



# DOE Bioenergy Technologies Office (BETO) 2019 Project Peer Review

## Process Analytical Technologies for Conversion Operations (Previously named Advanced SCADA for Biochemical Process Integration)

3/5/2019

Biochemical Conversion

**Jim Collett - PNNL**

This presentation does not contain any proprietary, confidential, or otherwise restricted information

# Goal Statement

**The Challenge:** BETO-supported bioconversion pathways assume deployment of fed-batch control systems for production of hydrocarbon fuels and commodity chemicals from crude, high-solids hydrolysates. Such control systems will require **real-time sensor systems for tracking hydrolysate components** that are not yet commercially available.

**The Goal:** We will, in collaboration with industrial partners, advance technology to market transfer of **Process Analytical Technologies (PAT) for bioreactor control systems** that enable profitable bioconversion of biomass feedstocks to advanced biofuels and commodity chemicals.

**The Outcome:** **Risk is reduced in achieving BETO objectives** by providing biorefineries with proven PAT equipment and methods that integrate seamlessly with commercial plant control systems.

# Quad Chart Overview

## Timeline

- Project Start: 10/1/2017
- Project End: 9/30/2020
- Percent Complete: 42%

	Total Costs Pre FY17**	FY 17 Costs	FY 18 Costs	Total Planned Funding (FY 19-Project End Date)
<b>DOE Funded</b>	\$900K	\$300K	\$250K	\$500K
<b>Project Cost Share*</b>				

## Partners:

**Industry: MarqMetrix, Sartorius**

**DOE: NREL, ABPDU, INL, ORNL, FCIC**

## MYPP Barriers Addressed

- Ct-N. Multiscale Computational Framework toward Accelerating Technology Development
- Ct-K. Developing Methods for Bioproduct Production
- Ct-D. Advanced Bioprocess Development
- Ot-B. Cost of Production

## Objective

Provide advanced process analytical technologies for bioconversion of variable, high-solids feedstocks that support online, real-time tracking of sugar substrates, targeted products, and untargeted byproducts within bioreactors, and enable the use of automated fed-batch control systems that are currently not feasible with the present state of process analytical technology.

## End of Project Goal

Demonstrate a PAT-enabled bioreactor chemostat that can automatically identify pH, DO, and temperature conditions that optimize microbial growth rate and/or specific productivity of a commodity chemical in hydrolysate medium over a 2-week period without manual adjustment of the controls.



# 1 - Project Overview



**New spectroscopy tools** for bioconversion of **high-solids feedstocks** are being developed with industrial partners to support advanced **fed-batch control systems** for production of fuels and chemicals.

FY15

FY16

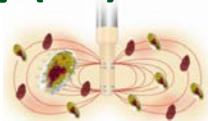
FY17

FY18

FY19

FY20

## Dielectric spectroscopy (DS)



In the FY15-17 AOP cycle, **dielectric spectroscopy** and **near infrared spectroscopy tools** were developed for tracking cell mass growth and for fed-batch control during high-solids bioconversion.

## Near-infrared spectroscopy (NIRS)



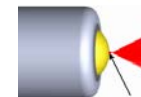
sartorius



In the FY18-20 AOP cycle, **Raman spectroscopy tools** and an **integrated data management system** are being developed for high-resolution prediction of the concentrations of bioconversion intermediates and products.

## Raman spectroscopy (RS)

MARQMETRIX



## Chemometric Data Integration



**Bench to  
Pilot Scale  
Bioreactors**

## 2 – Approach (Management)

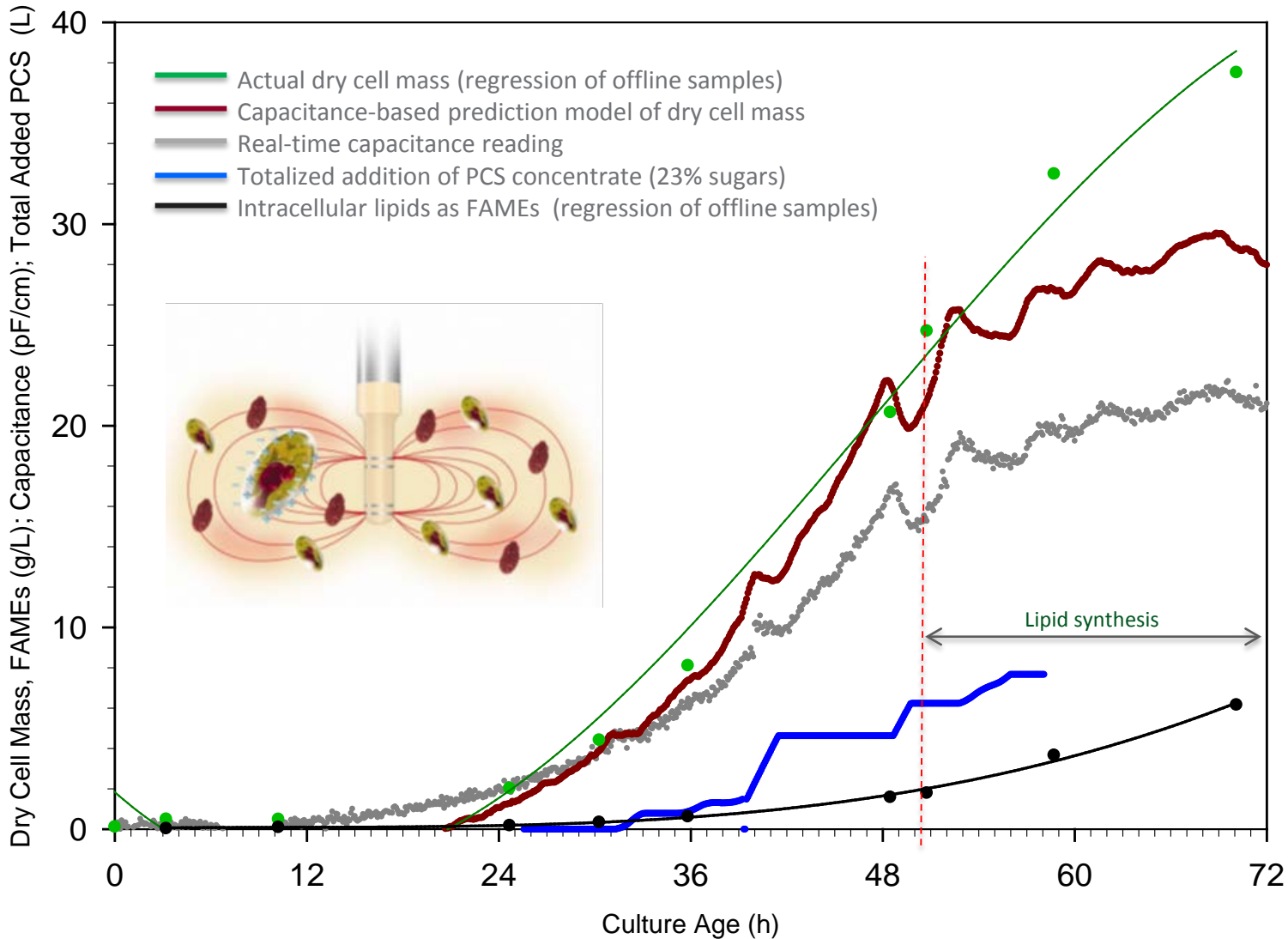
- Standard Project Management Good Practices
  - Statement of work and how it relates to DOE goals
  - AOP Project Management Plan
  - Quarterly milestones
  - Go/No Go decision point
- Frequent project communications
  - Teleconferences with BETO platform leads
  - Quarterly formal reporting to BETO
  - Site visits by industrial collaborators
- Draw on deep talent pool in analytical sciences, process control, machine learning, and chemical engineering at PNNL

### Critical Success Factors:

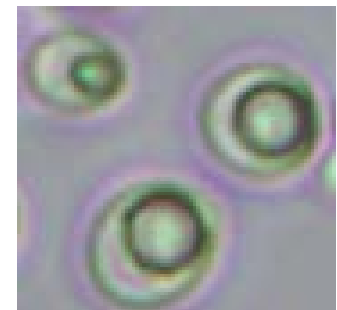
- Use **SCADA-compatible, commercial-off-the-shelf (COTS) equipment** when possible.
- Achieve **real-time tracking and control of critical process parameters**.
- Ensure results are **generalizable for a variety of biocatalyst organisms**.
- Validate performance in **crude, high-solids, biomass feedstocks**.

## 2 – Approach (Technical)

FY17: Dielectric Spectroscopy was validated for tracking growth of *Lipomyces starkeyi* yeast in coarse-filtered, pretreated corn stover (PCS) hydrolysates in a 30-liter bioreactor containing up to 42.5 g/L insoluble lignin.



PCS  
Feedstock

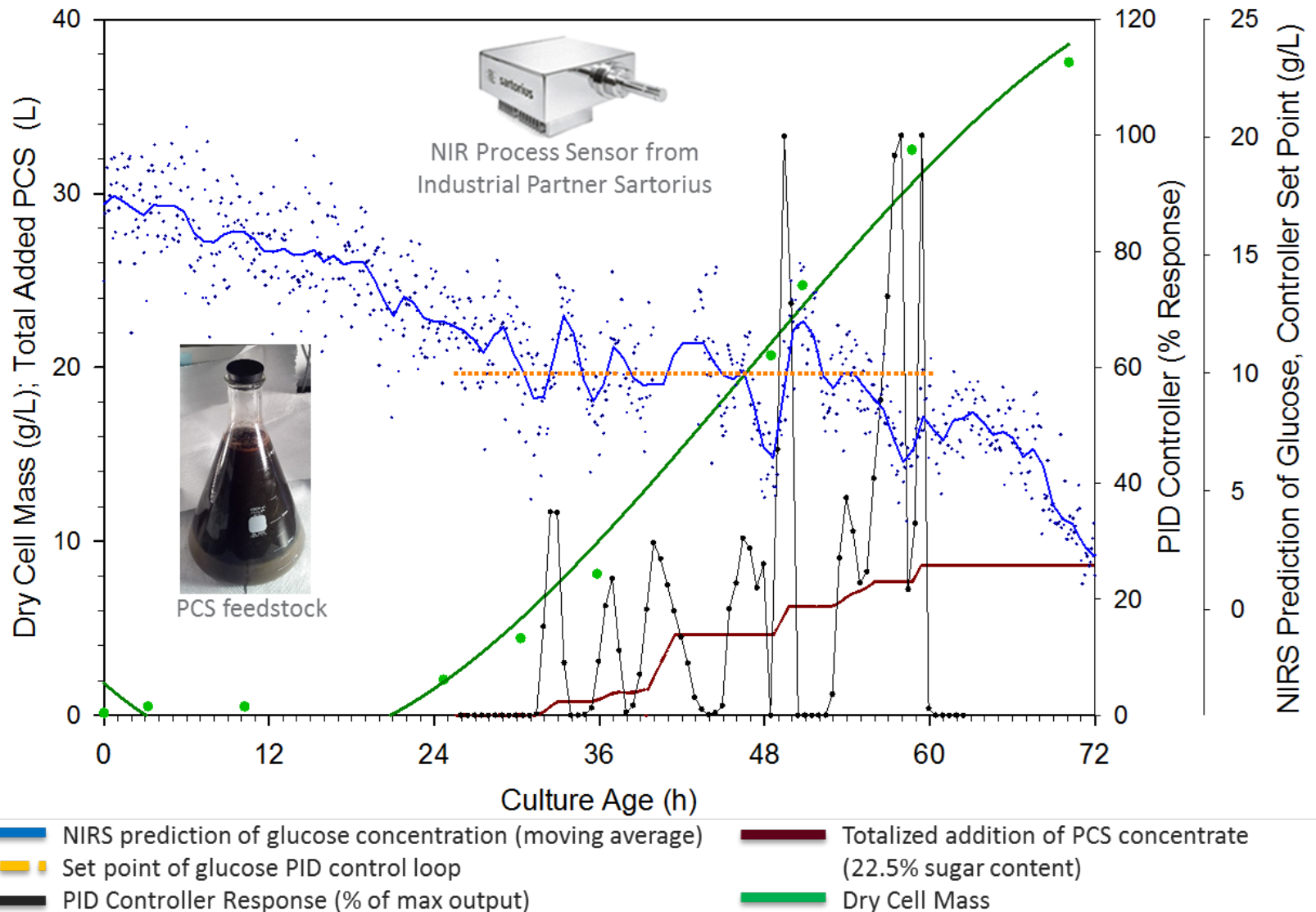


*L. Starkeyi*  
oleaginous  
yeast cells

## 2 – Approach (Technical)



FY17: Fed-batch control of glucose concentration was maintained in a 30-liter bioreactor cultivation of *Lipomyces starkeyi* yeast using near-infrared spectroscopy (NIRS) in deacetylated, dilute acid pretreated corn stover hydrolysate (DDA PCS) over a period of 36 hours.



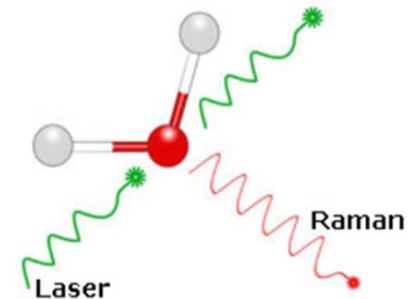
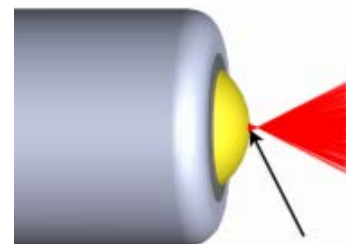
## 2 – Approach (Technical)

FY18: [Raman Spectroscopy](#) for tracking of substrates and products during hydrolysate bioconversion

We have partnered with [MarqMetrix](#), a University of Washington spin-off company, to adapt their [All-in-One \(AIO\) Process Raman Analyzer](#) for use in biomass conversion applications.

[Raman spectroscopy](#) measures inelastic scattering of photons from a monochromatic light source as they interact with the electron cloud and chemical bonds of a molecule.

The [MarqMetrix TouchRaman® BallProbe®](#) has a spherical sapphire lens that can continuously withstand [pressures of to 6000 psi](#) and [temperatures of up to 350°C](#). It focuses the 782 nm laser beam from the AIO at the tip of the lens to [virtually eliminate errors from focal length variability](#).





## 2 – Approach (Technical)

### Chemometric Data Integration

We have also partnered with [LabKey Software](#) (Seattle, WA) to adapt their [R&D workflow/LIMS database](#) for integration of process, sample, and event data with spectroscopy data to support [chemometric analysis](#).

LabKey CE Server is a free and open-source, web-enabled database server for [secure workflow and data management](#) and [multi-center collaboration](#).



LabKey was spun out of the Hutchinson Cancer Center in Seattle in 2004

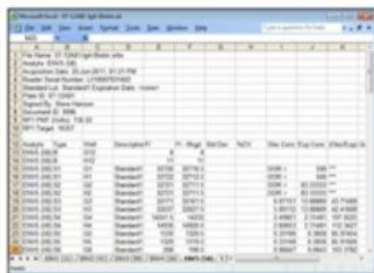
User base includes:

- Novo Nordisk
- Roche
- Merck
- Alder Pharmaceuticals
- **Lygos**
- Johnson & Johnson
- Stanford University
- University of Wisconsin
- University of Washington
- Harvard Medical School
- Genomics England (100,000 Genomes Project)
- ...and many others.

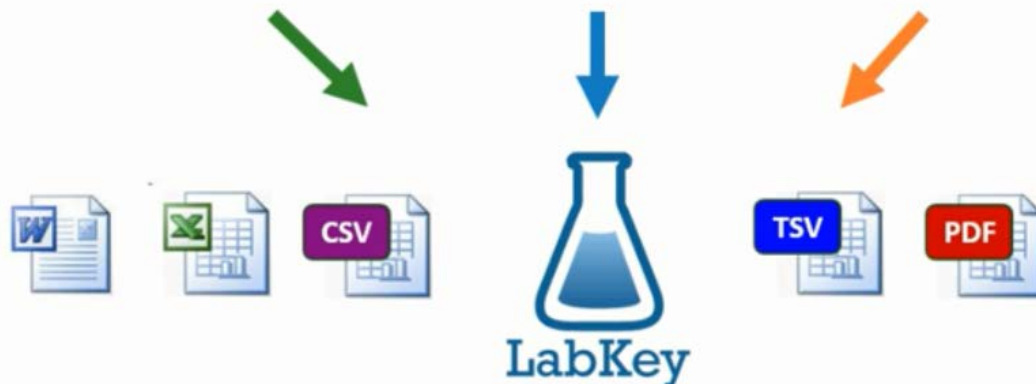
#### Process Sensor & Event Data

Day	Weight (kg)	Temperature (C)
0	86	36.0
79	84	37.0
108	83	37.0
190	80	36.7
246	79	36.8
276	79	37.3
303	79	38.0
335	78	40.0
364	77	39.0
394	75	39.0
0	55	37.0
42	54	37.0

#### HPLC Assay Data & Metadata



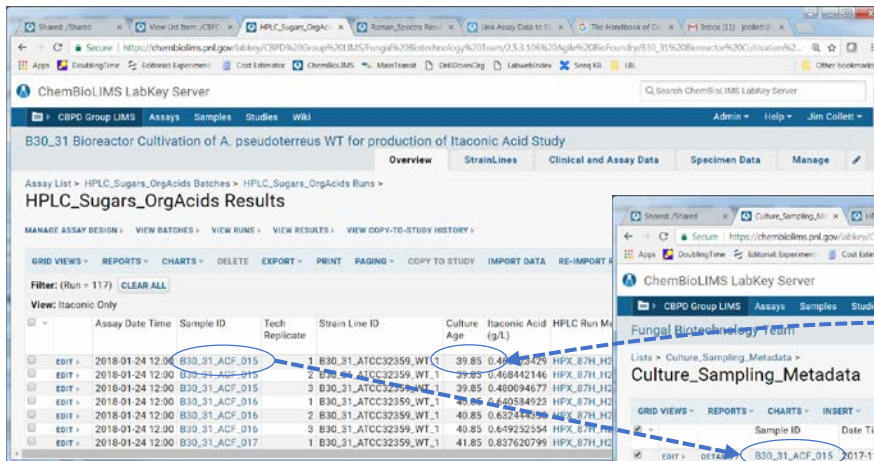
#### Sample Tracking & Metadata

## 2 – Approach (Technical)

### Chemometric Data Integration

Assay data may be easily uploaded into LabKey, and then immediately linked to metadata within a “normalized” database schema that minimizes data redundancy and promotes data integrity.



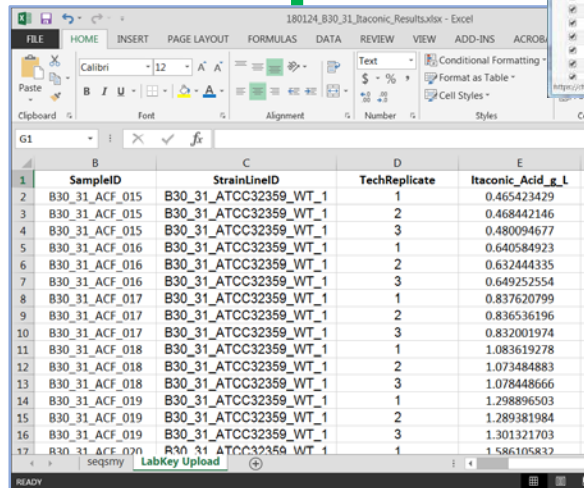
ChemBioLIMS LabKey Server

B30\_31 Bioreactor Cultivation of A. pseudoterrus WT for production of Itaconic Acid Study

Assay List > HPLC\_Sugars\_OrgAcids Batches > HPLC\_Sugars\_OrgAcids Runs > HPLC\_Sugars\_OrgAcids Results

Assay Date Time	Sample ID	Tech Replicate	Strain Line ID	Culture Age	Itaconic Acid (g/L)	HPLC Run Method
2018-01-24 12:00	B30_31_ACF_015	1	B30_31_ATCC32359_WT_1	39.85	0.465423429	HPX_87H_LH2
2018-01-24 12:00	B30_31_ACF_015	2	B30_31_ATCC32359_WT_1	39.85	0.46842146	HPX_87H_LH2
2018-01-24 12:00	B30_31_ACF_015	3	B30_31_ATCC32359_WT_1	39.85	0.480094677	HPX_87H_LH2
2018-01-24 12:00	B30_31_ACF_016	1	B30_31_ATCC32359_WT_1	39.85	0.640584923	HPX_87H_LH2
2018-01-24 12:00	B30_31_ACF_016	2	B30_31_ATCC32359_WT_1	40.85	0.63247799	HPX_87H_LH2
2018-01-24 12:00	B30_31_ACF_016	3	B30_31_ATCC32359_WT_1	40.85	0.649252554	HPX_87H_LH2
2018-01-24 12:00	B30_31_ACF_017	1	B30_31_ATCC32359_WT_1	41.85	0.837620799	HPX_87H_LH2

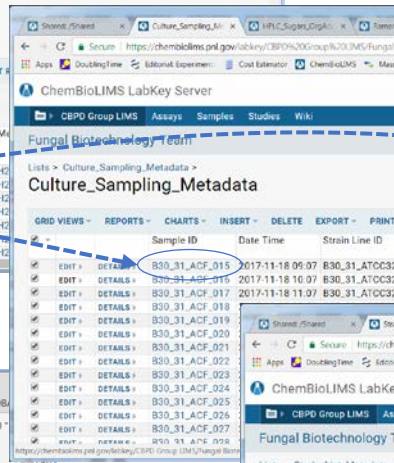
Assay data uploaded to LabKey



180124\_B30\_31\_Itaconic\_Results.xlsx - Excel

SampleID	StrainLineID	TechReplicate	Itaconic_Acid_g_L
B30_31_ACF_015	B30_31_ATCC32359_WT_1	1	0.465423429
B30_31_ACF_015	B30_31_ATCC32359_WT_1	2	0.46842146
B30_31_ACF_015	B30_31_ATCC32359_WT_1	3	0.480094677
B30_31_ACF_016	B30_31_ATCC32359_WT_1	1	0.640584923
B30_31_ACF_016	B30_31_ATCC32359_WT_1	2	0.632444335
B30_31_ACF_016	B30_31_ATCC32359_WT_1	3	0.649252554
B30_31_ACF_017	B30_31_ATCC32359_WT_1	1	0.837620799
B30_31_ACF_017	B30_31_ATCC32359_WT_1	2	0.836536196
B30_31_ACF_017	B30_31_ATCC32359_WT_1	3	0.832001974
B30_31_ACF_018	B30_31_ATCC32359_WT_1	1	1.083619278
B30_31_ACF_018	B30_31_ATCC32359_WT_1	2	1.073484883
B30_31_ACF_018	B30_31_ATCC32359_WT_1	3	1.078448666
B30_31_ACF_019	B30_31_ATCC32359_WT_1	1	1.298896503
B30_31_ACF_019	B30_31_ATCC32359_WT_1	2	1.289381984
B30_31_ACF_019	B30_31_ATCC32359_WT_1	3	1.301321703
B30_31_ACF_020	B30_31_ATCC32359_WT_1	1	1.586105832

LabKey Upload

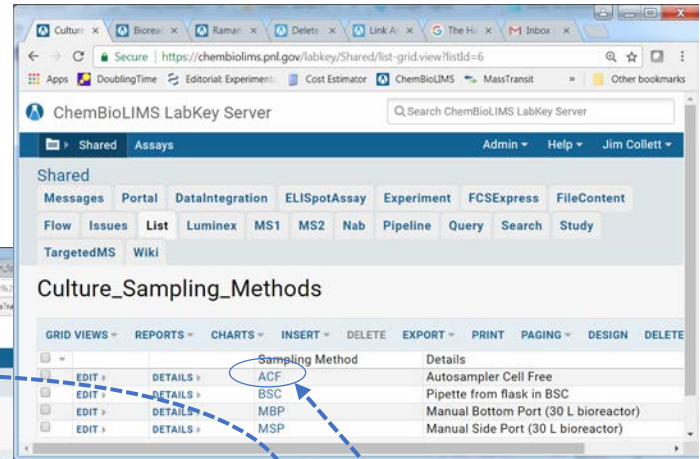


ChemBioLIMS LabKey Server

Fungal Biotechnology Team

Culture\_Sampling\_Metadata

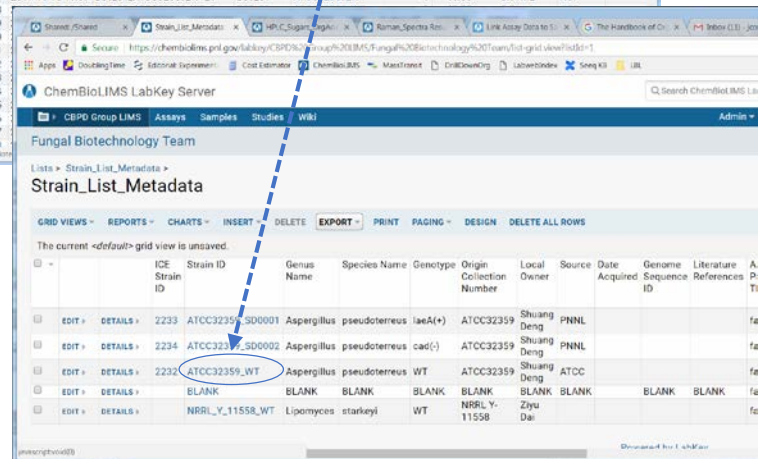
Sample ID	Date Time	Strain Line ID	Experiment	Strain ID	Bio Replicate	Culture Age	Sample Quantity	Sampling Method
B30_31_ACF_015	2017-11-18 09:07	B30_31_ATCC32359_WT_1	B30_31	ATCC32359_WT	1	39.85	5.0 mL	ACF
B30_31_ACF_016	2017-11-18 10:07	B30_31_ATCC32359_WT_1	B30_31	ATCC32359_WT	1	40.84	5.0 mL	ACF
B30_31_ACF_017	2017-11-18 11:07	B30_31_ATCC32359_WT_1	B30_31	ATCC32359_WT	1	41.85	5.0 mL	ACF



ChemBioLIMS LabKey Server

Culture\_Sampling\_Methods

Sampling Method	Details
ACF	Autosampler Cell Free
BSC	Pipette from flask in BSC
MBP	Manual Bottom Port (30 L bioreactor)
MSP	Manual Side Port (30 L bioreactor)



ChemBioLIMS LabKey Server

Fungal Biotechnology Team

Strain\_List\_Metadata

Strain ID	Genus Name	Species Name	Genotype	Origin Collection Number	Local Owner	Source	Date Acquired	Genome ID	Literature References	Annotations
2233	Aspergillus	pseudoterrus	laeA(+)	ATCC32359	Shuang Deng	PNNL				fa
2234	Aspergillus	pseudoterrus	cad(-)	ATCC32359	Shuang Deng	PNNL				fa
2232	Aspergillus	pseudoterrus	WT	ATCC32359	Shuang Deng	ATOC				fa
BLANK	BLANK	BLANK	BLANK	BLANK	BLANK	BLANK				fa
NRRL_Y_11558	Lipomyces	starkeyi	WT	11558	Zyu Dai					fa

### 3 – Technical Accomplishments/Progress/Results

**FY17 Annual Milestone Achieved** Fed-batch control using NIRS maintained predicted glucose levels within 10% of a target value of  $\leq 20$  g/l or less in PCS hydrolysate media with an insoluble lignin concentration of  $\geq 10$  g/l.



Fed-batch bioreactor setup with hydrolysate reservoir on a stir plate to keep lignin solids in suspension.

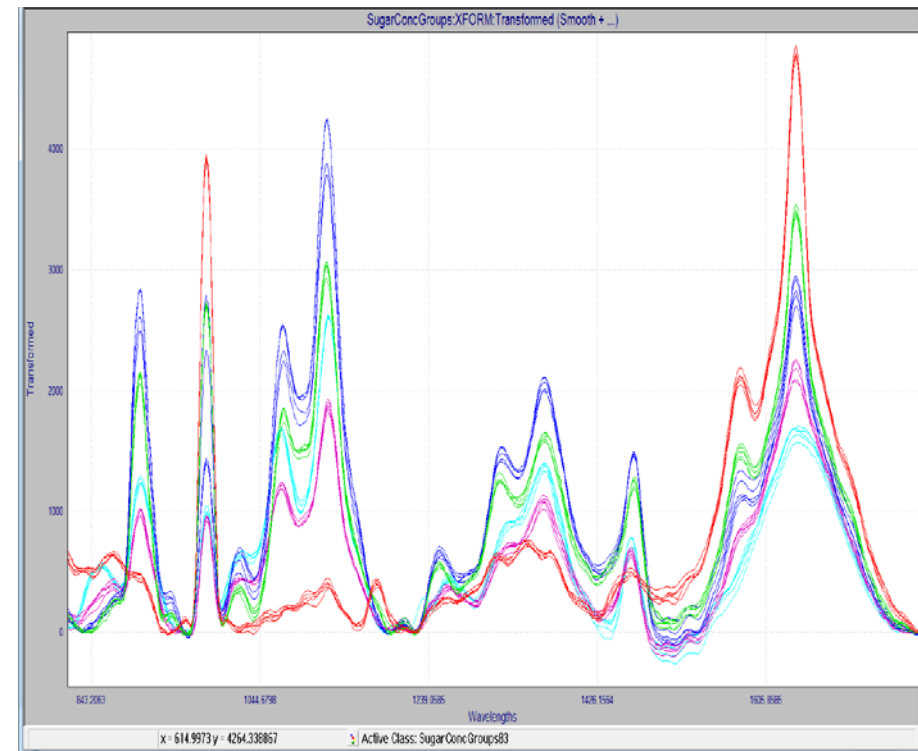
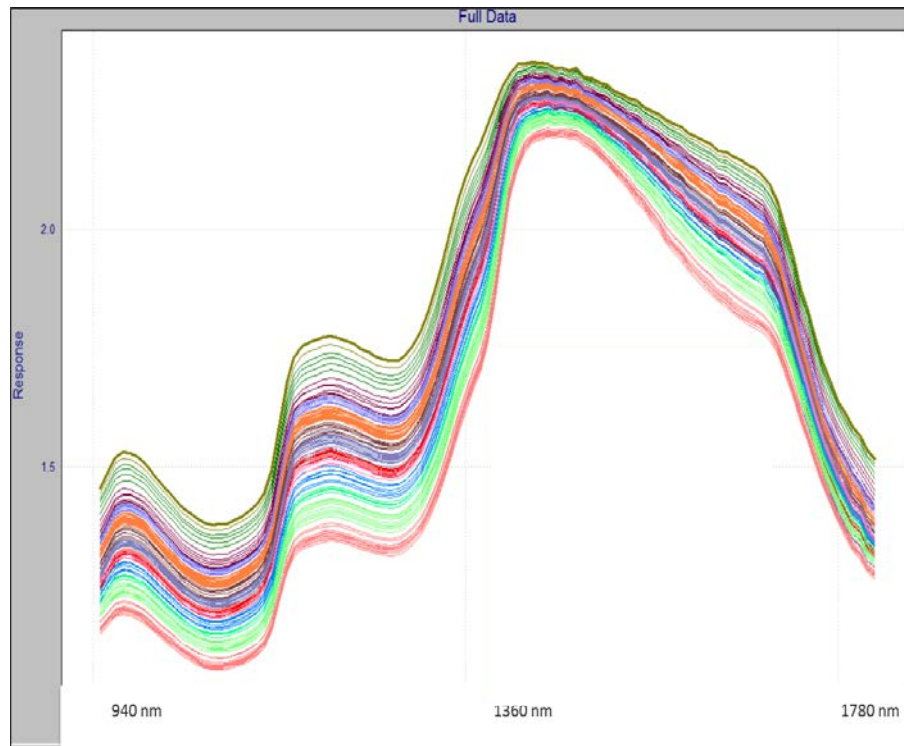




### 3 – Technical Accomplishments/Progress/Results

The high spectral absorbance of water in the NIR range appeared to limit the sensitivity and specificity of sugar concentration prediction in hydrolysates.

Raman spectroscopy is much less susceptible to background interference from water and yielded more obvious feature differences in spectra taken in hydrolysate.



Typical features of NIR spectra taken during fed-batch bioconversion of hydrolysate.

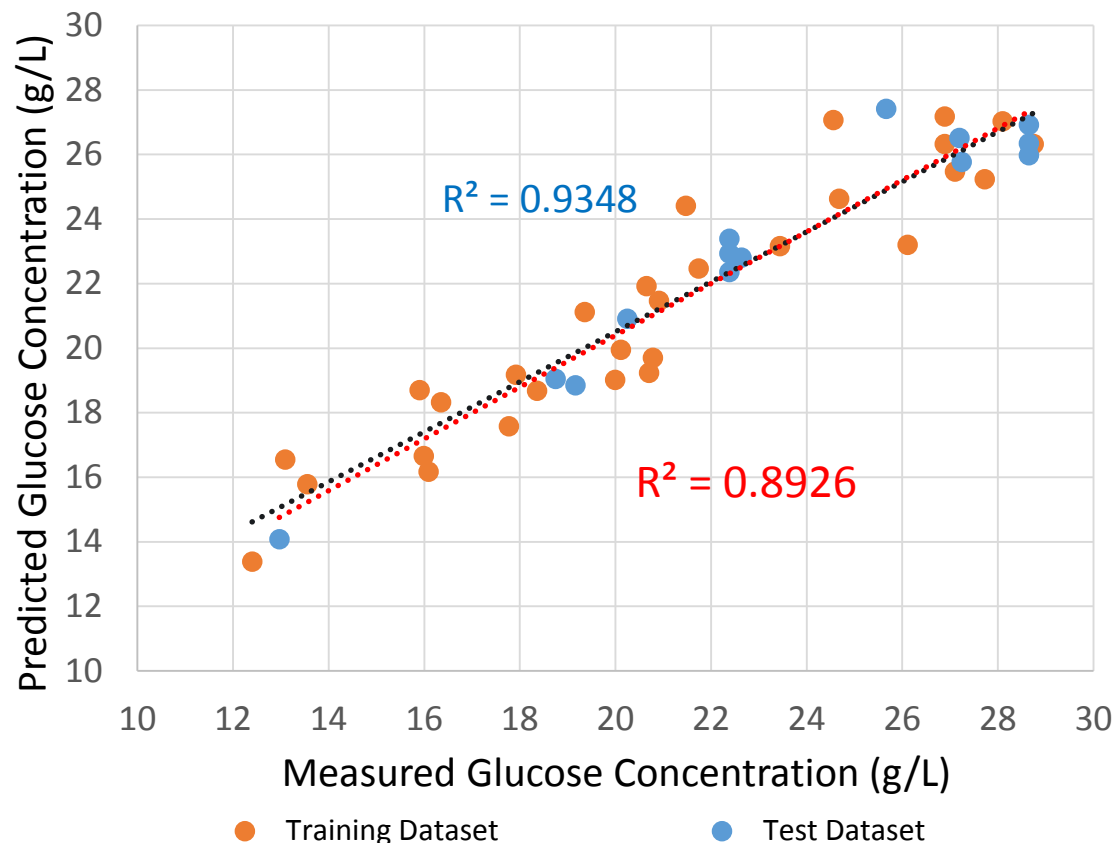
Typical features of Raman spectra taken during fed-batch bioconversion of hydrolysate.



### 3 – Technical Accomplishments/Progress/Results

#### FY18 Annual Milestone Achieved

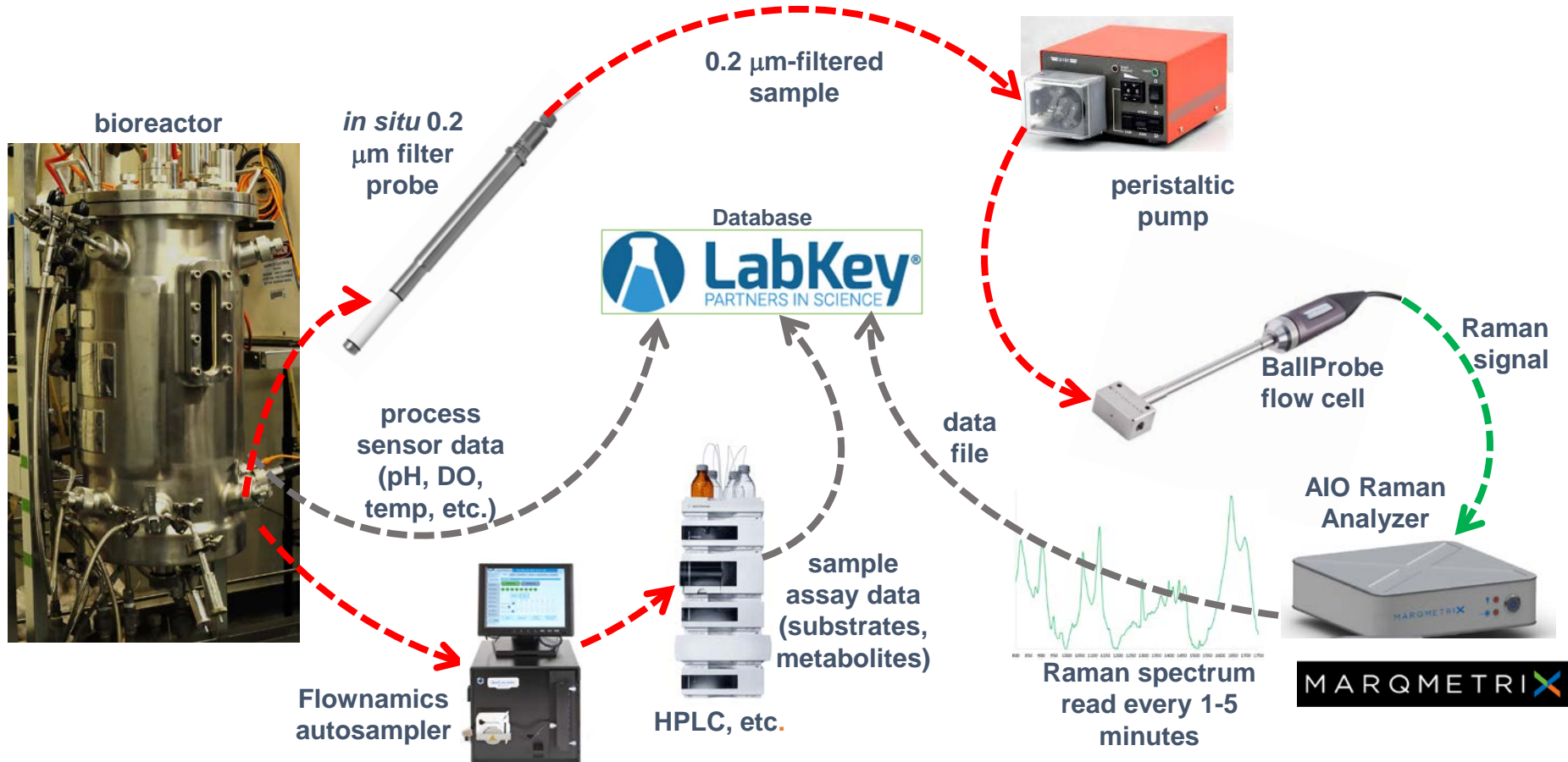
Raman spectral data and HPLC analysis were used to create a partial least squares (PLS) multiple regression model that predicted glucose concentrations over a range from 12.97 to 28.65 g/L with a relative root mean square error of prediction (RMSEP) of 9.15% during cultivation of *Aspergillus pseudoterreus* on deacetylated machine refined (DMR) hydrolysate.



Above, left: 1.75 L bioreactor culture of *A. pseudoterreus* used for Raman PLS model development. Below, left: FlowNamics autosampling system used for drawing 0.2 mm-filtered samples every 30 minutes over 36 hours for Raman and HPLC data acquisition. Above, right: Plot of glucose concentrations measured by HPLC vs. Raman training and test dataset predictions via PLS.

### 3 – Technical Accomplishments/Progress/Results

Automated Raman analysis of hydrolysate media during bioconversion was improved by continuously drawing a small stream of broth from our 30-liter bioreactor and passing it through the MarqMetrix BallProbe flow cell.



# 3 – Technical Accomplishments/Progress/Results

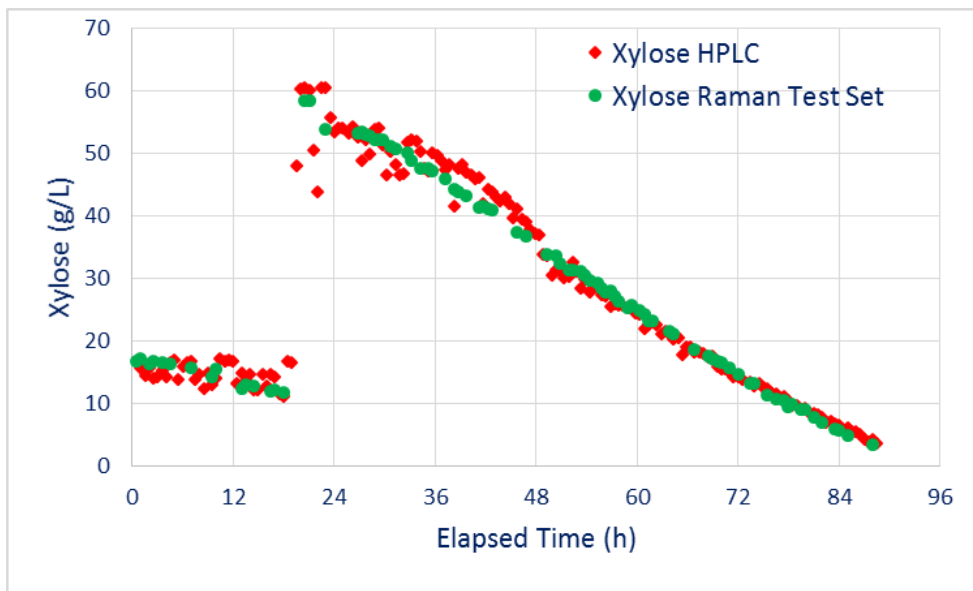
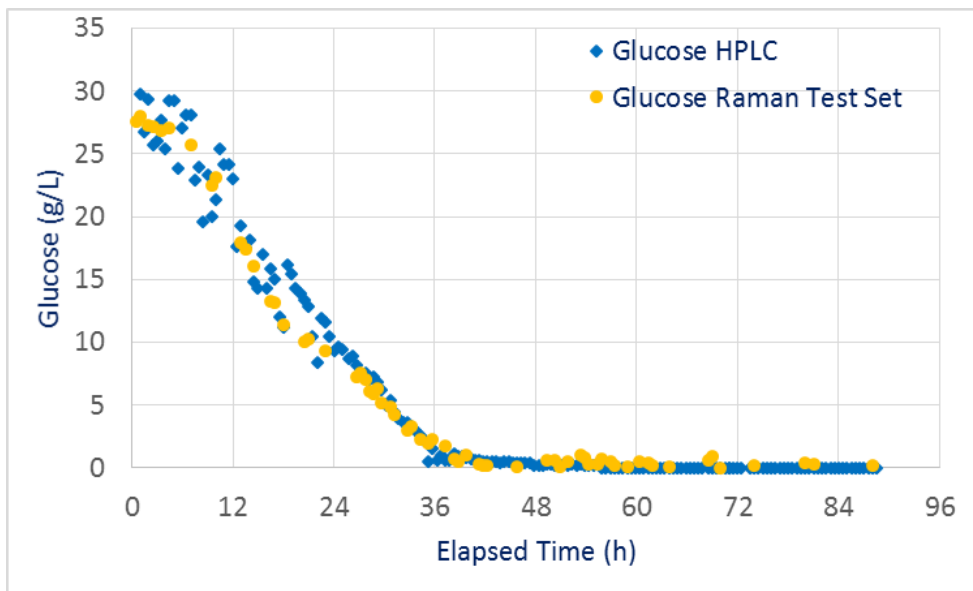


## FY19 Progress

Further refinements in Raman spectroscopy methods [enabled reduction of RMSEP](#) to 3.39% for glucose prediction and 3.77% for xylose prediction, [meeting our FY19 Q1 milestone](#).

Raman calibration models were constructed using ½ of 167 HPLC samples for training, and reserving the other ½ as a validation test set.

The demonstrated level of performance [achieved our Go/No-Go milestone](#) by validating the potential of Raman spectroscopy for enabling a 50% reduction in the frequency of bioreactor culture sampling and manual offline analysis.



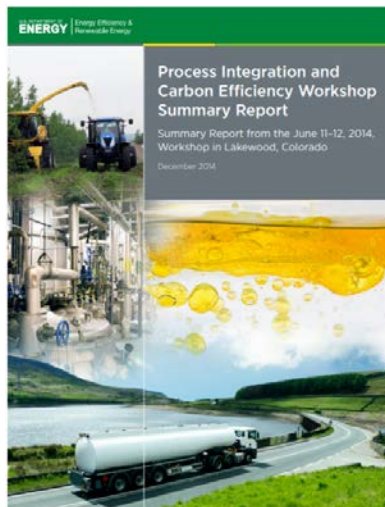
Above, right: Comparison of HPLC measurements to Raman-based PLS predictions of glucose and xylose concentrations in a culture of *A. pseudoterreus* fungi growing in DMR hydrolysate medium in a 30-liter bioreactor (far left image). Above, center: Photograph of an *A. pseudoterreus* grown in defined media shows “pellet” morphology typical of this species when grown in submerged culture.

## Responding to Industry Stakeholder Input



### BETO 2016 Biorefinery Optimization Workshop:

- **Robust control logic** is necessary for integrated operation.
- **Inter-machine communication** to prevent overflows is critical.
- **Well-designed sensors and control systems** are very important; a good control system can improve equipment and biorefinery up-time and system reliability.
- **Provide access to pilot-scale data** that was used to scale-up the process to better understand the reliability of extrapolating pilot data on comparable processes.



### BETO 2014 Process Integration and Carbon Efficiency Workshop:

- **Work with equipment manufacturers** to develop **COTS equipment** for biomass conversion that is efficient and can be **predictably scaled up or down**.
- Determine specifications for process integration, including **development of online monitoring capabilities and analytical tools** that can be used throughout the biorefinery.
- **Tailor technologies to expected particle sizes**.
- Focus optimization on **industrially relevant organisms**.





## 4 – Relevance

Site visits with both established and emerging companies and with other National Laboratories have been conducted for information exchange on PAT for industrial biotechnology.

LanzaTech 

novozymes 

 **IMPACT**  
BIOENERGY

 **NREL**  
NATIONAL RENEWABLE ENERGY LABORATORY

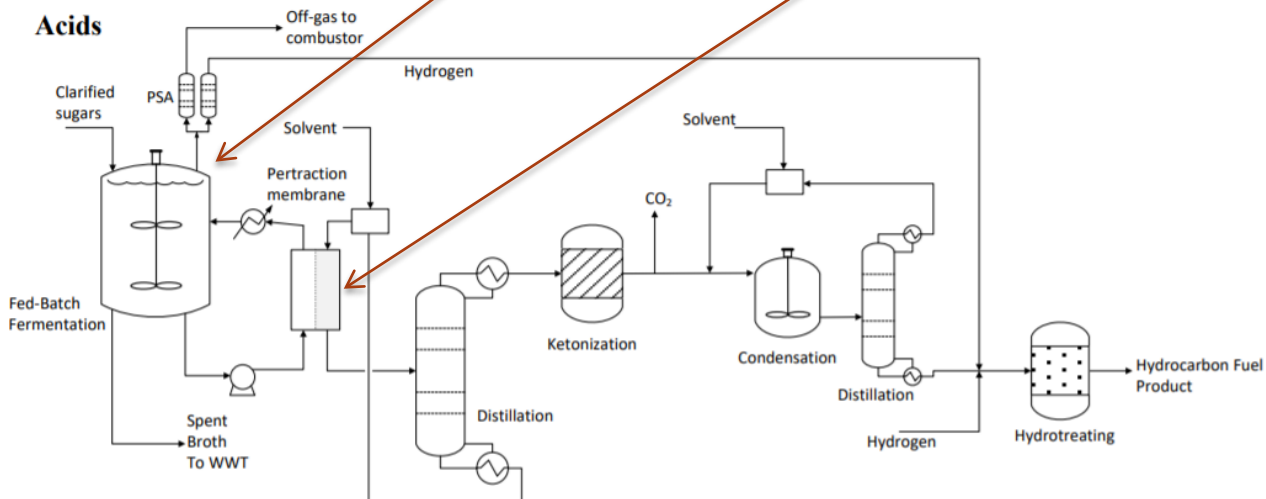


### Related AOP Projects

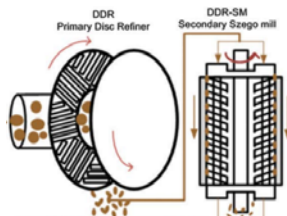
- 2.2.2.501 Fungal Biotechnology
- 2.1.0.301 Analysis and Sustainability Interface
- 2.5.3.106 Agile BioFoundry
- 2.6.1.101 Advanced Biofuels Process Demonstration Unit (ABPDU)
- 2.4.1.100 Bench Scale Integration
- 1.2.2.800 Feedstock Conversion Interface Consortium

# 4 – Relevance

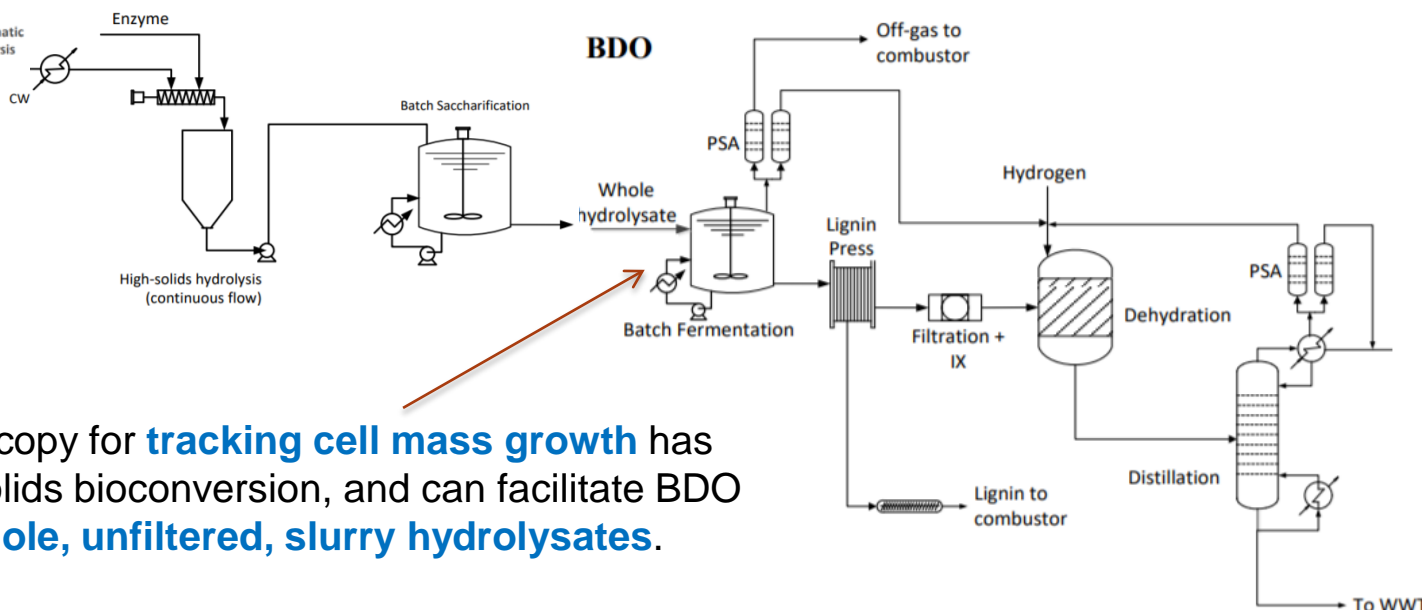
Raman and near-infrared spectroscopy may be useful for **carboxylic acid production** and **downstream separation**.



R&D on **PAT** is relevant to both the fed-batch acid production pathway and the high-solids bioconversion production pathway in the NREL 2030 Target Case.



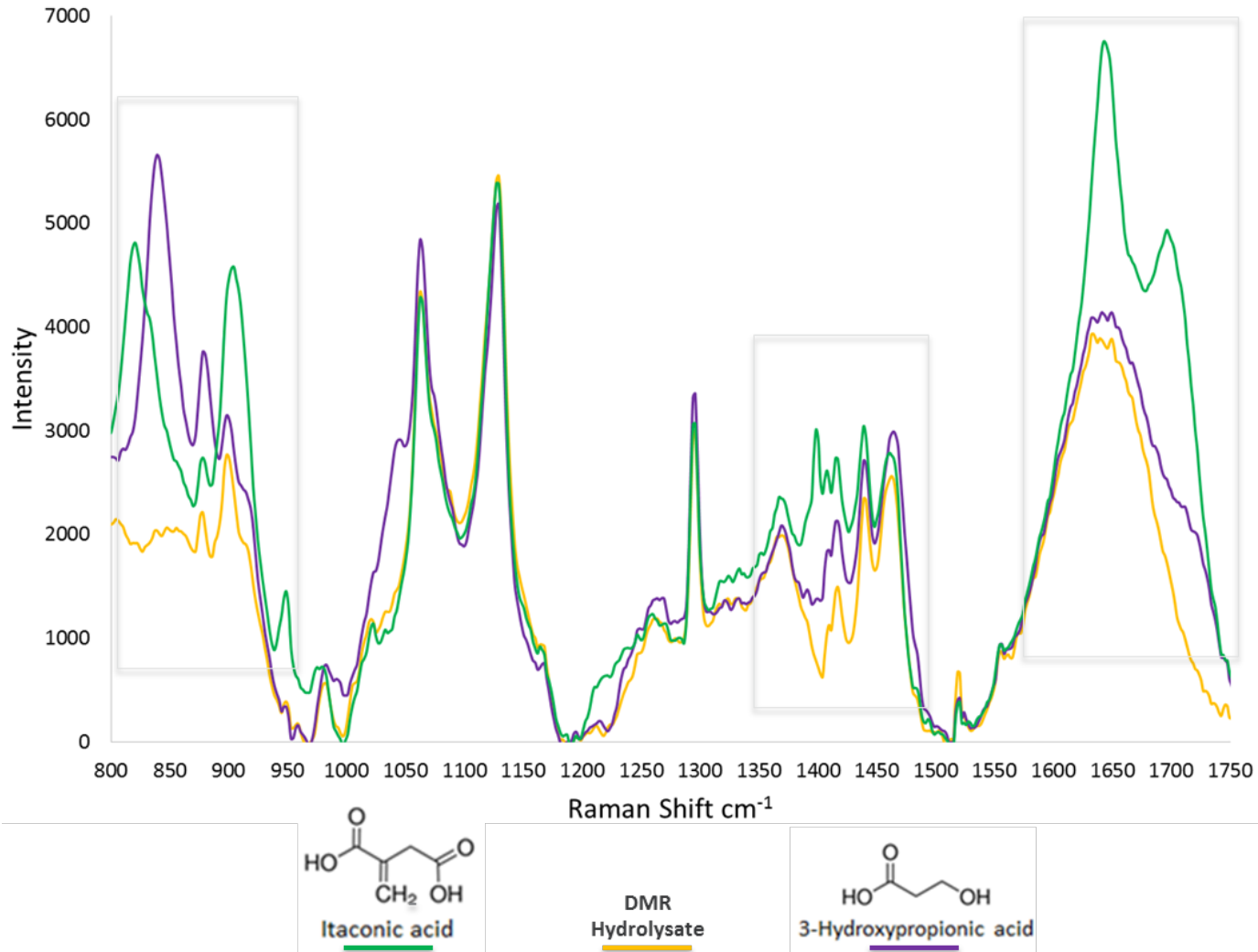
Deacetylation & Machine Refining (DMR)



Dielectric spectroscopy for **tracking cell mass growth** has potential in high-solids bioconversion, and can facilitate BDO fermentation **in whole, unfiltered, slurry hydrolysates**.

## 5 – Future Work

Raman methods for **real-time prediction of chemical bioproducts** will be developed. Preliminary tests of carboxylic acid prediction in DMR hydrolysates have been promising.



Raman spectra of 3 mL samples of itaconic acid and 3-hydroxypropionic acid mixed separately at a concentration of 33.4 g/L into a background solution of DMR-pretreated corn stover hydrolysate (NREL) with a glucose + xylose concentration of 30 g/L. A blank sample containing only DMR hydrolysate was included for comparison. The spectra were baseline corrected using a cubic polynomial fit for easier comparison of bands of interest (gray boxes).

## 5 – Future Work

We will reduce risk for BETO by developing Raman analysis methods for NREL's new **DMR hydrolysate** as well as **conventional DDA hydrolysate**.

**High background fluorescence in DDA** will require mitigation strategies to enable Raman-based prediction of bioreactor broth component levels.

MARQMETRIX MM All in one - AIO-M73-0032

Acquisition Recall Setup

Measurement

Integration time: 100 ms

Averages: 3

Time between samples: 180000 ms

Number of samples: 1

Laser power: 100 mw

Continuous sampling

Dark subtract

Auto dark sample

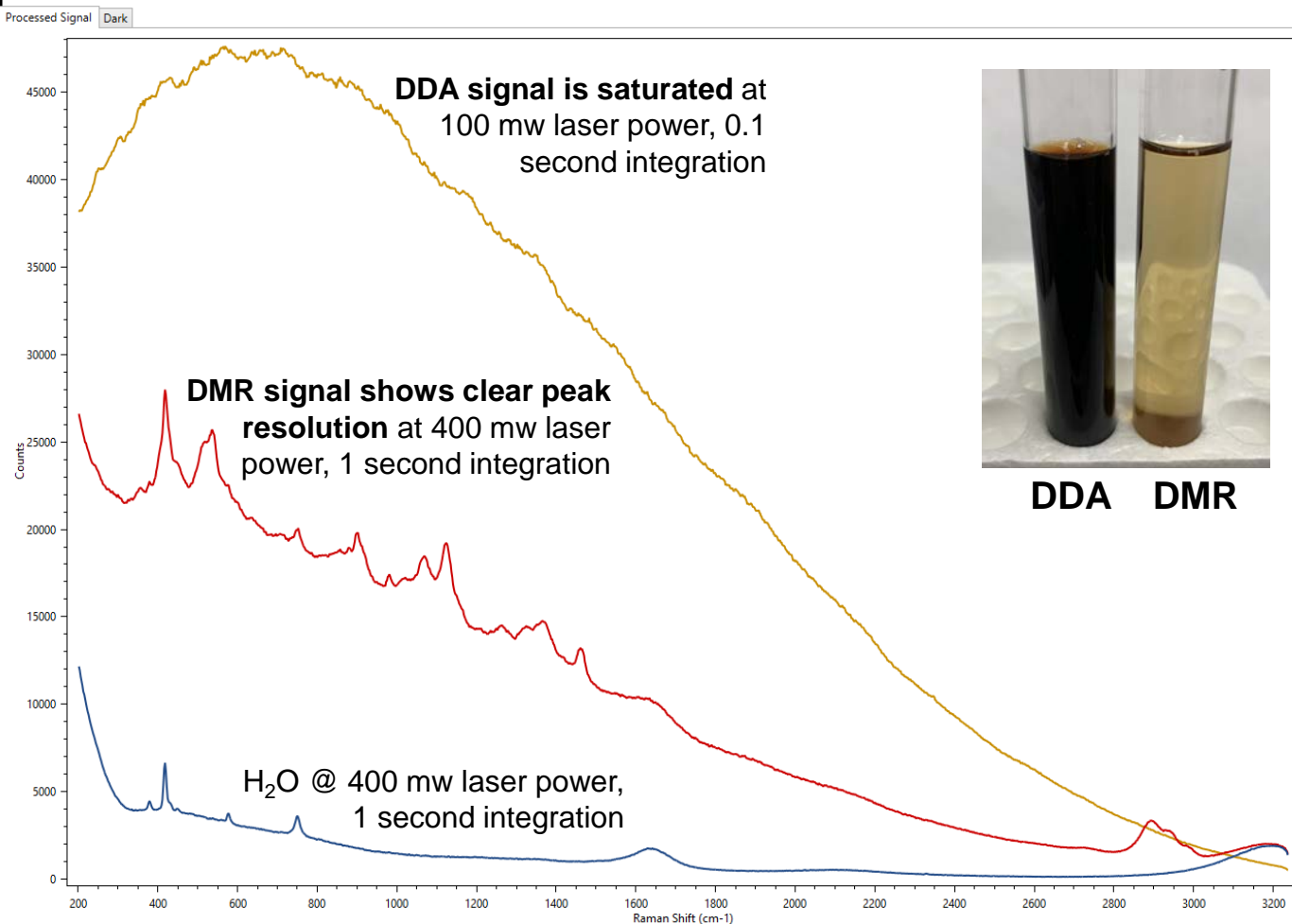
Acquire New Dark Cancel

Data Recording

Path: D:\ Browse

File name: Data Save

Auto Save



**DDA DMR**

*Right:* Raman spectra of 0.2  $\mu\text{m}$  filtered DDA and DMR hydrolysates (glucose + xylose concentration = 125 g/L)

DDA = deacetylated, dilute acid pretreated corn stover



## 5 – Future Work

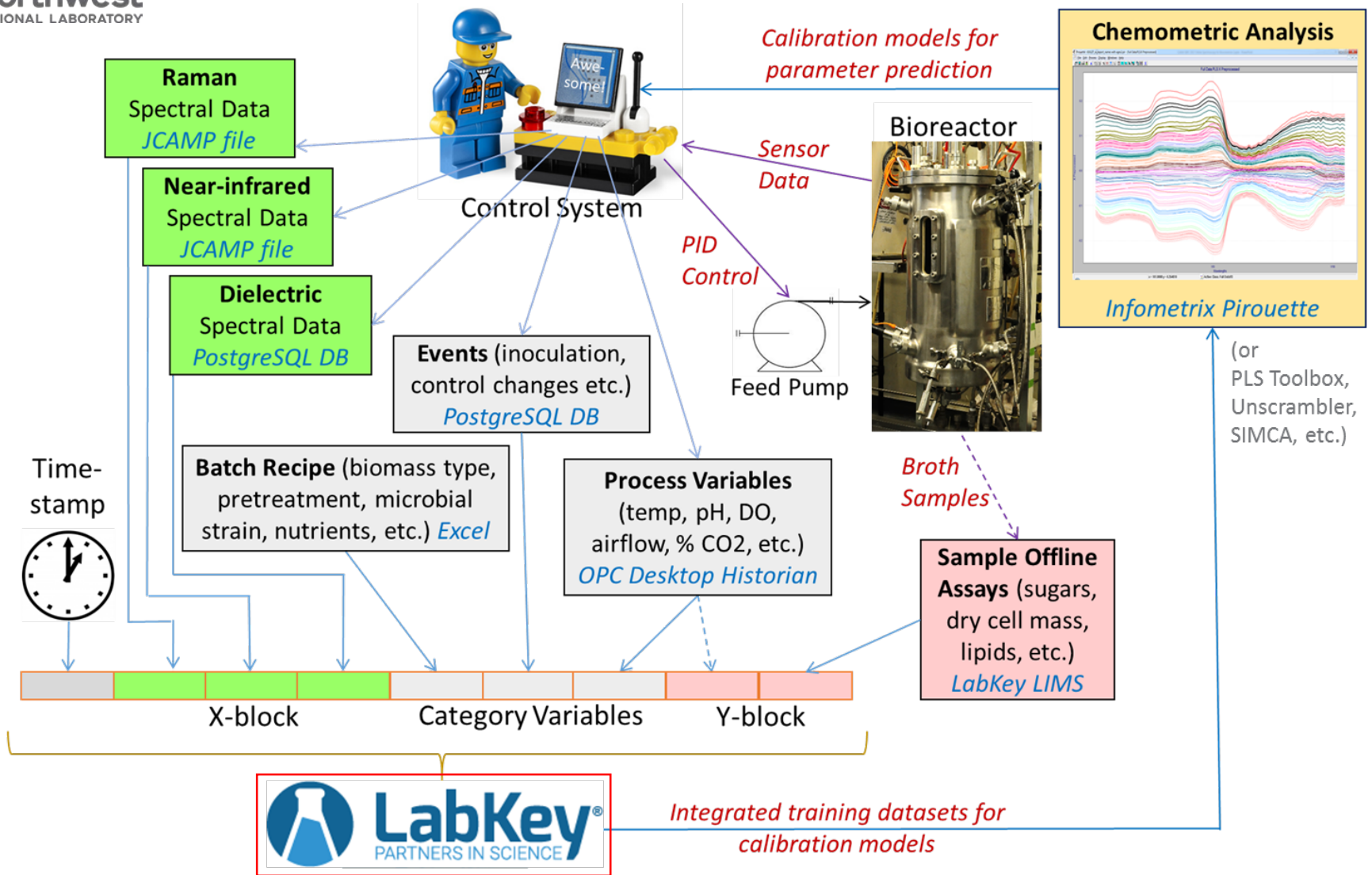
To mitigate background fluorescence in bioreactor broth, we are developing methods for using a **dialysis membrane probe** to only permit compounds with molecular weights below 10kDa to pass through the Raman signal detector.



# 5 – Future Work



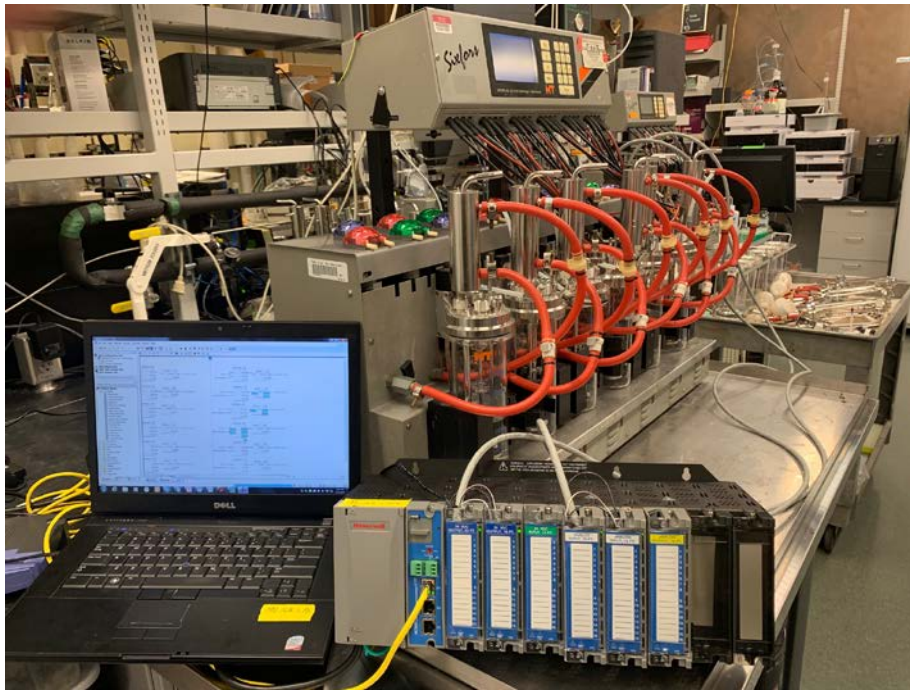
Integration of multiple modes of spectral data with process data within the LabKey database platform will facilitate iterative development of multi-block PLS calibration models for process control systems.



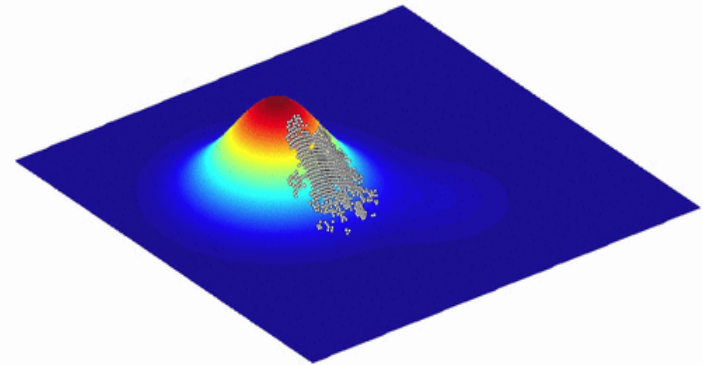
## 5 – Future Work



Integrated spectroscopy methods and the LabKey open source chemometric informatics platform developed via this project will enable creative approaches to **directed evolution of industrial microbes**



Dynamic fitness landscape



Population size,  $N = 2,304$   
Mutation rate,  $\mu = 0.5$  per trait

© Randy Olson and Bjørn Østman

A Sixfors multi-vessel bioreactor array is being retrofitted with an HC900 programmable logic controller to create a “**fitness landscape explorer**” for automatically **improving growth rate and productivity of industrial microbes** across multiple ranges of culture conditions and feedstock compositions.

## 5 – Future Work

**3-Year Project Outcome:** Industrial and R&D deployment of [online Raman, dielectric, and near-infrared spectroscopy tools for bioreactor process control](#), supported by an integrated LIMS and chemometric data analysis environment for bench-to-pilot scale, multi-team collaboration.

**18-month G/NG:** Demonstrate that the accuracy of Raman spectroscopy will result on a 50% reduction in the frequency of bioreactor culture sampling and offline analysis.

### **FY19 Milestones:**

**Q1 QPM:** Use Raman spectroscopy to track xylose with <10% relative standard error of prediction during bioconversion of simulated or actual hydrolysate.

**Q2 QPM :**Use Raman spectroscopy to track a product compound with <10% relative standard error of prediction during bioconversion of simulated or actual hydrolysate.

**Q3 QPM:** Use Raman spectroscopy to maintain a feedstock component within 15% of a set point target of 20 g/L or less for 24 hours during a fed-batch or chemostat bioconversion of a simulated or actual hydrolysate.

**Q4 Annual SMART:** Use Raman spectroscopy to maintain a feedstock component within 15% of a set point target of 10 g/L or less for 24 hours during a fed-batch or chemostat bioconversion of an actual corn stover hydrolysate.





# Summary

**Overview:** The goal of this project is to develop Process Analytical Technologies (PAT) that **enable profitable bioconversion of biomass feedstocks** to advanced biofuels and commodity chemicals..

**Approach:** **Dielectric, near-infrared, and Raman spectroscopy tools and methods** are being developed to support fed-batch and chemostat controls that will scale from the lab up to biorefinery SCADA systems.

## **Technical Accomplishments**

- **FY17 Milestone Achieved:** Demonstrated fed-batch control using near-infrared spectroscopy to maintain glucose levels within 10% of a target value of  $\leq 20$  g/l or less in PCS with an insoluble lignin concentration of  $>10$  g/l.
- **FY18 Milestone Achieved:** Raman spectroscopy enabled prediction of glucose concentrations in an *A. pseudoterreus* culture grown DMR corn stover hydrolysate with a RMSEP of 9.15%.
- **FY19 Q1 Milestone Achieved:** Raman spectroscopy enabled prediction of both glucose and xylose concentrations during bioconversion of DMR hydrolysate with respective RMSEP of 3.39% and 3.77%.
- **FY18-20 Go/No-go Milestone Achieved:** Demonstrated the potential for 50% reduction in manual sampling and analysis

**Relevance -** **This project responds to industry stakeholder interests and reduces risk in BETO R&D** by working with equipment manufacturers to develop efficient, scalable, Process Analytical Technologies for biomass conversion that can be predictably scaled up or down to reliably extrapolate biorefinery unit operation performance from bench and pilot scale data.

## **Future Work**

- Develop Raman methods for **real-time prediction of bioconversion products within bioreactors**.
- Develop **multi-block PLS models** that fuse data from multiple spectroscopy modes to improve accuracy.
- Expand the applicable range of feedstocks and intermediates via **fluorescent background signal mitigation**.
- Design and build **chemometric workflows within LabKey** for higher productivity and efficient classification of sensor and event data to improve training datasets for multiple regression models.
- Develop **PAT-enabled chemostat controls** to support directed evolution of microbial biocatalysts.



# Acknowledgements



BETO Technology Manager: Beau Hoffman



Giora Proskurowski



Ryan Luce



Dan Kopec



Beth Hofstad, Kyle Pomraning, Pavlo Bohutskyi, Ziyu Dai, Diana Rodriguez, Shuang Deng, Ellen Panisko, John Magnuson, Marie Swita, Teresa Lemmon



Dan Schell, Rick Elander, Dave Sievers, Nancy Dowe  
(for hydrolysate supply, advice, and feedback)

## Additional Slides

- Publications and Presentation
- Responses to Previous Reviewers' Comments
- Efficient Workflows in LabKey

## Publications and Presentation

Collett J.R., J.M. Billing, P.A. Meyer, A.J. Schmidt, A.B. Remington, E. Hawley, and B.A. Hofstad, et al. 2019. "Renewable diesel via hydrothermal liquefaction of oleaginous yeast and residual lignin from bioconversion of corn stover." *Applied Energy* 233-234.

Collett, J.R. "Online near infrared and dielectric spectroscopy for real-time tracking of critical process parameters during bioconversion of pretreated corn stover". 2017 Symposium for Biotechnology for Fuels and Chemicals, San Francisco, CA.



# Responses to Previous Reviewers' Comments

## 2017 Peer Reviewer's Comments

The project has addressed an important area that has attracted some attention in industry but is certainly a growing field and currently under-exploited. The team can provide great value to the biorefinery industry (and also further afield) by testing out the validity of the three technologies under study when applied to the complex mixtures seen in cellulosic bioprocess streams.

A further benefit is the software side of this project, testing out tools that can be integrated to bring together the data from online and of at-line monitoring, process conditions and set points, and starting material. Demonstrating a system and advising companies (particularly small companies) on configurations they can build would be a key output of the project. Looking at the other presentations at the review, there are some obvious collaborations within the national laboratory system that could be beneficial here, particularly with the NREL analytical effort (developing a broader chemometric model for near-infrared spectroscopy on feedstocks for instance), and the computational modeling effort also at NREL. Overall, this is a really great idea producing very promising results.

As bioprocesses move from the laboratory to industrial use, successful scale-up is critical. However, understanding and monitoring these larger-scale operations is equally important. This work is developing some nice tools for following processes in real time and providing rapid feedback regarding the operation of a process. The integration of these techniques with laboratory management software and their utility in helping process automation is a real strength. As the project goes forward, better integration with the existing large-scale facilities in the BETO portfolio is recommended.

The objective of this program is to develop complex control systems for fermentations and other unit operations in bioconversion. This is important for second-generation biofuels in particular because of variability in feedstocks and intermediate streams that affect downstream processes. The team has taken an innovative approach, implementing off-the-shelf instrumentation for a particular program (oleaginous yeast fermentation) with the eventual goal of generalizing to any unit operation. Due to support from instrument vendors, this program was operated with relatively low DOE funding. So, the return on investment for BETO was high.

# Responses to Previous Reviewers' Comments

The over- all system will likely not be ready to contribute to 2022 goals, but it may be essential for commercial viability of biorefineries. The use of machine learning to incorporate analytical data and metadata is especially promising. Development will take a long time (in particular, collecting enough data to build the model), so now is the right time to start introducing commercially proven technology to improve monitoring and automated control has significant potential impact on performance and cost. Reducing the need for manual intervention, sampling, and technical expertise may improve efficiency and reproducibility. There are many variables in these bioenergy processes, and reducing a few could have big impact!

This is a great and much-needed approach. The team has realistic and achievable goals: reducing manual sampling, improving process control, minimizing loss runs, and enhancing operation excellence. This is very much desired and can drive many BETO's sponsored laboratories to accelerate their R&D time and improve de-risking strategies. I would heartily recommend integrating this effort across other pilot facilities and, where possible, small-scale and bench-scale facilities as well.

## 1.PI Response to Reviewer Comments

We thank the reviewers for their constructive feed- back and encouragement.

We appreciate the reviewers' multiple recommendations that we share details of our methods for integrating online spectroscopy into our bioreactor operations with the NREL Process Development Unit and the ABPDU, and we hope to assist them in deploying advanced PAT within their pilot-scale operations as well. We have already reached out to these teams and look forward to working with them more closely during the next 3-year cycle of this project.

We will coordinate our further development of chemometric predictive models for biorefinery operations with the NREL Biomass Compositional Unit to take advantage of their well-established LAPs for rapid analysis of biomass feedstocks. Our collaboration could assist future biorefineries with a consistent, plant-wide approach for chemometric analysis and PAT integration. Moreover, we will share with them our progress in deploying the open-source LabKey Laboratory Information Management System database so that they may consider using it for managing their data and analytical workflows as well.



# Additional Slide – Efficient Workflows in LabKey

Simple SQL queries in LabKey enable facile construction of reduced datasets for chemometric analysis from large pools of raw data.

The query on the right matched HPLC sugar concentration data from each of 167 broth samples with the single Raman spectral reading (out of >1600) whose timestamp most closely matched the time at which each sample was drawn from the bioreactor (as shown in the dataset below).

Query Schema Browser / study Schema  
 Edit Raman\_HPLC\_Join\_3 B30\_34\_Apst\_ZD6\_DMR

Source | Data | XML Metadata

Save & Finish | Save | Execute Query | Edit Properties | Edit Metadata | Help ▾

```

1 SELECT DISTINCT Raman_Spectra_ET_StrnID_Date.StrainLineID,
2 Raman_Spectra_ET_StrnID_Date.Read_ID,
3 Raman_Spectra_ET_StrnID_Date.ElapsedTime,
4 HPLC_Sugars_OrgAcids_BPD.SampleID,
5 HPLC_Sugars_OrgAcids_BPD.SampleID.DateTime,
6 HPLC_Sugars_OrgAcids_BPD.SampleID.CultureAge,
7 Raman_Spectra_ET_StrnID_Date.IntegrationTime,
8 Raman_Spectra_ET_StrnID_Date.AverageOf,
9 Raman_Spectra_ET_StrnID_Date.LaserPower,
10 HPLC_Sugars_OrgAcids_BPD.Glucose_g_L,
11 HPLC_Sugars_OrgAcids_BPD.Xylose_g_L,
12 Raman_Spectra_ET_StrnID_Date.Intensities
13 FROM HPLC_Sugars_OrgAcids_BPD
14 LEFT JOIN Raman_Spectra_ET_StrnID_Date ON
15 Raman_Spectra_ET_StrnID_Date.StrainLineID =
16 HPLC_Sugars_OrgAcids_BPD.StrainLineID
17 WHERE abs(Raman_Spectra_ET_StrnID_Date.ElapsedTime
18 -HPLC_Sugars_OrgAcids_BPD.SampleID.CultureAge) <0.025
19
  
```

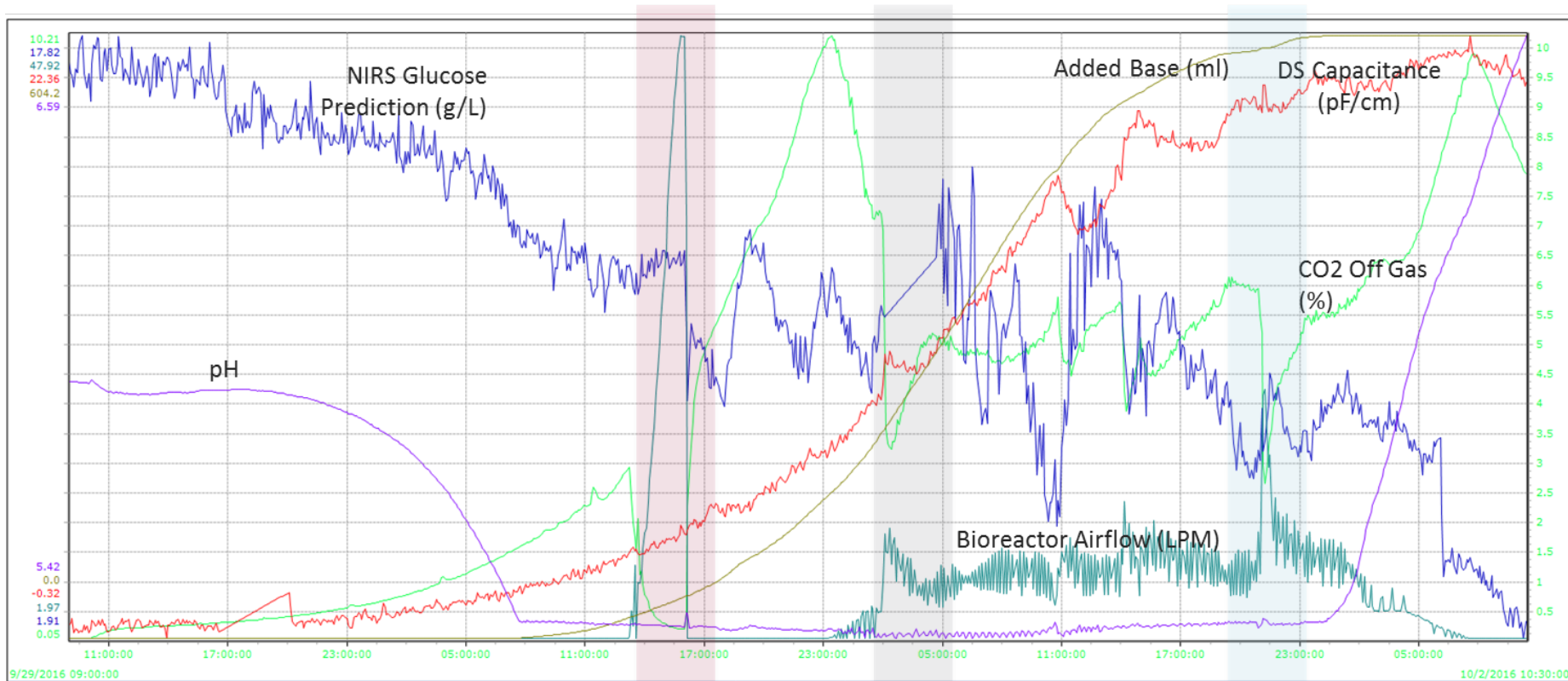
Query Schema Browser / study Schema  
 Edit Raman\_HPLC\_Join\_3 B30\_34\_Apst\_ZD6\_DMR

Source | Data | XML Metadata

Strain Line ID	Read ID	Elapsed Time	Sample ID	Date Time	Culture Age	Integration Time	Average Of	Laser Power	Glucose (g/L)	Xylose (g/L)	Intensities
B30_34_ATCC32359_ZD0006_1	B30.34_104.95	104.950	B30_34_ACF_051	2018-12-09 00:00	104.961	2000.000	3.000	405.000	33.590	12.880	33767.3333,33284.6667,32
B30_34_ATCC32359_ZD0006_1	B30.34_105.48	105.483	B30_34_ACF_052	2018-12-09 00:30	105.461	2000.000	3.000	405.000	29.760	15.650	34260.3333,33944,33425.3
B30_34_ATCC32359_ZD0006_1	B30.34_105.97	105.967	B30_34_ACF_053	2018-12-09 01:00	105.961	2000.000	3.000	405.000	26.760	14.380	34558,34122.3333,33518.3
B30_34_ATCC32359_ZD0006_1	B30.34_106.47	106.467	B30_34_ACF_054	2018-12-09 01:30	106.461	2000.000	3.000	405.000	29.360	15.740	34463.6667,33889.6667,33
B30_34_ATCC32359_ZD0006_1	B30.34_106.95	106.950	B30_34_ACF_055	2018-12-09 02:00	106.961	2000.000	3.000	400.000	25.760	13.910	34598.3333,34148.6667,33
B30_34_ATCC32359_ZD0006_1	B30.34_107.48	107.483	B30_34_ACF_056	2018-12-09 02:30	107.461	2000.000	3.000	405.000	26.060	14.260	34916,34343.3333,33861.3
B30_34_ATCC32359_ZD0006_1	B30.34_107.98	107.983	B30_34_ACF_057	2018-12-09 03:00	107.961	2000.000	3.000	405.000	27.650	15.320	35075.6667,34611.3333,34

## Additional Slide

Integrating sample, process, and event data within LabKey will facilitate classification of training data to improve chemometric model performance.

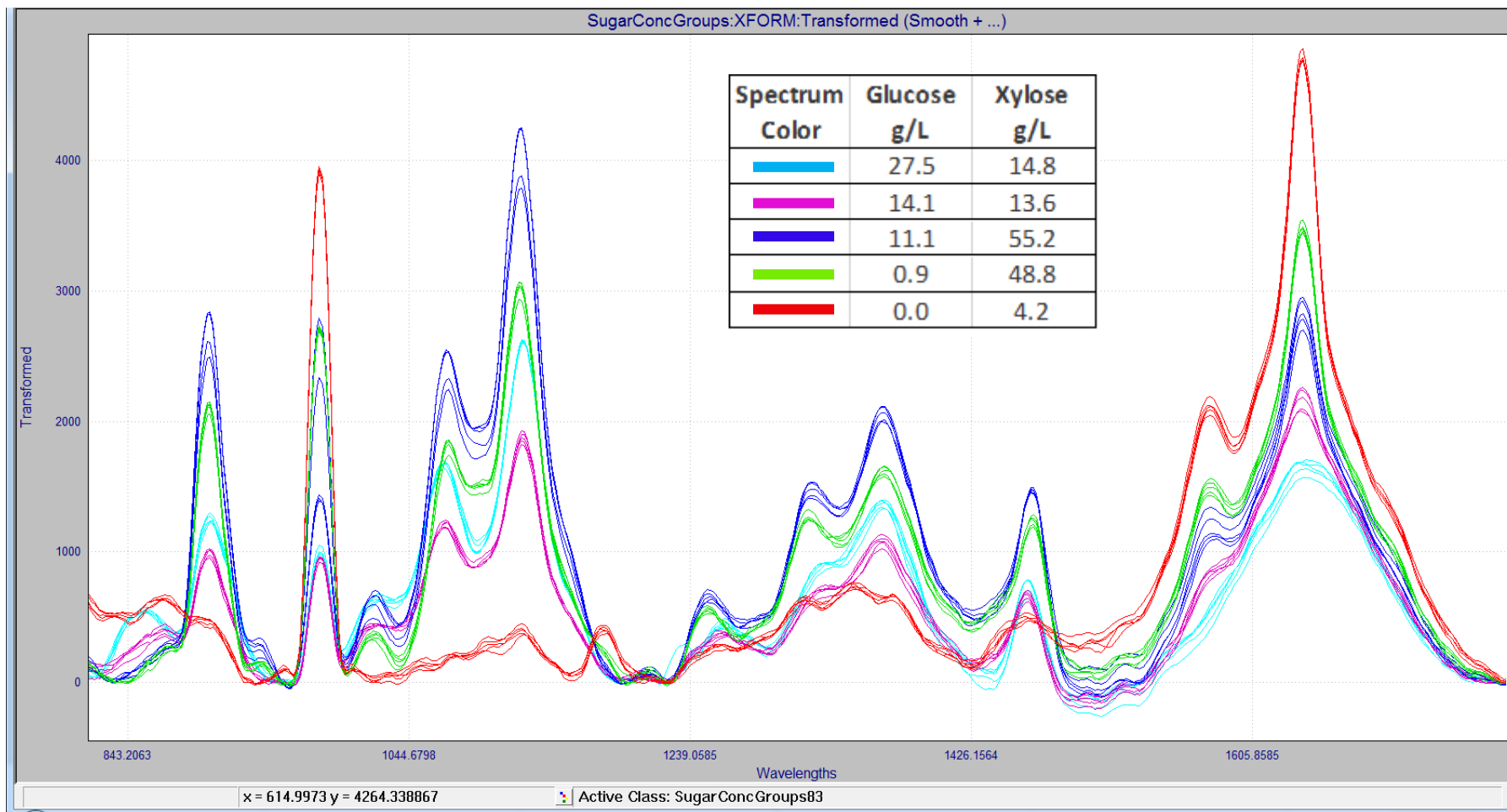


Shaded areas highlight periods of process instability that may be easily flagged within LabKey to scrutinize their effects on chemometric calibration model performance.



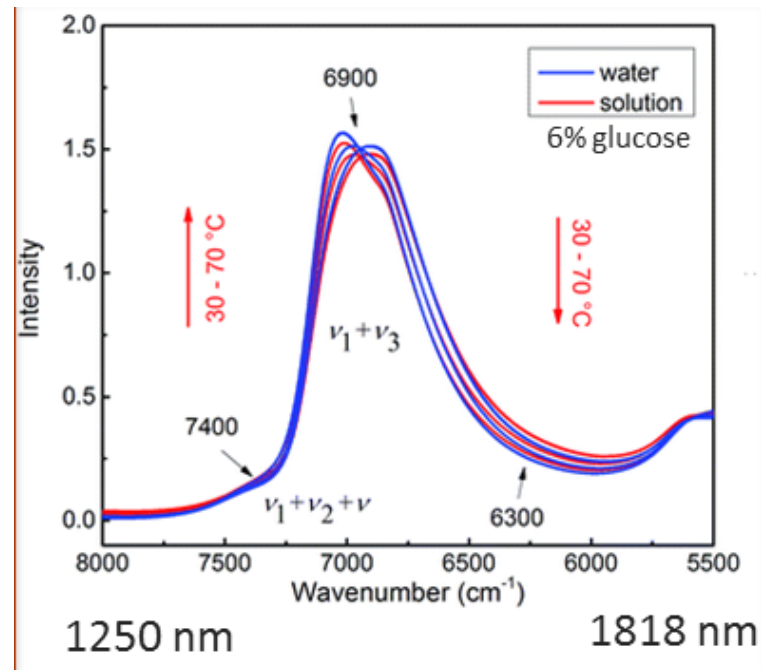
## Additional Slide

Raman spectral peak height changes were easily associated with changes in bioreactor broth component concentrations in an *A. pseudoterreus* culture grown on DMR hydrolysate.



## Additional Slide

The high spectral absorbance of water in the NIR range appeared to limit the sensitivity and specificity of sugar concentration predictions.



NIR spectral features in 6% glucose vs. water are hard to distinguish.

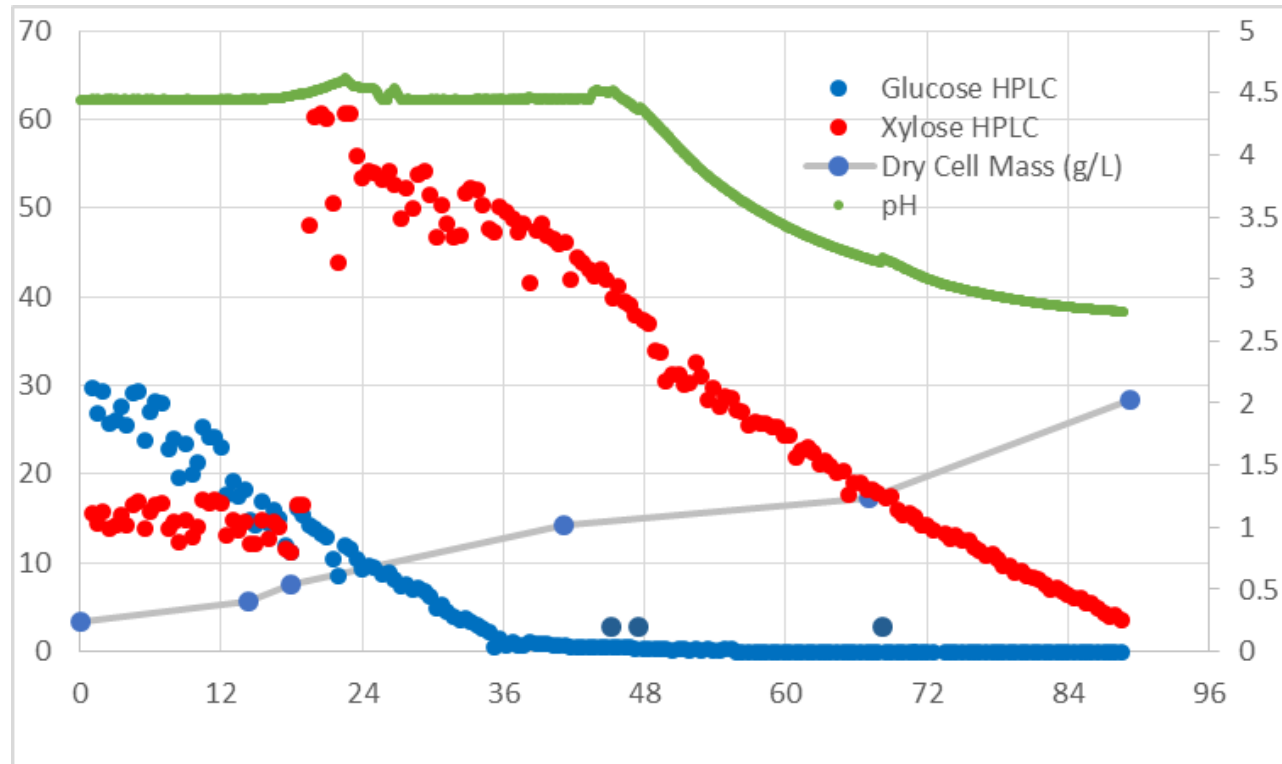
<http://dx.doi.org/10.1039/C6RA18912A>

## Additional Slides



MarqMetrix BallProbe Raman flow cell detector and Flownamics FISP probe for drawing 0.2  $\mu\text{m}$ -filtered bioreactor broth samples.

## Additional Slides



Additional data from bioreactor cultivation of *A. pseudoterreus* in DMR hydrolysate described on Slide 15.