results in more effective products for the advanced biofuel market.

Introduction

Novozymes and others have effectively used biotechnological methods to improve enzyme cocktails for conversion of pretreated materials to sugars. Cellic® CTec and HTec cocktails are robust enough to convert variously pretreated feedstocks, and support a range of biomass types, with strong performance in a broad range of process contexts. By delivering individual enzyme components, and optimized novel enzyme cocktails containing these components on a diverse set of pretreated feedstocks, Novozymes has delivered an exceptionally performing, robust enzyme platform for the industry in Cellic® CTec3/HTec3 (hereafter called CTec3/HTec3). Novozymes has recently worked with specific customers to further reduce costs by way of individualized enzyme development programs. Beyond delivering enzyme dose reductions, economic improvements may be achieved by delivering enzymes that are specifically designed to reduce operating costs. Examples of value generation through enzyme innovation include enabling milder pretreatments to allow for lower capital expenditures, reducing salts in process streams, and permitting the use of innovative pretreatment processes which produce lower levels of biological inhibitors.

Synthetic biology offers a route to allow for rapid "enzyme tailoring." The knowledge regarding mechanisms and paradigms exploited by biomass degrading enzymes, and the vast collection of genetic diversity that has been assembled by us and others over the past decade is powerful, but the diversity of processes that have been developed for converting biomass to product dictates that enzyme innovation must occur more rapidly than it has in the past if tailored enzyme solutions that "match" the diversity of processes. In addition, the explosion in number of available biological "parts" that must be screened, and the combinatorial expansion of potential groupings of these catalytic parts which may be assembled to look for effective synergies begs for a high throughput solution for examining the vast biological store of diversity.

Our concept for the SynTec platform was to develop a tool that permits direct screening of an array of different combinations of synergistic enzyme domains. Rather than rational selection of candidates, and individual production of gene products and an exhaustive assay of their properties, we developed a streamlined, modular system for direct screening of cellulolytic cocktails. The system allows for a more streamlined discovery timeline, and rapid discovery of cocktails that can meet process-specific requirements and thereby dramatically reduce overall process costs.

Approach

The key goal was to deliver a synthetic screening tool that enables rapid assessment of unexplored natural diversity to deliver cost effective enzyme solutions tailored to specific industrial biorefineries. As proof of concept for viability of the novel screening approach, we aimed to deliver a novel enzyme cocktail for hydrolysis of AFEX™ pretreated corn stover (AFEX™ PCS), allowing production of sugars that are clean enough to be used as intermediates for a diverse range of products.

The work falls under the DOE/EERE Biomass Program's biochemical conversion R&D goals, which focus on reducing the cost of converting biomass to sugars. Specifically, the technical barrier "Bt-G. Cellulase Enzyme Loading: Reducing the cost of enzymatic hydrolysis" (Biomass Multi-Year Program Plan 2012) was addressed in this project. The focus on technology development is aligned with the DOE's directive that innovations in the area of synthetic biology will be harnessed toward accelerating the pace of cost reductions for production of advanced biofuels.

The proof of concept deliverable (cost effective enzymes tailored for AFEX™ PCS) addresses the particular need for biomass fractionation processes which produce cleaner sugar streams, with the expectation that production of some advanced biofuels and chemicals other than ethanol may be more sensitive to salts and other inhibitors in process streams, due to limitations in downstream conversion of sugars to product, and product recovery.

Project Milestones and Results

Key Project Goals

The overall project objectives were as follows:

- Construct a screening platform that can be used to deliver highly effective, synergistic multi-enzyme cocktails through plug-and-play high throughput screening (HTS) of biomass deconstructing gene modules.
- Use the method to screen a large number of enzymes, thereby delivering an enzyme cocktail that enables cost-effective production of sugars from AFEX™ PCS.

Overview of project structure & Project timeline

The timeline of the R&D project activities are illustrated in Fig.1, including planned timeline and actual task durations.

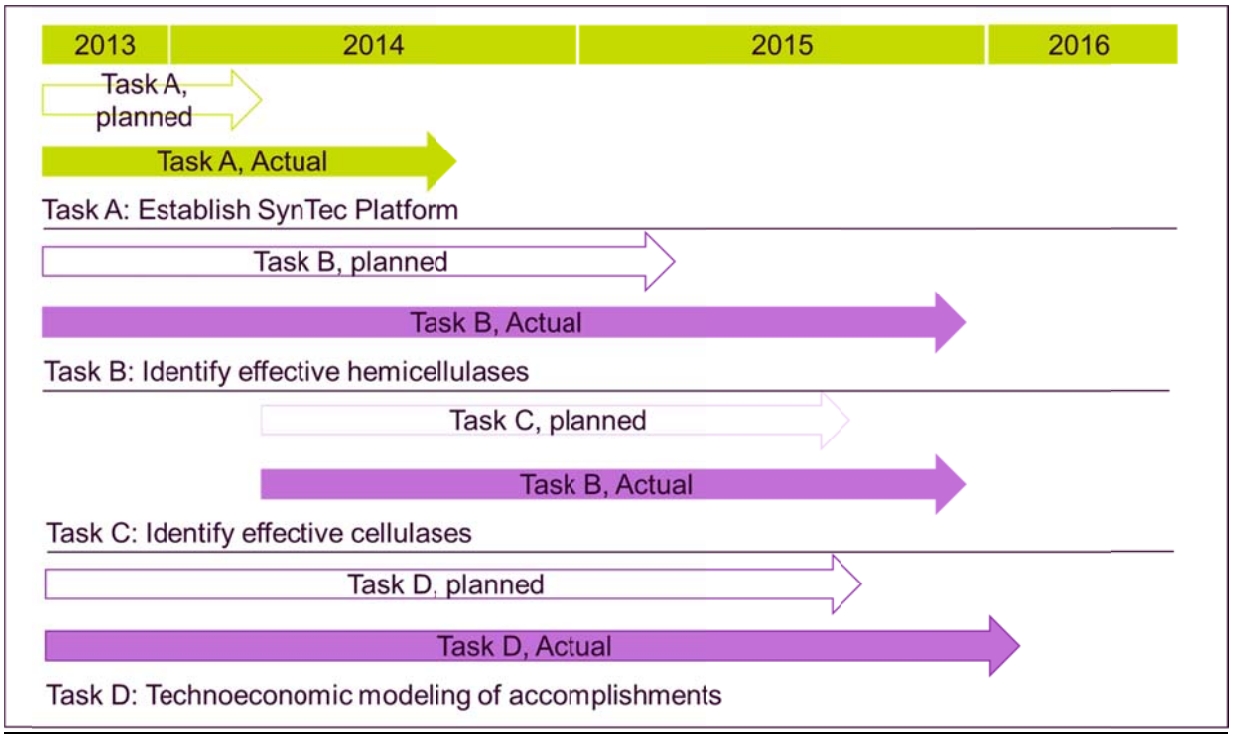


Figure 1. Timeline of the SynTec project activities. Task 1, Establish SynTec Platform, was planned from 0-6 months, and included molecular biology, enzyme discovery, assay development and automation activities to develop an integrated and validated screening system. Actual duration of Task A was 12 months. Tasks B & C, Identify effective hemicellulases/cellulases, respectively, focused on discovering and optimizing hemicellulase/cellulases cocktails for AFEX™-PCS. The tasks were initially planned to run in a staggered fashion, with duration of 18 months; actual duration on both tasks was longer. Finally, Task D focused on technoeconomic validation of accomplishments, demonstrating proof of concept of the screening methodology. The task encompasses initial validation, work on the technoeconomic analysis (TEA) throughout the project, intermittent testing of top enzyme candidates, and the final validation. We received a month no-cost extension to allow for additional candidate enzyme testing cycles, and finished the R&D work in January 2016.

Task A. Establish SynTec platform

Establish SynTec Platform: Assembly of parts into an integrated and validated system that can be used to dramatically shorten enzyme discovery times

A schematic of the screening platform delivered in the project is shown in Fig. 2.

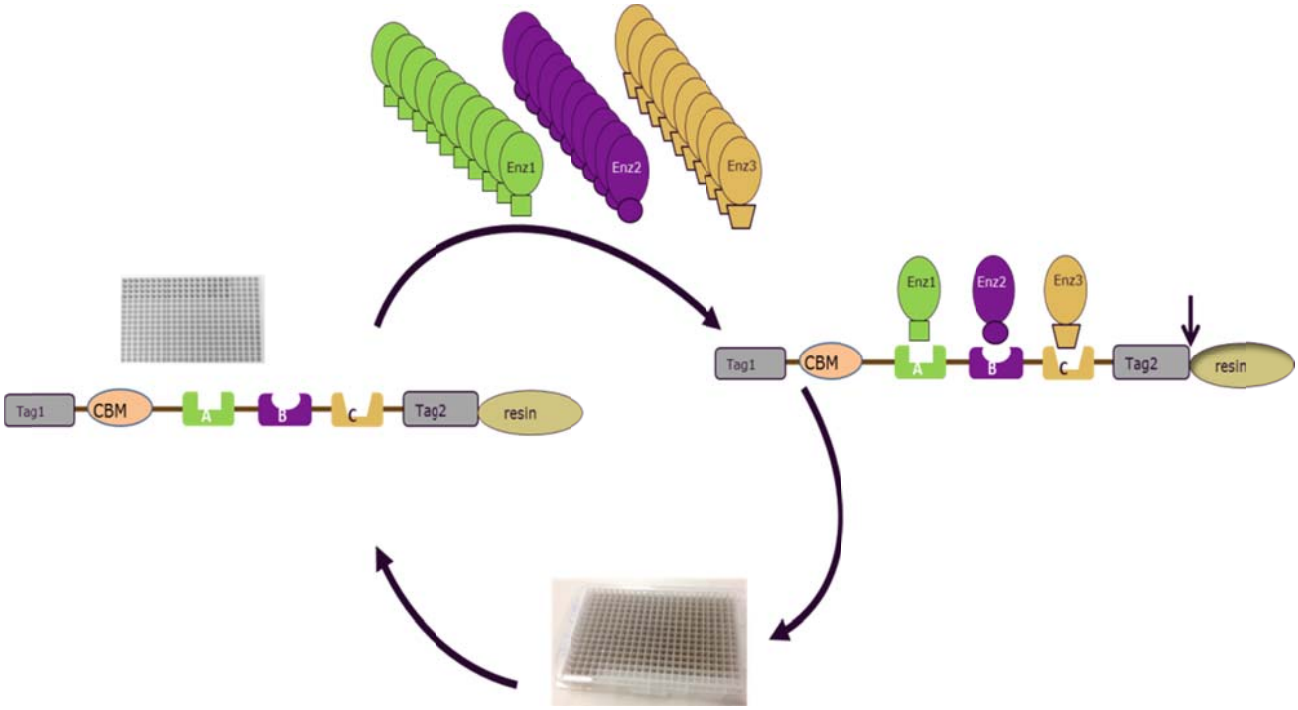


Figure 2. SynTec screening method. Left-Scaffoldin is depicted, containing an affinity tag for purification, a carbohydrate binding domain (CBM), three individual cohesin domains (A, B, C), and a second affinity tag (shown bound to affinity resin). The scaffoldin is attached to a resin and loaded into 348 well plates. **Top-Libraries** of candidate enzymes, organized by functional classes, are shown (Enz1- enzymes of functional class 1, etc.). Enzymes are expressed as fusions to cognate dockerin domains that bind specific sites of the scaffold. **Right- "Loaded" scaffold** is shown with enzymes of each functional class bound via dockerin:cohesin protein interaction. Dockerin enzymes are added to scaffoldin in individual wells in ≥3X molar excess, to saturate binding sites. After unbound proteins are washed away, the enzyme-scaffoldin complex is eluted from the resin. **Bottom-screening** A 384 microtiter plate is shown. The scaffoldin-bound dockerin enzymes are added to AFEX™-PCS biomass, with complementary enzyme classes, to form "complete enzyme cocktails." Hydrolysis of is monitored through indirect measure of glucose and xylose. The top candidates are selected, and subjected to further rounds of screening.

Construction of the scaffoldin was accomplished under SubtaskA.1, as detailed below.

Subtask A.1, Construct scaffoldin

Goals: Construct scaffoldin, with cohesin modules, CBM, and develop pull-down methodology. Demonstrate robustness and specificity of dockerin-cohesin binding under intended screening conditions.

Initially, we weren't able to produce a scaffoldin that showed suitable complex stability at 55°C for 3 days. After testing 12 different cohesin domains and 25 different scaffoldin constructs we identified an optimal synthetic scaffoldin for our screening conditions. The construct is thermostable, shows specific binding to cognate dockerins, and retains all bound dockerins over a 3 day time course at 55°C, under hydrolysis conditions (as

evidenced by native gel analysis; Fig. 3). A second challenge we encountered was generating scaffoldin containing all necessary domains; the synthetic multi-domain polypeptide led to a range of different length proteins upon heterologous expression (Fig 4B). We required a homogenous population of scaffoldin to allow for binding of all of our candidate enzymes. To enable easy "clean up" of expression broths resulting in purification of only the full-length scaffoldin, two different affinity tags were attached to the N and C-terminus of the protein (Fig. 4).

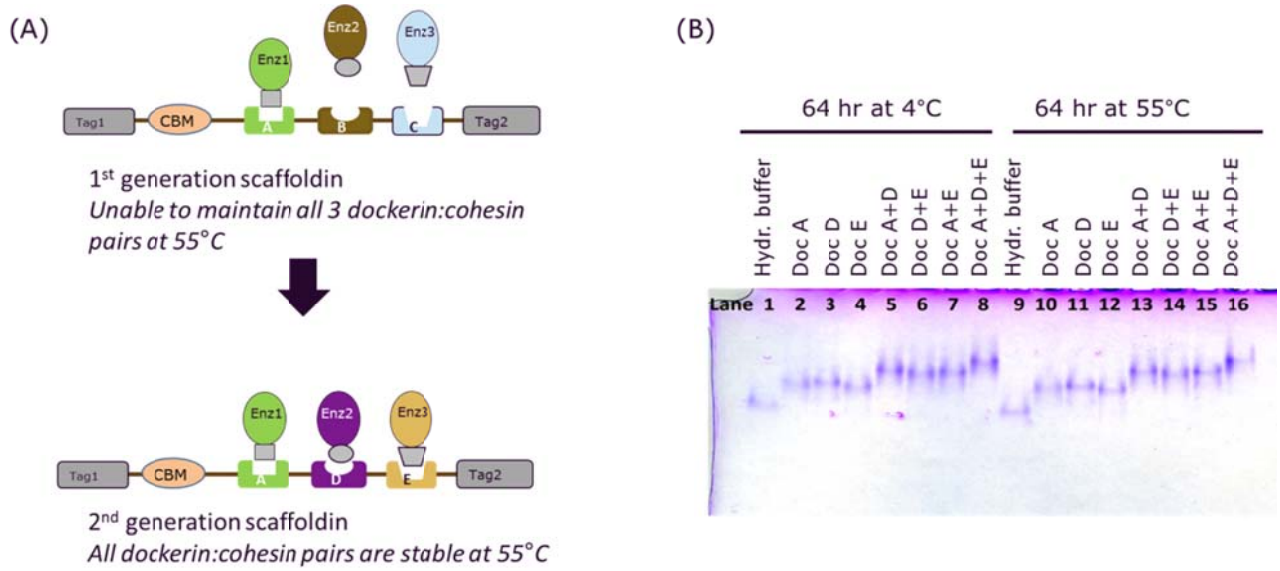


Figure 3. Generation of a stable scaffoldin. (A) Schematic illustration of the establishment of a stable scaffoldin. Interaction between each pair of the dockerin-cohesin is specific. (B) The native gel shows stable interaction between scaffold and docked proteins up to 55°C for 64 hr.

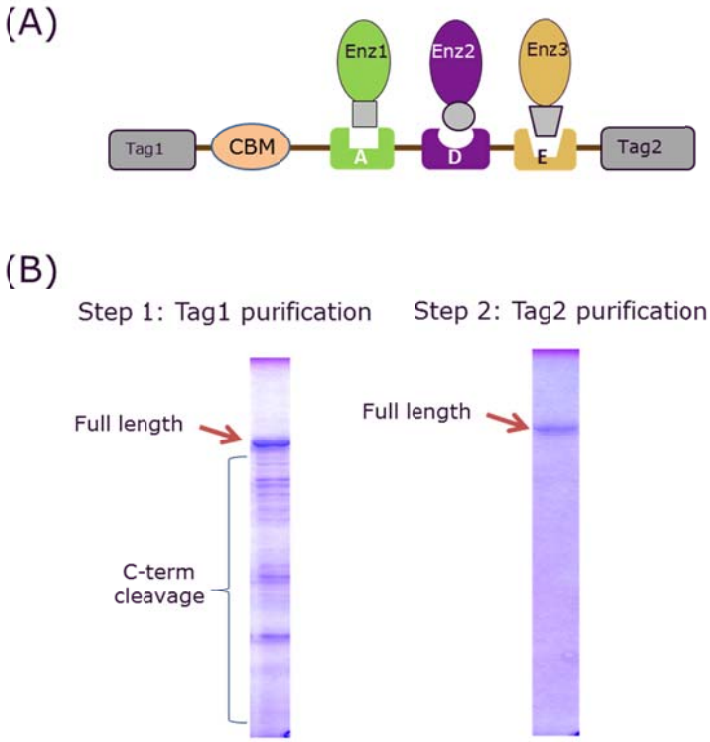


Figure 4. Scaffoldin expression. (A) Dual tags of scaffoldin construct were designed to ensure purification of full-length protein. (B) Two step purification of scaffoldin.

A second sub-task in this work package is construction of vectors, and host systems for production of libraries of genes of interest. The accomplishments under this subtask are described under Subtask A.2.

Subtask A.2 Dockerin library vector construction.

Goals: Molecular construction of vectors for wild-type enzyme module libraries, including cognate dockerins that bind to cohesins from subtask A.1, and additional tags for ease of purification.

Vectors for expression of library enzymes were completed. To assess expression and activity rapidly, we added a His-tag to enzyme-dockerin fusion proteins so that we can often ensure recovery of the enzyme after affinity purification. Expression of most test dockerin-fused enzymes was sufficient (Fig. 5). Specific activity measurements of the dockerin fusions were conducted in the context of a xylanase. Dockerin fusions do not appear to disrupt activity compared to the xylanase domain-only constructs. To test for potential limitations with high molecular weight enzymes binding to our scaffoldin, we looked at a 65 kDa enzyme with beta-xylosidase (bX) activity, and showed that the scaffoldin construct had the capacity to bind this protein in combination with other library enzymes.

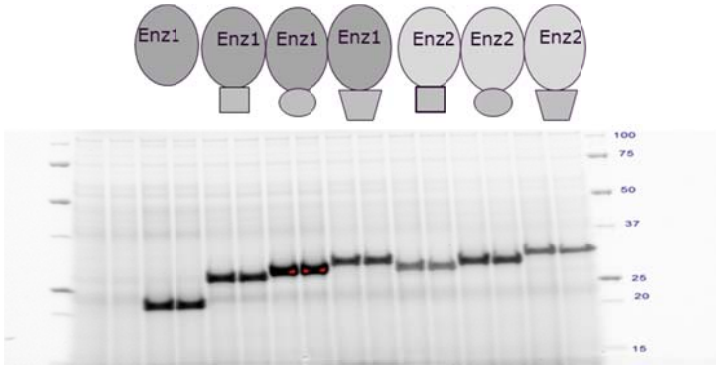


Figure 5. Enzyme-dockerin expression. Expression of two enzymes (Enz1 and Enz2) fused to three different domains, illustrated by three different shapes.

Under this subtask we demonstrated viability of two expression hosts, *Saccharomyces cerevisiae* and *Bacillus subtilis*.

We contemplated expressing multiple candidate genes in a single host cell, through operon based cloning. A specific subtask, A.3 addressed this decision point.

Subtask A.3. Evaluate two approaches for library construction of dockerin-enzymes
Goals: Determine whether operon expression or separate expression of individual library fusions is superior. Select based on throughput/effective combinatorial expression.

After a series of scoping experiments, we decided on individual transformations for each gene in the library ("No-go" on operon based cloning). The throughput was sufficient to meet the demands of the project.

The final subtask for this work package was assay development, including automation, and the project accomplishments are described below.

Subtask A.4 Assay development.
Establish pull down procedure, and assay for combinatorial screening of hemicellulases on AFEX™ PCS. Automate screens and develop data handling tools.

Automation of several steps was required for our screening, including: a) arraying enzyme candidates into multicomponent mixtures, b) isolating captured enzymes on the scaffoldin in the "pull down" step, and c) combining the scaffoldin/enzyme complexes with AFEX™-PCS in the hydrolysis step. All of these steps were established for both pH 6.2 (close to the pH of unadjusted AFEX PCS) and pH 5 high throughput screens. As our primary target was unadjusted pH, we report primarily on the candidates from the pH 6.2 screens.

A key aspect of the hydrolysis assay utilized for SynTec is that enzyme classes are assayed in the context of a complete enzyme mix. After working on various pretreated biomass substrates over the past years, we have been able to define the required enzyme classes needed for optimal hydrolysis of substrates. For hemicellulase screens, we queried performance of "loaded" scaffoldin to test enzyme library combinations in the presence of complete cellulase mixtures (also containing lytic polysaccharide monoxygenases, LPMOs). For cellulase screens, we tested combinations of cellulases (and in some cases containing LPMO) in the presence of bG and hemicellulase cocktails.

We optimized the unit operations in the screening process to reduce the noise in output sugar from the AFEX™ PCS hydrolysis assays. We observed <5% CV for controls where multiple repeats of a given enzyme domain loaded and isolated on scaffoldin were tested. Later use of the assay in screening of enzyme libraries showed consistent ranking of top candidates among independent screens. Finally, 384 well high throughput screens were established to minimize need for protein input. The screening flow is summarized in Fig. 2.

Variance, Task A

We originally anticipated that the task A should run for 6 months, while it in fact took 12 months. We generated the initial scaffoldin that was fully functional in capturing all 3 dockerins in 3 months. However it took us another 9 months to obtain a scaffoldin that maintained dockerin:cohesin binding throughout hydrolysis. The design of the stable scaffoldin ensured that co-localized domains screened stayed bound throughout the hydrolysis. The potential for improving enzyme performance through co-localization of individual enzyme domains was a direction we aimed to pursue with this screening system.

Because design of a the stable scaffoldin-enzyme complex required iterative testing of various cohesin modules, we worked with a series of generations of scaffoldin structures, each improved over the previous generation. This required parallel re-design of dockerin vectors, and re-cloning/ and re-expression of libraries of enzyme candidates; as we changed cohesin “landing pads” in the scaffoldin, this required that we coordinated with change of dockerin modules. Thus, additional time was needed complete dockerin-enzyme library generation, which impacted the timelines for Task B in particular (see below). Finally, the automation of the sub-steps required to screen libraries took more time than we initially planned.

Task B. Identify effective hemicellulases.

This task focused on discovering and optimizing hemicellulase cocktails for AFEX™-PCS

Subtask B.1 Hemicellulase library

Goals: Bioinformatic identification of hemicellulase candidates in libraries, codon optimization and gene synthesis, cloning into vector libraries, transformation into hosts.

Our hemicellulase genes are split into three libraries, namely xylanases (Xyn), beta-xylosidases (bXs), and “other” hemicellulases. Genes of each of the three libraries are constructed to express as fusions to a library-specific dockerin domain, to allow capture of all three types of enzyme activity on to scaffoldin (Fig. 6). Because Novozymes’ prior enzyme screening had focused primarily on fungal hemicellulases, we skewed the selection under SynTec to prokaryotic domains.

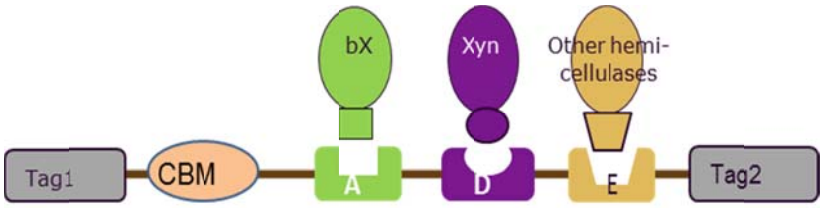


Figure 6. The scaffoldin/enzyme complex for hemicellulase screens.

For the bX, xylanase, or other hemicellulases, expression success rates ranged from of 50-80%, depending on class of enzyme. In addition to newly expressed enzymes, testing included candidates from Novozymes internal collection (screened as "free," that is, non-scaffoldin associated enzymes).

Subtask B.2 Hemicellulase screening.

Goals: Identify best hemicellulase modules. Perform iterative diversity generation for selected enzyme classes, with refined cocktails resulting from screening loops.

After several rounds of screening, we selected the best bX and another hemicellulase. Verification of the improved performance of these enzymes was obtained by assaying in the context of complete enzyme cocktails, comparing total sugar yield to controls where the same enzyme loading of benchmark cocktail (CTec3/HTec3) was tested.

Clear trends emerge in the hemicellulase screen data. Across the range of xylanases screened, there was a broad range in glucose release during hydrolysis. This finding is in line with a range of published studies that have demonstrated that xylanases are critical for effective glucose conversion in pretreated biomass. In contrast to xylanases, β-xylosidases candidates did not impact the glucan conversion, but instead could be differentiated by comparing the xylose released. Accessory enzymes tested to date have variable impact, where some classes only impact xylan conversion, while others also have some impact on glucan conversion.

Secondary screens were run to evaluate performance of top hits from the screen. We examined whether enzymes show superior performance as "free" enzymes (not bound to the scaffold complex), or whether improved activity is detected when enzymes are co-localized with partner enzymes on the scaffold complex. We have only seen one case where one specific enzyme mixture showed better activity when scaffold co-localized domains were compared to the same combination of free enzyme. Follow-up experiments using the protein thermal shift assay showed that this difference appears at least partially due to the lower thermostability of the xylanase in the enzyme mixture in its "free" form.

Variance, Task B

Significant work was expended to re-cloning domains of interest as we deployed different dockerins to match the designed scaffoldin cohesin structure (as we evaluated different dockerin/cohesin pairs, see Task A). It took us longer than expected to complete Subtask B.1., but we eventually provided the required diversity in each of the three hemicellulase libraries.

Task C. Identify effective cellulases

This task focused on discovering and optimizing hemicellulase cocktails for AFEX™-PCS.

Subtask C1 Cellulase library

Goals: Identify best cellulase modules. Perform iterative diversity generation for selected enzyme classes, with refined cocktails resulting from screening loops

The cellulases screened in this project include the following enzyme classes: cellobiohydrolases I (CBHIs), CBHIIs, endoglucanases (EGs), and beta-glucosidases (bGs); though they are not true cellulases, lytic polysaccharide monooxygenases (LPMOs) were included in this set. Additionally, a few other enzyme classes not belonging to the hemicellulases or above cellulase categories were also screened. The dockerin fusion enzyme libraries were split into three libraries, namely CBHI, CBHII/LPMO, and EG. Some of the EGs tested are likely to work processively in addition to cleaving internally. Because Novozymes' prior enzyme screening had focused primarily on fungal hemicellulases, we skewed the selection under SynTec to prokaryotic domains, as we did for the hemicellulases. Genes of each of the three libraries were constructed to express as fusions to a library-specific dockerin domain, to allow capture of all three types of enzyme activity on to scaffoldin (Fig. 7).

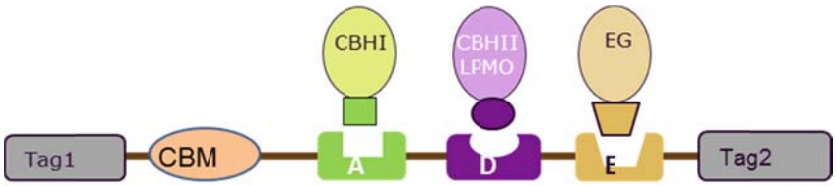


Figure 7. The scaffoldin/enzyme complex for cellulase screens.

For cellulases/ "other enzymes," the rate of successful expression was above 80%. Full length CBHI, CBHII, LPMO, and EGs were screened after pull-down by scaffoldin, while bG and truncated CBHI, CBHII, LPMO, and EGs were screened as free enzymes (not bound to scaffoldin). In addition, enzymes from Novozymes collection were screened as free enzymes. Among the enzyme classes, we were able to achieve the desired library size, with the exception of CBHII/LPMO class, where obtaining a large diversity set for combinatorial screening was challenging. There are very few bacterial CBHII/LPMO enzymes that fulfilled our screening criteria (predicted high temperature active with known biomass hydrolysis activities in source organism). In contrast, numerous bacterial EGs were available for screening.

Subtask C.2 Cellulase screening

Goals: Identify best cellulase modules. Perform iterative diversity generation for selected enzyme classes, with refined cocktails resulting from screening loops.

A total of six cellulase screens were run to complete this subtask. For all cellulase screens performed, we consistently observed similar performance among CBHI library candidates. It was difficult to differentiate among the CBHII candidates at pH 6.2, though one, designated H002, was slightly better than the other CBHII candidates at pH 5. Among the LPMOs, one candidate, called H005, performed slightly better than the other LPMOs at pH 6.2. The performance of the LPMOs looked more uniform at pH 5, overall, while one, coded H007, performed slightly better than others under this condition. In contrast to these enzyme classes, EG library candidates showed significant variability in the screens. The bacterial bGs showed poor activity overall, and were screened off-line (*e.g.,* as free enzymes, not bound to the scaffoldin).

As a result of the screening in the project, we found new CBHI, CBHII, LPMO, EG, and two enzymes from other classes to have higher hydrolysis activity under target conditions when compared to 100% CTec3 enzyme broths when single components were used to replace part of the CTec3 enzyme broths.

Variance, Task C

Cellulase screening was delayed as a result of our extending hemicellulase screening (which, as discussed above, was mainly due to our need to re-clone library candidates with appropriate dockerin domains).

Task D. Technoeconomic modeling of accomplishments.

This task allowed us to verify the feasibility of using the SynTec method to rapidly tailor enzyme cocktails. The proof of concept work, tailoring of cocktails for specific use in conversion of AFEX-™to sugars, was evaluated by MBI and Novozymes.

Subtask D.1 Production of AFEX™-PCS for screening and benchmarking.

A single batch of AFEX™ PCS pellets was used to generate proposal data, and also used for initial and final performance validations (called "validation substrate"). Use of a single batch reduces variability between comparative tests, allowing us to better isolate the differences in enzyme performance. However, insufficient quantities of this particular batch were available to support the screening activities for the full duration of the project. MBI was tasked with delivering a similarly pretreated batch of AFEX™ PCS in the first few months of work, which they accomplished. While the PCS was pretreated under identical conditions as for our validation substrate, we noted that the same CTec3/HTec3 baseline enzyme preparation consistently showed higher conversion on the "screening substrate" relative to our proposal and initial validation results with the "validation substrate." All technoeconomic analysis for initial and final validation was based on results from the "validation substrate."

Subtask D.2 Benchmarking of improved enzyme cocktails.

To inform the screen design and to validate the concept of performing hydrolysis without pH adjustment, MBI and Novozymes engaged in testing different enzyme loadings in our standard validation assay, testing different blends of CTec3/HTec3 at both pH 5 and unadjusted pH, assessing pH drift at 3, 5, and 7 days hydrolysis, and understanding contamination risk at both pH conditions. Briefly, the profile of pH drift during hydrolysis at unadjusted pH was captured, informing the screening assay design parameters (Task A.4). Contamination risk was tested in range of set ups and it was concluded that control of infection would be feasible in a commercial scale facility.

The primary focus for this task was collaborative testing to verify the top candidate enzymes for improved performance in hydrolysis. Novozymes performed smaller scale assays, at low solids and at high solids, to refine enzyme mixtures and select top candidates for larger scale tests at MBI. Typical assays employed shake flask scale analysis, 18% solid loading, unadjusted and pH 5 buffered conditions, with samples taken at 24, 48, 72 and 168 hr for sugar analysis. Varying doses were applied to estimate required protein needed to achieve a given level of biomass conversion to sugar on a glucan and xylan basis ("dose response curves"). Because top enzyme candidates showed superior or equivalent performance as free enzymes relative to performance when co-localized to the scaffoldin, we assayed performance as free enzymes under Task D. MBI performed assays where candidates were incorporated into a CTec3/HTec3 blend, or fully synthetic blend (comprised of purified proteins), and compared to the benchmark CTec3/HTec3 blend at the same total protein dose. Initially, we saw dramatic improvements in xylan conversion relative to the benchmark at unadjusted pH, due to incorporation of key hemicellulases. Improving glucan conversion was more challenging, but as we began to test cellulases in addition to hemicellulases, we gradually improved C6 sugar yield in the validation style assays. Towards the end of the project, MBI and Novozymes collaborated to optimize the ratios of components in SynTec cocktails. We employed response surface methodology at small scale (plate based assays), and followed up on a limited set of high performing mixtures, using the validation SOPs, to identify optimal ratios of components for the final performance validation.

Subtask D.3. Validations with external validation team.

Our validations of the starting and final enzyme performance were essential in evaluating the viability of our new screening platform for rapid tailoring of synthetic cocktails. Both validations (initial and final) were performed at MBI with an external validation team (CNJV/NREL/DOE) along with Novozymes and MBI team members.

The initial validation visit was completed in the week of September 23rd, 2013. Prior to the visit all of the relevant SOPs, detailed schedule for the visit, and list of the planned experiments were sent to the validation team for their review and approval. The initial validation verified the experimental data and cost information provided in the SynTec proposal; this serves as a baseline of commercially relevant enzyme performance on AFEX™-PCS at pH 5. Additionally, a baseline of performance for CTec3/HTec3 was generated at unadjusted pH. Specifically, the validation experiments included triplicate runs of:

1. Avicel (internal control, for comparison to historical runs with CTec3/HTec3)
2. No substrate +enzyme
3. No enzyme + substrate
4. Experiment at pH adjusted to 5.0, with buffer
5. Experiment without pH adjustment and without buffer

In the resulting initial validation report, the external team confirmed the reasonableness of Novozymes' technoeconomic model, which is used to estimate resulting sugar price. Evaluations of process and cost data, process engineering were accomplished through observation of the process, review of supporting documentation, and critical discussions with the team members.

During the course of the project, the MBI team refined the technoeconomic model to ensure an accurate assessment of the current state of technology. As a result of this work, a new model, based on the Humbird 2011 NREL economic model, was constructed. In addition to the NREL model, internal Novozymes numbers were used for the cost of enzymes, internal MBI values were used for the cost of the AFEX™ pellets, and a modified saccharification section based on MBI's experience with large-scale liquefaction of AFEX™ pellets was used instead of the base NREL saccharification model.

The model boundary was assumed to be sugar production, as was used in the proposal. Because sugar is an intermediate product, and that the form of that intermediate product would depend on the final product, it is impossible to have one model that is directly comparable across all scenarios. To simplify matters, the sugar consumer was assumed to be co-located with the sugar producer, and thus the intermediate product was not refined. It was assumed that the sugar stream would not be concentrated, although the lignin would be removed. As such, it was assumed that no wastewater treatment would be performed in the model. However, the soluble components were not included as potential sources of electricity either. Given that all of these extra steps would be identical for all scenarios, this approach was deemed adequate.

The variables changed throughout the model, based on experimental costs, are as follows:

- The amount of enzyme used, which impacts enzyme operating cost and capital cost of storing enzymes
- Whether or not pH is controlled, which impacts capital and operating costs of storing and handling sulfuric acid
- The residence time, which impacts capital cost of saccharification
- The hydrolysis temperature, which impacts energy use and therefore the electricity credit
- The sugar yield, which impacts the cost per lb sugar

For simplicity, secondary effects, such as the amount of solids recovered in each scenario, were not taken into account, as they were not measured in the process. It is assumed that all downstream operations are identical in order to focus the model on the saccharification process.

The final performance validation was held at MBI, starting 11 Jan 2016, where the MBI team hosted Novozymes along with external validation team (DOE, Allegheny Science & Technology, NREL). The specific cost assumptions of the resulting refined model, as developed by MBI over the project, were discussed in detail at the final performance validation, including specific capital and operating cost assumptions. During these discussions it was evident that we had not fully captured the expected cost benefits of removing a pH adjustment step prior to hydrolysis. Some of the expected benefits are difficult to model given the boundaries of the economic model, which essentially ends after sugar production step. Following the validation, we delivered a document detailing potential areas where further cost savings might be realized, depending on the processes following sugar production. Some of the issues evaluated included waste water pretreatment, recovery of products from process streams at with reduced salts, ability to ferment with *Z. mobilis* without requirement for a pH adjustment step, and cost savings associated with eliminating the need for pH control and monitoring in the hydrolysis and fermentation tanks.

To assess the final enzyme cocktail, comprised of top candidates from Tasks B and C, the SynTec team provided two improved enzyme cocktails for assay during the final performance validation. These enzyme cocktails were compared to the starting benchmark, CTec3/HTec3, as internal controls to compare with proposal data and initial validation data. The two new cocktails were "semi-synthetic" cocktails, meaning that they included a complex *T. reesei* expressed enzyme (CTec3), augmented with newly delivered individual purified enzymes.

SynTec cocktails and the CTec3/HTec3 benchmark were compared. The same mass of CTec3/HTec3 product blend was dosed in the final validation experiment for pH 5 condition as in the initial validation and as in the proposal. The key assumption for benchmarking purposes is that SynTec cocktails could be produced at same cost per gram protein as CTec3/HTec3. BCA assay of the desalted enzyme preparations were performed, utilizing the method provided prior to the initial validation (Aug 2013). As improved performance was expected for the SynTec cocktails *vs.* CTec3/HTec3 benchmark, the new cocktails were loaded at lower enzyme dose relative to the benchmark commercial enzyme preparation. Each enzyme was assayed at its optimal temperature (50^{o} C for CTec3/HTec3; 53^{o} C for the SynTec cocktails). As the targets for the project are measured against CTec3/HTec3 performance at the benchmark cocktail's optimal pH, a key experiment is CTec3/HTec3 assayed at pH 5. CTec3/HTec3 was also run at unadjusted pH, at two loadings; this allows "apples to apples" comparison of what improvement in enzyme performance has been made under the challenging condition where pH is not controlled. For all conditions, 3 replicates were run, at 18 % TS, in 50 g samples, with 24 hr, 72 hr, and 168 hr time points taken. Additional analysis was run to measure background sugars in each of the enzyme preparations. Further, a control reaction was run with Avicel as the substrate, to allow direct comparison of CTec3/HTec3 performance on a model substrate between the initial and final validation. Finally, a no-enzyme control set was run to measure free sugars coming from feedstock.

Results, shown as average yield obtained glucan + xylan (G+X) conversion basis, at 72 hr is plotted in Figure 8. Yields obtained at the higher dose for CTec/HTec3 at pH 5 were comparable to yields obtained at a lower dose for both SynTec blends at unadjusted pH (a similar trend was seen at 168 hr, data not shown). This allows direct quantification of dose reduction achieved relative to the project baseline condition, and significant performance improvement of the SynTec cocktails is evident. SynTec cocktails assayed without pH control performed as well as the CTec3 commercial benchmark at the optimal controlled pH, even with significantly less protein. The comparison of SynTec blends to CTec3/HTec3 where both were challenged with no pH control in the reactor

showed dramatic improvement of the SynTec blend over the CTec3/HTec3 benchmark. As the CTec3/HTec3 blend never achieves the same level of conversion as the SynTec blend at the unadjusted pH, it is not possible to provide a fold performance (dose reduction) estimate.

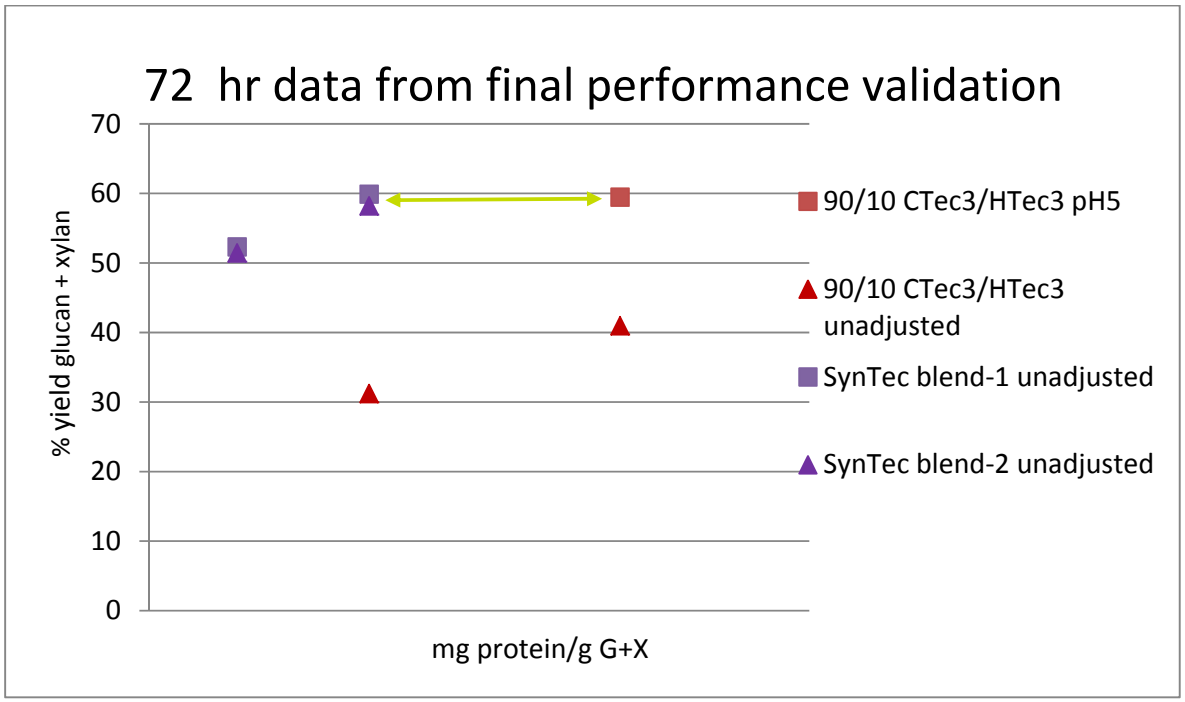


Figure 8 Dose conversion plots of CTec3 benchmark and SynTec blends in the final performance validation assays. Green arrow denotes equivalent performance between SynTec blends with no pH adjustment and CTec3/HTec3 at optimal pH.

In addition to allowing a measure of performance improvement through measured dose reduction to achieve a given level of conversion, the data obtained in the final performance validation was evaluated to determine the conformance between initial and final validations. CTec3/HTec3 were used in both validations, using the same lot of enzymes, and the same dose was applied (mass product basis). Comparison between glucose yields observed in Aug 2013 and in Jan 2016 showed less than 5% variance at both 72 hr and 168 hr time points. However, the xylose released was shown to be lower in the final validation, with ~11-14% variance observed. Comparison of the Avicel control, also run in both validations and with the same enzyme prep, showed very good reproducibility (2% variance), suggesting that there is good stability of the cellulase preparation. We considered whether hemicellulase instability over 2.5 years storage at -20 ^{o}C might have contributed to the instability in benchmark. Another source of variability is that the working assay volume changed, from 100 g to 50 g, due to limitations in enzyme availability. Finally, we speculate that changes in the AFEX™ PCS pellets might have occurred over the 2.5 year, changing enzymatic susceptibility, despite the fact that no change in composition can be observed over years of storage of AFEX™ PCS (MBI compositional data).

Variance, Task D

Our project deliverable is measured in reductions in sugar price achieved through enzymatic performance improvements. Given the change in CTec3/HTec3 baseline enzyme performance (xylan activity), it was necessary to look at the relative cost reduction in modelled sugar price within each given experiment. We utilized the TechFin tables to evaluate the extent that cost reductions were delivered through SynTec enzyme innovations.

Our specific object was to reduce the modelled production costs per lb of sugar, from $0.31 at optimal pH (maintained at pH 5), to $0.21 at the more challenging unadjusted pH condition. This represents a 1.55X reduction in production costs. Using the conversion yields achieved in the final validation, and assuming an equivalent production cost per mg enzyme as for CTec3/HTec3, we were able to model the net sugar production costs. Comparing the sugar price to the benchmark price delivered through use of CTec3/HTec3 at pH 5 (from final validation experiment conversion data), we saw a reduction in price of sugars from 1.2-1.4X (depending on whether 72 hr or 168 hr conversion data formed basis of the comparisons).

This falls short of our ambitious target for reducing the modelled price of biomass derived sugars, but is a significant achievement for 2.5 year of R&D, especially considering that a long period of the project was dedicated to developing a new screening method (~ 1 year). To place the magnitude of the improvements delivered in context, it is useful to consider that our previous DECREASE award delivered similar performance, though this project focused on a different substrate and used different hydrolysis conditions.

Acknowledgements

Key research contributors for SynTec include Novozymes employees Johnnie Hahm, Paul Harris , Sumati Hasani , Ian Haydon, Tia Heu , Aubrey Jones, Michael Lamsa , Janine Lin, Fang Liu, Ronald Mullikin, Ani Tejirian, Carly Todd, William Widner, Elizabeth Znameroski , James Broering, Grace Cooley, Kurt Creamer, Jesper Frickmann, and Sarah Teter. Key contributors and MBI include Farzaneh Teymouri and Bryan Bals.

Publications / Presentations

Lin, Janine. “Discovering optimally tailored enzyme cocktails using a synthetic screening tool”, Oral presentation at Symposium on Biotechnologies for Fuels and Chemicals, April 27-30, 2015, San Diego, CA.

Hahm, J.B. “Development and application of a synthetic cellulosome-based screening platform for enhanced enzyme discovery”, Poster at 11th Carbohydrate Bioengineering Meeting, May 10-13, 2015. Espoo, Finland.

Haydon, Ian. “Development of modular, self-assembling protein complexes as platforms for enzyme discovery”, poster at AIChE’s 2nd Synthetic Biology: Evolution, Engineering, and Design Conference. June 10-13, 2015. Boston, MA.

Teter, S. and Creamer, K. “R&D at Novozymes in Biomass Conversion: Advancing towards a Biobased Economy” Oral presentations at Great Lakes Bioenergy Research Center Annual Retreat. May 21, 2015, South Bend, IN.

Patent applications

1. Polypeptides having xylanase activity and polynucleotides encoding same –13145-US-PRO, 13145-WO-PCT.
2. Polypeptides having arabinofuranosidases activity and polynucleotides encoding same – 13147-US-PRO, 13147-WO-PCT.
3. Polypeptides having beta-xylosidase activity and polynucleotides encoding same –13157-US-PRO, 13157-WO-PCT.
4. Polypeptides having endoglucanase activity and polynucleotides encoding same – 14152-US-PRO.

5. Endoglucanase for neutral pH, high-temperature lignocellulose hydrolysis - Filing report: First Patent Application (14153-US-PRO)

Biocontainment challenges, unexpected outcomes, and potential societal implications

SynTec delivered a tool for enzyme discovery. Containment of genetically modified organisms is tightly controlled at Novozymes, Inc, and all recombinant *S. cerevisiae* and *B. subtilis* libraries generated through use of the SynTec screening method were maintained on site. We have procedures in place that prevent release of viable GM organisms.