Proceedings of the

# Inter-Agency Conference on Metabolic Engineering

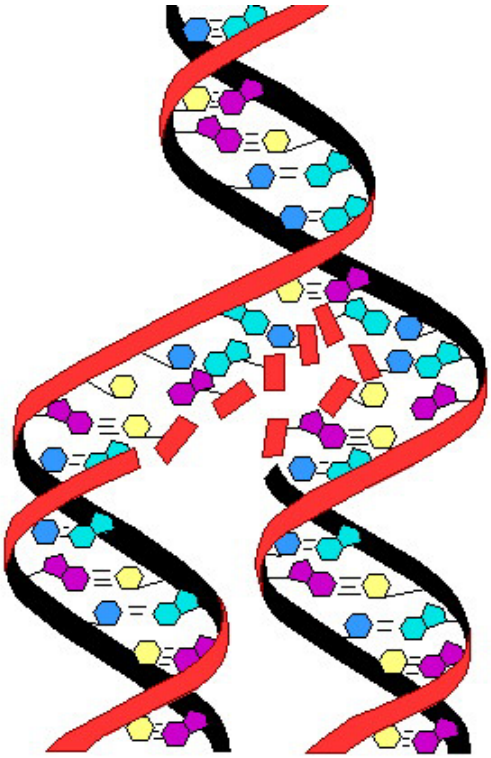
# 2008

---

**Metabolic Engineering Working Group**

February 11, 2008

---



Workshop facilitation, logistics, & report editing and publishing provided by the World Technology Evaluation Center, Inc.

WTEC's mission is to inform U.S. scientists, engineers, and policymakers of global trends in science and technology in a manner that is timely, credible, relevant, efficient and useful. WTEC leads in international assessments of research and development with over 60 such studies. WTEC has been sponsored by most of the Federal research agencies. WTEC is a private, nonprofit corporation. For other WTEC reports, see http://www.wtec.org. For more information contact:

Dr. R. D. Shelton, President
4800 Roland Avenue
Baltimore, MD 21210
Email: shelton@wtec.org
Tel: 717-299-7130

# INTER-AGENCY CONFERENCE ON METABOLIC ENGINEERING 2008

PROCEEDINGS

**February 11, 2008**

**Metabolic Engineering Working Group**

This document was sponsored by the National Science Foundation (NSF) under NSF Grant BES-0508384, awarded to the World Technology Evaluation Center, Inc. (WTEC). The government has certain rights in this material. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Government, the authors' parent institutions, or WTEC.

**NATIONAL SCIENCE FOUNDATION**
4201 WILSON BOULEVARD
ARLINGTON, VIRGINIA 22230

August 1, 2008

Dear Colleague:

We wish to present to you Proceedings from the recent Inter-Agency Conference on Metabolic Engineering, which was held at the Bethesda North Marriott Hotel & Conference Center, North Bethesda, Maryland on February 11, 2008 in conjunction with the Joint Genomics: GTL Contractor-Grantee Workshop VI. As part of the larger workshop, the session on Metabolic Engineering provided presentations which sought to identify opportunities for leveraging genomic technologies for Metabolic Engineering.

This breakout session, conducted jointly between the DOE genomics: GTL program and the MEWG Inter-Agency Conference on Metabolic Engineering 2008, addresses how metabolic engineering facilitates the development of the biorefinery approach to chemical production. Biorefineries are a major focus of the drive to replace fossil fuels with sustainable biologically derived ones. The same biological feedstocks that can be used for fuel production can also be the starting materials for value added chemical products, from acids to polymers. Biorefineries have the potential to produce multiple product lines at the same plant. To do so, however, often requires engineering of the organisms involved in the various production stages. This session, therefore, focuses on work that addresses some of the issues associated with such chemical production, from modification of organisms for fuel and value added chemical production to techniques for measuring and manipulating the necessary metabolic networks. Specific topics for discussion included:

- Foundational Advances in RNA Engineering Applied to Control of Biosynthesis
- Large-Scale Modeling of the Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in *Zymomonas mobilis*
- Development of Tolerant and other Complex Phenotypes for Biofuel Production
- Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant *Synechocystis* PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas
- Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol

Presentations were provided by Grantees of Awards resulting from the activities of the Inter-Agency Metabolic Engineering Working Group (MEWG). The electronic version of this report is available at the following URL: http://www.metabolicengineering.gov/me2008/Report.pdf. This website has color slides information from each of the presentations.

For current information on the activities of the MEWG, please refer to the MEWG web site located at http://www.metabolicengineering.gov/.

We thank you for your interest in Metabolic Engineering and invite your inquiries regarding the Inter-Agency activities.

With best regards, I am

Sincerely yours,
Fred G. Heineken
Chair

# Metabolic Engineering Working Group (MEWG)


***Department of Agriculture***

| | | |
|---|---|---|
| Liang-Shiou Lin | (202) 401-5042 | llin@csrees.usda.gov |
| Gail McLean | (202) 401-6060 | gmclean@csrees.usda.gov |

***Department of Commerce***

| | | |
|---|---|---|
| Robert Goldberg | (301) 975-2584 | robert.goldberg@nist.gov |

***Department of Defense***

| | | |
|---|---|---|
| Walter Kozumbo | (703) 696-7720 | walter.kozumbo@afosr.af.mil |

***Department of Energy***

| | | |
|---|---|---|
| Leslie Pezzullo | (202) 586-1514 | Leslie.Pezzullo@ee.doe.gov |
| Amy Miranda | (202) 586-6471 | amy.miranda@ee.doe.gov |
| Valerie Sarisky-Reed | (202) 586-8014 | valerie.sarisky-reed@ee.doe.gov |
| David Thomassen | (301) 903-9817 | david.thomassen@science.doe.gov |
| Sharlene Weatherwax | (301) 903-6165 | sharlene.weatherwax@science.doe.gov |

***Environmental Protection Agency***

| | | |
|---|---|---|
| Michael Broder | (202) 564-3393 | broder.michael@epa.gov |
| Mark Segal | (202) 564-7644 | segal.mark@epa.gov |

***National Aeronautics and Space Administration***

| | | |
|---|---|---|
| Jitendra Joshi | (703) 599-3674 | jitendra.a.joshi@nasa.gov |

***National Institute of General Medical Sciences***

| | | |
|---|---|---|
| Warren Jones | (301) 594-3827 | jonesw@nigms.nih.gov |

***National Science Foundation***

| | | |
|---|---|---|
| Fred Heineken | (703) 292-7944 | fheineke@nsf.gov |
| Wilfredo Colon | (703) 292-8171 | wcolon@nsf.gov |
| Semahat Demir | (703) 292-7950 | sdemir@nsf.gov |
| Wilson Francisco | (703) 292-5029 | wfrancis@nsf.gov |
| Robyn Hannigan | (703) 292-7163 | rhanniga@nsf.gov |
| John Regalbuto | (703) 292-8320 | jregalbu@nsf.gov |

# Table of Contents

Cover Letter ..... ii

Metabolic Engineering Working Group (MEWG) ..... iii

**1. Background**
Metabolic Engineering ..... 1
Scope ..... 1
Introduction ..... 1
Conference Theme ..... 2

**2. Agenda** ..... 3

**3. Abstracts of Expert Presentations**
Christina D. Smolke ..... 5
Dhinakar S. Kompala ..... 5
E. Terry Papoutsakis ..... 6
Roger L. Ely ..... 7
John Frost ..... 7

**4. Summary of the Panel Session**
Introduction ..... 9
Presentation Summaries ..... 10
General Discussion ..... 18

**5. Agency Activities in Metabolic Engineering**
Department of Agriculture (USDA) ..... 21
Department of Commerce (DOC) ..... 22
Department of Defense (DOD) ..... 22
Department of Energy (DOE) ..... 24
Environmental Protection Agency (EPA) ..... 24
National Aeronautics and Space Administration (NASA) ..... 25
National Institutes of Health (NIGMS/NIH) ..... 26
National Science Foundation (NSF) ..... 26

**6. Workshop Presentations**
Fred Heineken ..... 27
Christina D. Smolke ..... 31
Dhinakar S. Kompala ..... 46
E. Terry Papoutsakis ..... 56
Roger L. Ely ..... 69
John Frost ..... 79

# 1. Background

## Metabolic Engineering

An emerging approach to the understanding and utilization of metabolic processes is Metabolic (or pathway) Engineering (ME). As the name implies, ME is the targeted and purposeful alteration of metabolic pathways found in an organism in order to better understand and utilize cellular pathways for chemical transformation, energy transduction, and supramolecular assembly. ME typically involves the redirection of cellular activities by the rearrangement of the enzymatic, transport, and regulatory functions of the cell through the use of recombinant DNA and other techniques. Much of this effort has focused on microbial organisms, but important work is being done in cell cultures derived from plants, insects, and animals. Since the success of ME hinges on the ability to change host metabolism, its continued development will depend critically on a far more sophisticated knowledge of metabolism than currently exists.

This knowledge includes conceptual and technical approaches necessary to understand the integration and control of genetic, catalytic, and transport processes. While this knowledge will be quite valuable as fundamental research, per se, it will also provide the underpinning for many applications of immediate value.

## Scope

The Metabolic Engineering Working Group is concerned with increasing the science and engineering community's level of knowledge and understanding of ME. The Working Group strives to encourage and coordinate research in ME within academia, industry, and government in order to synergize the Federal investment in ME.

## Introduction

In November 1995, Science Advisor John H. Gibbons of the Office of Science and Technology Policy (OSTP) released the report, "Biotechnology for the 21st Century: New Horizons." This report was a product of the Biotechnology Research Subcommittee (BRS) under OSTP, and identifies priorities for federal investment and specific research opportunities in biotechnology. These priorities include agriculture, the environment, manufacturing and bioprocessing, and marine biotechnology and aquaculture. The BRS formed several working groups to facilitate progress on some of these key priorities. The Metabolic Engineering Working Group (MEWG) was created to foster research in Metabolic Engineering, an endeavor that can contribute to all of the key priorities in the aforementioned report. The Working Group is composed of Federal scientists and engineers who participate as part of the activities of OSTP, and represent all of the major agencies involved in Metabolic Engineering research.

## Conference Theme
## Metabolic Engineering for Biorefineries: Alternative Fuels and Byproducts

The Metabolic Engineering Working Group (MEWG), in pursuit of its goals to promote the advancement of metabolic engineering and coordination of the Federal metabolic engineering research activities for maximum productivity, organized its eighth Inter-Agency Conference held on February 11, 2008.

The goals of any ME program comprise the conceptual and technical approaches necessary to understand integration and control of genetic, catalytic, and transport processes in cellular metabolism. The ability to modify biological pathways extends the fundamental knowledgebase of predictive systems biology towards driving practical applications and sustainable resources for bioenergy solutions.

This breakout session, conducted as part of the jointly-held DOE Genomics: GTL Contractor-Grantee Workshop VI, and the Metabolic Engineering Working Group Inter-Agency Conference on Metabolic Engineering 2008, seeks to identify opportunities for leveraging genomic technologies for metabolic engineering.

In particular, the breakout session addresses how metabolic engineering facilitates the development of the biorefinery approach to chemical production. Biorefineries are a major focus of the drive to replace fossil fuels with sustainable biologically derived ones. The same biological feedstocks that can be used for fuel production can also be the starting materials for value added chemical products, from acids to polymers. Biorefineries have the potential to produce multiple product lines at the same plant. To do so, however, often requires engineering of the organisms involved in the various production stages. This session, therefore, focuses on work that addresses some of the issues associated with such chemical production, from modification of organisms for fuel and value added chemical production to techniques for measuring and manipulating the necessary metabolic networks. Specific topics for discussion included:

- Foundational Advances in RNA Engineering Applied to Control of Biosynthesis
- Large-Scale Modeling of the Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in *Zymomonas mobilis*
- Development of Tolerant and other Complex Phenotypes for Biofuel Production
- Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant *Synechocystis* PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas
- Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol

# 2. Agenda

**Metabolic Engineering Breakout Session**

Tuesday, February 11, 2008, 2:00 to 5:00 pm

Moderator: Fred Heineken (NSF)

Opening Remarks – Fred Heineken (NSF)

Welcome - Gail McLean (USDA)

2:00 pm Foundational Advances in RNA Engineering Applied to Control of Biosynthesis – Christina D. Smolke (California Institute of Technology)

2:25 pm Large-Scale Modeling of the Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in Zymomonas mobilis
- Dhinakar S. Kompala (University of Colorado, Boulder, Colorado)

2:50 pm Development of Tolerant and Other Complex Phenotypes for Biofuel Production
- E. Terry Papoutsakis (University of Delaware)

3:15-3:40 Break

3:40 pm Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant Synechocystis PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas
- Roger L. Ely (Oregon State University)

4:05 pm Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol
- John Frost (Michigan State University)

4:30 pm General Discussion

5:00 pm Adjournment

# 3. Abstracts of Expert Presentations

**Foundational Advances in RNA Engineering Applied to Control of Biosynthesis**

Christina D. Smolke
California Institute of Technology, Pasadena, California

Project Goals: The goals of this project are to: (i) generate RNA-based molecular sensors to key metabolites along the BIA pathway; (ii) engineer *Saccharomyces cerevisiae* to synthesize early BIAs; (iii) integrate the molecular sensors into the transgenic yeast strain for the noninvasive, real-time detection of key metabolite levels along this heterologous pathway; (iv) expand the utility of the engineered molecular sensors to the dynamic regulation of enzyme levels in response to metabolite accumulation; and (v) solidify a strong base in biomolecular and metabolic engineering and encourage its advancement by training and educating scientists through cutting-edge, integrated research and educational plans.

Recent progress in developing frameworks for the construction of RNA devices is enabling rapid advances in cellular engineering applications. These devices provide scalable platforms for the construction of molecular communication and control systems for reporting on, responding to, and controlling intracellular components in living systems. Research that has demonstrated the modularity, portability, and specificity inherent in these molecules for cellular control will be highlighted and its implications for synthetic and systems biology research will be discussed. In addition, tools that translate sequence information to device function to enable the forward design and optimization of new devices will be discussed. The flexibility of the specified framework enables these molecules to be integrated as systems that perform higher-level signal processing based on molecular computation strategies. The application of these molecular devices to studying cellular systems through non-invasive *in vivo* monitoring of biomolecule levels and to regulating cellular behavior will be discussed, in particular in the control and optimization of the biosynthesis of alkaloids in *Saccharomyces cerevisiae.*

**Large-Scale Modeling of the Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in *Zymomonas mobilis***

Dhinakar S. Kompala
University of Colorado, Boulder, Colorado

Project Goals: To investigate the complex network of metabolic pathways in engineered bacterium *Zymomonas mobilis.* To develop a quantitative understanding of the metabolic fluxes along the newly engineered pathways for pentose fermentation. *Zymomonas mobilis* has been engineered with four new enzymes to ferment xylose along with glucose. A network of pentose pathway enzymatic reactions interacting with the native glycolytic Entner Doudoroff

pathway has been hypothesized. We have analyzed the complex interactions between the pentose phosphate and glycolytic pathways in this network by developing a large-scale kinetic model for all the enzymatic reactions. Based on the experimental literature on *in vitro* characterization of each of the 20 enzymatic reactions, the large-scale kinetic model is numerically simulated to predict the dynamics of all the intracellular metabolites along the network of interacting metabolic pathways. This kinetic model takes into account all the feedback and allosteric regulations on the enzymatic reaction rates and is better suited to the systems level analysis of interacting metabolic pathways compared to the standard linearized methods of metabolic flux analysis and metabolic control theory.

This nonlinear kinetic model is simulated to perform numerous *in silico* experiments by varying different enzyme concentrations and predicting their effects on all the intercellular metabolic concentrations and the ethanol production rates in continuous fermentors. Among the five enzymes whose concentrations were varied and given as input to the model, the ethanol concentration in the continuous fermentor was optimized with xyloscisomerase was needed at the highest level, followed by the transaldolase. Predictions of the model, that inter-connecting enzyme phosphoglucose isomerase, does not need to be overexpressed, were recently confirmed through experimental investigations. Through this kinetic modeling approach, we can develop efficient ways of maximizing the fermentation of both glucose and xylose, while minimizing the expression of the heterologous enzymes.

**Development of Tolerant and Other Complex Phenotypes for Biofuel Production**

E. Terry Papoutsakis

Department of Chemical Engineering and the Delaware Biotechnology Institute,
University of Delaware, Newark, Delaware

Project Goals: Develop solvent tolerant phenotypes in prokaryotes for biofuel production. In bioprocessing, in addition to maximizing the flux for a desirable product, the robustness and prolonged productivity of the biocatalyst (the cells) under realistic bioprocessing conditions is an equally important issue. Thus, the ability of cells to withstand “stressful” bioprocessing conditions without loss of productivity is a most significant goal. Such conditions include: toxic substrates accumulation of toxic products and byproducts, high or low pH, or high salt concentrations as encountered inmost applications for the production of chemicals and biofuels as well as in bioremediation applications. The difficulty—but also the intellectual and biotechnological challenge—is that the desirable phenotypic trait is determined by several genes or a complex regulatory circuit. Complex phenotypes are also encountered when one desires to develop a *de novo* capability or pathway in a particular cell type. For example, how do cells put together the regulatory elements of a sequence of genes to make a pathway or program possible? Yes, it is an evolutionary process, but if we are to "imitate" the process, what would we do? What tools could one possibly use and strategies to facilitate the development of complex phenotypes in microbial cells? From omics-based analysis to synthesis, all selection based, or hybrid? Knowledge-based and mechanistic or not? This will be the focus of this presentation, together with some data from early efforts to demonstrate some key concepts that we explore in my laboratory.

## Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant *Synechocystis* PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas

Roger L. Ely

Department of Biological & Ecological Engineering, Oregon State University, Corvallis, OR 97331

This talk will describe results from an ongoing GTL project in which we are using the cyanobacterial species *Synechocystis* PCC 6803 to address two main factors affecting $H_2$ production in PCC 6803: NADPH availability and O2 sensitivity. $H_2$ production in PCC 6803 requires that the NADP pool be highly reduced, which can be problematic because several metabolic pathways potentially can act to raise or lower NADPH levels. Also, the [NiFe]-hydrogenase ($H_2$ase) in PCC 6803 is reversibly inactivated at very low O2 levels due to binding of O2 at the active site. Largely because of this O2 sensitivity and the requirement for high NADPH levels, much of the overall $H_2$ production occurs under anoxic conditions in the dark, supported by breakdown of glycogen or other organic substrates accumulated during photosynthesis. Also, other factors, such as N or S limitation, pH changes, presence of other substances, or deletion of particular respiratory components, can affect light or dark $H_2$ production. Therefore, we have used $H_2$ production profiling and metabolic flux analysis to examine light and dark $H_2$ production under a number of culture conditions with wild-type (WT) PCC 6803 cells and with mutant strains. Also, some of the mutants we have created have shown themselves capable of increased $H_2$ production. Specific project tasks are as follows:

1. Evaluate the effects of various culture conditions (N, S, or P limitation; light/dark; pH; exogenous organic carbon) on $H_2$ production profiles;
2. Conduct metabolic flux analyses for enhanced $H_2$ production profiles using selected culture conditions and inhibitors of specific pathways;
3. Create PCC 6803 mutant strains with modified $H_2$ases exhibiting increased O2 tolerance and greater $H_2$ production;
4. Integrate enhanced $H_2$ase mutants and culture and metabolic factor studies to maximize 24-hour $H_2$ production.

## Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol

John Frost

Michigan State University / Frost Chemical Laboratory

1,2,4-Butanetriol trinitrate (BTTN) is manufactured by the nitration of 1,2,4-butanetriol (BT). The challenges associated with chemical synthesis of BT will be discussed along with the creation of a biosynthetic pathway that allows a single microbe to catalyze the conversion of D-xylose into D-BT. Central to this created pathway is the discovery of the ability of *Escherichia coli* to catabolize D-xylonic acid and the role that the enzyme D-xylonate dehydratase plays in this catabolism. The BT biosynthetic pathway was assembled in an *E. coli* host and begins with oxidation of D-xylose to D-xylonic acid. D-

Xylonate dehydrogenase, which is heterologously expressed in an *E. coli* host from the *Caulobacter crescentus xdh* locus, is recruited for this purpose. Two xylonate dehydratases encoded by *xjhG* and *yagF* loci, which were discovered to be native to *E. coli*, catalyze the conversion of D-xylonic acid into 3-deoxy-D-*glycero*-pentulosonic acid. Decarboxylation of 3-deoxy-D-*glycero*-pentulosonic acid to form 3,4-dihydroxy-D-butanal is mediated by heterologously expressed *mdlC* isolated from *Pseudomonas putida*. Final reduction of 3,4-dihydroxy-D-butanal to BT is catalyzed by an alcohol dehydrogenase native to the BT-synthesizing *E. coli*. BTTN is more stable than nitroglycerin and mixes effectively in a solvent-free process with nitrocellulose. These characteristics make BTTN an ideal replacement for nitroglycerin and a useful plasticizer in single-stage rocket motors.

# 4. Summary of the Panel Session

Gail McLean, USDA

**Introduction**

Dr. Fred Heineken (NSF) welcomed participants to the session and provided an overview of the Metabolic Engineering Working Group (MEWG) and its activities. He listed the eight participating agencies, which are the Department of Agriculture (USDA), Department of Defense (DOD), Department of Energy (DOE), Environmental Protection Agency (EPA), National Aeronautics Space Administration (NASA), National Institute of General Medical Sciences (NIGMS, NIH), National Institute of Standards and Technology (NIST), and National Science Foundation (NSF), and provided the goals of MEWG:

- Promote the Advancement of Metabolic Engineering
- Coordinate Federal Metabolic Engineering Research
- Identify and Address Gaps in Metabolic Engineering Research

Dr. Heineken then discussed the accomplishments of the MEWG over the past 13 years and reviewed the four recent topics of interagency interest:

- Instrumentation, Tools, and Methods to facilitate the Study of Metabolic Pathways in Cells
- Quantitative and Conceptual Models
- Bioinformatics
- Engineering of Metabolic Pathways for a Desired Output

The MEWG has released six interagency announcements and reviewed a total of 236 proposals. The 54 awards to date have totaled $34 million. Dr. Heineken noted that a call for new research proposals is anticipated for this spring, with awards potentially announced in the fall and funding in the winter. (The new announcement is now posted at: metabolicengineering.gov/1announce.html.) MEWG will continue to sponsor an annual interagency conference on metabolic engineering and maintain the interagency Web site.

Dr. Gail McLean of the USDA also extended a welcome to session participants from the USDA and its partners. She noted that the U.S. is facing many challenges in regard to energy production and use. With the increased efforts to develop and improve biofuels, there is particular interest in research relating to biorefineries including how metabolic engineering could contribute to the production of bioenergy. Metabolic Engineering will play a critical role in helping the U.S. to respond to the need for new energy sources and new methods for energy production. The topic of this year's session, **Metabolic Engineering for Biorefineries: Alternative Fuels and Bioproducts**, is thus very timely.

Dr. Heineken moderated the session which consisted of five speaker presentations followed by general audience discussion. The speaker presentations were:

- "Foundational Advances in RNA Engineering Applied to Control of Biosynthesis"
  Dr. Christina Smolke, California Institute of Technology
- "Large-Scale Modeling of the Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in *Zymomonas mobilis*"
  Dr. Dhinakar Kompala, University of Colorado, Boulder
- "Development of Tolerant and other Complex Phenotypes for Biofuel Production"
  Dr. Terry Papoutsakis, University of Delaware
- "Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant *Synechocystis* PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas"
  Dr. Roger Ely, Oregon State University
- "Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol"
  Dr. John Frost, Michigan State University/Frost Chemical Laboratory

## PRESENTATION SUMMARIES

### "Foundational Advances in RNA Engineering Applied to Control of Biosynthesis"

Christina D. Smolke
California Institute of Technology

The first presentation addressed a fundamental aspect of metabolic engineering, the study of complex gene regulatory networks and development of gene expression technology. Dr. Smolke focused on RNA engineering and the control of biosynthesis and how this knowledge can be applied to challenges of metabolic engineering. She noted that feedback control plays an important role in biosynthesis. Through layers of networks, both functional and control, the cell has evolved mechanisms and responses important for metabolism and cellular processes. Regulation mediated by a control network consists of sensor to computation to actuator functions (the "controller") which then make up the system (the "process"). Metabolic engineering to this point has largely ignored building in this level of regulation.

Engineered biology systems have been applied to many products: biofuels such as isobutanol, drugs such as artemisinin, and scents and flavors such as isoamyl acetate. Dr. Smolke's presentation focused on the engineering of benzylisoquinoline alkaloid (BIA) biosynthesis and how her research is attempting to construct the different components and products of the pathway in yeast. There are no natural sources for intermediates in the BIA pathway and many of the conversion steps and enzymes are not characterized or even identified, thus complicating reconstruction of this pathway. Consequently, yeast is being used for studies of the pathway in general as well as for reconstruction of the pathway. The ability to reconstruct such a pathway in yeast could result in a broad range of applications such as anti-malarial drugs.

Dr. Smolke's laboratory constructed in yeast a synthetic BIA pathway, in some cases using enzymes from other hosts. Early and Late Synthetic BIA pathways were constructed separately, characterized, and then combined together. The MEWG session presentation focused on the Late Synthetic BIA pathway. Intermediates such as reticuline were able to

be produced in yeast, in the case of reticuline by expressing three methyl transferases. Once they were able to see production of the desired product, such as reticuline, in the engineered yeast strain, experiments then focused on ways to optimize production. Research showed that the combination of enzymes makes a difference in conversion. The expression of enzymes in regard to what product is needed is also important. For example, maximal expression of the product could be reach at approximately 0.5% glucose. At higher levels of glucose, the enzymes were more highly expressed but were not making more product. Studies also showed that enzymes had to be partnered properly. An Arabidopsis enzyme may not necessarily work with the yeast partner and may require its Arabidopsis partner. Tools for optimizing BIA production involved layering controls on top of the pathway. The types of controls or tools included regulating flux around branch points, noninvasive sensing of metabolite production, and temporal regulation of enzyme cascades.

General biological control systems include RNA with secondary structure. A ribozyme switch platform can be developed to up- or down-regulate expression. RNA devices can also be integrated as noninvasive sensors of metabolite concentration. The RNA can serve all three functions of control, as actuator to affect responses such as transcription or translation, for computation to define and control action, and as sensor to receive information. Another challenge for control is scalability and specificity of nucleic acid sensors and aptamers.

Dr. Smolke's research shows that metabolic engineering requires a host of tools for optimizing flux and product accumulation. Flux through pathways can be controlled dynamically through user-programmed feedback control systems. The development of new genetically-encoded tools for receiving, processing, and transmitting molecular information is an important part of flux control in metabolic engineering. These tools will allow response properties to be programmed to fit performance specifications of a given application and will aid development of more robust cellular systems for metabolic engineering.

Dr. Smolke was asked how important it was to have transporter activity for the molecules produced so they are transported out of the cell. She responded that it depends on the pathway. They did not specifically research the transporters in yeast. They did find that BIA does get transported out of the cell and did not notice any toxicity on the cells when it was exported into the media.

**"Large-Scale Modeling of the Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in *Zymomonas mobilis*"**

Dhinakar S. Kompala
University of Colorado, Boulder

In the session's second presentation, Dr. Kompala described research to optimize engineered pentose fermentation in *Zymomonas mobilis*. This microbe is unlike *Escherichia coli* and yeast in that it doesn't ferment pentose to ethanol. Dr. Kompala

described how *Zymomonas mobilis* has been engineered with new enzymes to ferment the pentose sugar xylose along with glucose and hypothesized a network of pentose pathway enzymatic reactions. To test for this proposed
reaction network, they took a unique approach and developed a kinetic model for all the enzymatic reactions of the pentose phosphate and glycolytic pathways. They utilized kinetic data on different sugar metabolism rates and enzymatic activity data available in the literature to refine the parameters of the model and to validate the proposed reaction network. The objectives of their research are to investigate the assumed network of enzymatic reactions linking the pentose metabolism and glycolysis pathways; incorporate non-linear rate expressions for the feedback regulation of enzymatic reactions; and find an optimum combination of enzymes for maximizing ethanol concentration. As proof of principle, organisms having the four enzymes can now ferment pentose. Research, through use of kinetic modeling, is now optimizing expression of these enzymes for best production of ethanol.

Kinetic modeling involves mechanistic rate equations for each of the enzymatic reactions occurring inside the cell and mass and rate equations for the transport of major substrates into cells and of major products out of the cells. For *Zymomonas*, the kinetic model is comprised of 24 rate expressions and 24 balance equations. In this approach, the derivatives of the balance equations were set to zero to calculate a steady state and the resulting non-linear equations were solved numerically to calculate steady state metabolite concentrations. The amounts of five enzymes, phosphoglucose isomerase, xylose isomerase, xylulokinase, transaldolase, transketolase, were varied so that the total of Entner Doudoroff and Pentose Phosphate pathway enzymes ranged from 42%-66% of the total cellular protein. The $k_{cat}$ values were determined when necessary for the temperatures chose for production. These values are available from various sources in the literature, and in this case, the $k_{cat}$ values reported at varying temperatures were normalized by using the 'glucose consumption vs. temperature' table presented by Scopes and Griffiths-Smith (1986).

Dr. Kompala described the assumptions and conditions used for the model and then provided sample rate expressions and representative balance equations. Every intracellular metabolite undergoes synthesis and consumption, resulting in growth. The balance equations take into account both rate expression and dilution due to growth. The kinetic model can predict dynamics of metabolite concentrations and accounts for those that are fast reactants, those that are slow, etc. Thus, the kinetic model simulation at each enzyme concentration provides information on both steady state and dynamic variations in terms of metabolite concentrations. Metabolic flux analysis will not provide the amount of information that can result from kinetic modeling. Using the computer to estimate reactions, modeling can provide researchers with critical information on how to maximize ethanol concentration and/or to optimize ethanol production.

Using the kinetic model, Dr. Kompala looked at five different concentrations of the five enzymes under study and found that they gave about the same amount of ethanol. He then looked beyond the amount of ethanol produced to see what the cost was on the cells. It was determined that phosphoglucose isomerase doesn't need to be more than native amounts;

collaborators varied amounts of this enzyme and found that increased amounts made cells sicker due to metabolic burden. Dr. Kompala's modeling research showed that essentially same amount of ethanol could be produced with less enzyme expression, indicating that proteins don't have to be overexpressed to get maximum product. He noted how these mathematical predictions mimic what was described by Dr. Smolke as her experimental observations in her talk.

The kinetic model simulations at each enzyme combination produced a large amount of information, ranging from steady state metabolite concentrations to dynamic variations in the individual reaction rates for each enzymatic reaction and metabolite. Amongst the five enzymes whose amounts were tested in terms of the ethanol production, it was found that transaldolase, xylose isomerase, and xylulokinase were the most important enzymes in terms of effect on the extracellular ethanol concentration. Producing these enzymes at sufficient levels eliminates pathway bottlenecks and significant accumulation of intermediate metabolites. Thus, optimal concentrations of key enzymes should be determined for ethanol production. Modeling showed that the relatively low amounts of native phosphoglucose isomerase and heterologous transketolase are sufficient to enable maximal ethanol production. Results also show the reaction rates in the glycolysis pathway are higher than the rates in Pentose Phosphate pathway so that 'glyceraldehyde 3-P' is channeled quickly into the glycolysis pathway. With the flow of 'glyceraldehyde 3-P' into the glycolysis pathway, the amounts of 'erythrose 4-P' and 'fructose 6-P' drop dramatically when transaldolase and transketolase enzymes are available. The relatively slow rate of reactions in the Pentose Phosphate pathway increases importance of the enzymes xyloseisomerase and xylulokinase in controlling the reactions immediately following xylose transport into the cell.

Thus, the kinetic modeling strategy allows researchers to incorporate known information using non-linear rate equations and regulation by other metabolites. This approach can be a vital tool in determining optimum combinations of heterologous enzymes to maximize pathway efficiency. The model also enables comparison of enzymatic reaction rates between pathways and can calculate the metabolic flux for each enzymatic reaction at all the assumed levels of heterologous enzymes.

Dr. Kompala was asked how the model accounts for perturbation by a stressor such as oxygen. He responded that in this case, the particular reaction is anaerobic. However, one would need to take such variables into account when setting up model. There was also a question about enzymes in a sequential pathway. Many enzymes required in sequential pathway are localized in the same area in the cell so substrate is "floating around" but moving from one enzyme to the next. Dr. Kompala noted that it may be difficult to do kinetic modeling of this type of pathway. A more accurate way of modeling would have to account for heterology in the cell but this model would be one of the simplest ways to approach it.

**"Development of Tolerant and other Complex Phenotypes for Biofuel Production"**

E. Terry Papoutsakis
University of Delaware

Potential toxicity of a product in the cell is an issue for metabolic engineering as noted by the first two presentations of the session. In the third presentation, Dr. Papoutsakis discussed how metabolic engineering deals with potential toxicity by developing tolerant phenotypes for production. He noted that complex phenotypes involve many genes. These genes and their interactions and behavior may not be well-characterized or even identified. This lack of information makes it difficult alter or develop phenotypes for bioprocessing.

The challenge is to develop cells that can tolerate harsh bioprocesing conditions not typically seen in nature, for instance tolerating 30% ethanol, producing cellulosic biomass 2-3X faster than currently possible, or utilizing cellulose or xylans 5-10X more effectively than currently possible for biofuel production. Dr. Papoutsakis is studying stress responses to address this challenge for metabolic engineering. The goal of his research is to identify genes and cellular programs in Clostridia in response to solvent, such as butanol, and carboxylic acid stress. This will allow identification of general and specific stress regulons and of genes which may be used to impart solvent or acid tolerance to engineered cells. The resulting knowledge could be applied to production of solvents via fermentations, providing a green alternative to petrochemicals.

Currently, toxicity with butanol is a significant problem in fermentation. Solvent toxicity inhibits growth and glucose uptake and disrupts membrane integrity. A genome scale microarray was used to assay genes up- and down-regulated by solvent (acetate, butanol, and butyrate) stress and provide an overview of multi-stress responses. Three broad classes of genes were identified: energy production and metabolism; amino acid transport and metabolism; and post-translational modification, protein turnover, and chaperones. This analysis also showed which genes were involved in universal or common stress responses as well as which genes were associated with response to a specific stressor such as butanol. The study showed differential expression of branched chain amino acid genes involved in conversion of branched chain amino acids into branched chain fatty acids. Using observations of cold-shock adaptation in *B. subtilis* and butanol stress in *C. acetobutylicum*, it was suggested that synthesis of branched chain fatty acids in *C. acetobutylicum* may help the cells avoid membrane disruption and thus tolerate solvents such as butanol. A gene specifically induced by butanol stress is *glpA* which may regulate solvent tolerance by altering membrane composition, consistent with the regulation of osmotolerance by yeast *glp* genes. Genes universally up-regulated by acetate, butyrate, and butanol included the *groESL* chaperones. Dr. Papoutsakis' laboratory has demonstrated the benefits of *groESL* overexpression in enhancing solvent tolerance, suggesting that establishing protein stability, and thus functionality, is key in responding to toxic stress.

While stress responses can sometimes predict stress tolerant phenotypes, there is no obvious high throughput method to test and use stress responses to achieve strong solvent tolerant phenotypes. Tools are needed that can be used broadly for development of

complex phenotypes. Expression or genomic library screening followed by deep sequencing has been applied to identify single genes or loci that play a role in development of a selectable stress tolerant phenotype. It is not known how well this approach works for single loci, let alone multiple loci that interact, and if these genes or loci are components of a stress regulon. This method typically selects for growth under a particular condition which may not be optimal for production. The library approach was used to identify genes involved in butanol and butyrate tolerance. The identified genes were mostly transcriptional regulators. Several imparted solvent tolerance when expressed alone work but, in general, they worked better as a group in a mixed population rather than in pure recombinant strains. The studies of butyrate tolerance found a consistent enrichment of IR (non-coding) DNA of the rRNA locus. While the regulon has been determined by microarray analysis, the mechanism for butyrate tolerance is not yet known. Tolerance to butyrate imparted tolerance to all tested carboxylic acids and led to higher butanol formation, suggesting increased tolerance to butanol.

Dr. Papoutsakis noted that while the library approach was useful for identifying many loci, it did not identify all involved in tolerance. It appeared to primarily identify transcriptional regulators and did not identify chaperones. The library data showed little overlap with the butanol stress regulon, suggesting the method missed many genes and loci important for stress tolerance. In general, the library approach missed large, multigenic program. The library approach depends on a selection assay as well as on insert size and regulation of gene expression. The stress regulon approach works best when there is a dormant adaptive response such as in clostridial adaptation to butanol and butyrate accumulation. This approach also works in identification of potential target programs and pathways.

Current and proposed work in Dr. Papoutsakis' laboratory is continuing to address methods to improve tolerance and to develop strains that simultaneously, quickly, and efficiently ferment cellulose and xylans without enzymatic pre-treatment. Studies include development of high throughput methods to capture and combine distant multi-locus effects and allogenic traits. Approaches will be coupled with enhanced or temporarily induced recombination. Research is also exploring the relationship between differentiation (sporulation) and tolerance. Both mature and immature spores are extremely resistant to chemicals, radiation, and other harmful agents. The question being pursued is whether differentiation engineering can be use to freeze differentiation or change the differentiation state of the cell so that it is at a stage where the cell is tolerant and yet a good producer.

At the conclusion of the presentation, Dr. Heineken asked if the status of the BP and DuPont butanol production efforts was known. Dr. Papoutsakis responded that BP is using traditional methods for butanol production in Europe and South America. Not as much information is available on the efforts by DuPont, but they are generating patents.

**"Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant *Synechocystis* PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas"**

Roger L. Ely
Oregon State University

Dr. Ely's research focuses on hydrogen production in both dark and light in the cyanobacterium *Synechocystis*. During the day, photosynthesis is taking place but during night, *Synechocystis* breaks down the sugars and releases hydrogen. In the light, hydrogen drops down to essentially zero and in the dark the hydrogen starts coming back up again as oxygen comes up. In the mutant strain, however, oxygen comes up slower in the dark and hydrogen drops slower in the light. The peak of hydrogen production also differs between the wildtype and mutant strains. The research is attempting to answer the question whether differences in hydrogen production result are due to oxygen or due to electron transfer.

The approach examined both the light and dark pathways by investigating reductant supply and oxygen sensitivity. The study of reductant supply involved initial screening followed by biological hydrogen production profiling, metabolic flux analyses, and specific inhibitor. Studies of oxygen sensitivity utilized rational enzyme modifications and directed evolution. The project required a method to screen a lot of cultures very quickly. To meet this need, Dr. Ely and colleagues developed a high throughput screening assay for assessing biological hydrogen production. They have shown that the assay is semi-quantitative, sensitive to 20 nmol of hydrogen or less, and largely unaffected by oxygen, carbon dioxide, or volatile fatty acids at levels appropriate to biological systems.

They then attempted to optimize different variables such as media nutrient composition (N, S, or P limitation) or pH on hydrogen production profiles. Sulfur-deprived cultures produce more hydrogen than cultures grown on complete media. It was found that the highest hydrogen production occurred with low concentrations of sulfur and nitrogen instead of in the absence of sulfur or nitrogen.Hydrogen profiling experiments were used to characterize further the significant variables for hydrogen production. The experiments indicated that for hydrogen production, optimized amounts of nitrogen and sulfur in the nutrient media were superior to total deprivation of these nutrients. These experiments showed that by optimizing conditions up to 60-fold greater hydrogen production could be achieved. Metabolic flux analyses were also used to understand the biochemical reactions underlying hydrogen production, linking genomic, proteomic and metabolomic information. Dr. Ely conducted flux analyses on both wildtype and high hydrogen-producing strains with impaired NADH-dehydrogenase function and used these analyses to develop mathematical models of the metabolism underlying hydrogen production. The models were consistent with behavior of both the wildtype and mutant, thus validating the approach and leading to use of the models to predict possible effects of different mutant phenotypes on hydrogen production.

The sensitivity of the *Synechocystis hydrogensase* to oxygen limits the utility of Synechocystis for high yield hydrogen production. To attempt to overcome this, Dr. Ely

replaced the wild-type *Synechocystis* hoxH subunit of hydrogenase with the homologous, structurally similar hoxH subunit from *Ralstonia eutropha*, generating the strain REHX. The *Ralstonia eutropha* hydrogenase has significant sequence homology with the *Synechocystis* enzyme but is oxygen tolerant. The *hypX* gene that encodes an accessory protein essential for oxygen tolerance in *R. eutropha* was also inserted. Hydrogen production was observed with the appropriate timing of oxygen production, implying the protein functioned with the subunits of the wildtype hydrogenase. However, the hybrid hydrogenase had a much lower total activity, limiting this approach. Dr. Ely noted that it was difficult to sort out reductant availability from oxygen effects in these experiments.

In an alternative approach, rather than putting in an enzyme from another organism, the residues in the native *Synechocystis* enzyme were modified based on sequence comparisons with other related organisms. A 300 bp region near the active site of the enzyme was knocked out to generate the PSS1 mutant. Hydrogen production was then evaluated and effects of different specific inhibitors assayed. Interestingly, different measurement methods gave different results. When using the methyl viologen assay, no hydrogen was detected in PSS1. However, using a different assay, hydrogen production in the mutant appeared to be greater than in wildtype cells. Other mutations (deletions) in conserved regions were also analyzed. The various mutants still produced hydrogen and reacted differently to inhibitors. While the only hydrogenase appears to be knocked out, hydrogen is still being produced, indicating there is still a lot to learn about hydrogenase and hydrogen production.

An audience member asked if Dr. Ely could tell if the enzyme was oxygen sensitive after the *R. eutropha* enzyme was added. He replied that it looks like it is still oxygen sensitive, but there is more to the story. Another audience member asked if the same electron donor was used in both cases for the hydrogenase and was told that it was NADPH in both cases.

**"Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol"**

John Frost
Michigan State University/Frost Chemical Laboratory

Dr. Frost began his presentation by explaining that the project on 1,2,4-butanetriol trinitrate (BTTN) started while he was at Michigan State University and is now being continued at the company Draths. BTTN is less volatile than nitroglycerine, more thermally stable than nitroglycerine, and has lower shock sensitivity relative to nitroglycerine. Butanetriol (BT) is the very expensive starting material for BTTN. BT has not been produced in the US since 1994, and now all BT is imported from China. Thus, new less expensive methods need to be developed for BT manufacture. The basic chemistry of BT manufacture is problematic and requires a series of distillation columns to clean up the product. To circumvent this issue, Dr. Frost is developing a microbial catalyst.

Dr. Frost started with *E. coli* and brought in genes from two other microbes, *Pseudomonas fragi* and *Pseudomonas putida*, creating a biosynthetic pathway that does not exist in nature. The enzymes catalyze reactions on either native or non-native substrates to

generate BT. Basically, the pathway starts with a C-5 sugar and ultimately ends up with a simple aldehyde reduction. However, *E. coli* has another biosynthetic pathway which competes for components of BT pathway. A *Pseudomonas fragi* D-xylonate dehydratase was then purified to elaborate on the pathway. In *E. coli*, aldolase is the competing enzyme, not a dehydratase, but catalytic genes are clustered in *E. coli* so this type of approach can be used to elucidate the d-xylonic acid catabolic pathway. The partial sequence of dehydratase was then used as a rosetta stone, and using ERGO database, a series of D-xylose dehydrogenase were identified, cloned, and assayed for selectivity for xylose, glucose, etc. Based on this analysis, the *Caulobacter crescentus* xylose dehydrogenase was inserted into the *E. coli* chromosome to generate an *E. coli* strain that has plenty of xylose dehydrogenase and can no longer grow on xylose. Strains were then used for fermentor-controlled synthesis of BT from xylose.

Dr. Frost noted that during BT biosynthesis, several other byproducts are produced such as pentanoic acid and derivatives. The decarboxylase isn't active enough to push the reaction more towards BT. He noted that the finding that butyric acid is also a byproduct was somewhat surprising. The engineering of this pathway, however, now enables production of pure BT without having to go through a series of distillation columns.

The microbial catalysis method enables green synthesis of BT. The synthesis utilizes nontoxic, renewable xylose, reduces salt waste streams, and avoids elevated temperatures and pressures. As intended, it is a single step process that will allow domestic manufacture of BT. Important dual-use markets also come into play with the metabolic engineering of this microbial catalytic pathway. The process is important not only for C-4 based molecules such as 1,4 butanediol, erythritol, and 1,2,4-butanediol but also for C-3 based products such as glycerol, 1,3 propanediol, 1,2 propanediol. The C-3 biobased products are the furthest along in terms of commercialization and are either being commercialized or on the cusp of commercialization. The C-4 biobased products have unique market potential.

Dr. Frost was asked how far the engineered organism is from commercial availability. He indicated that a ton of BTTN is scheduled for delivery in 2009. He was also asked if patents on glycerol purification research can provide information on separating organics from BTTN. He responded that this is not trivial, but scientists in this field know how to do it.

**General Discussion**

Dr. Heineken moderated the audience discussion. He began the session by requesting input on the topic areas of interagency interest and to consider for proposal solicitations. The last Interagency Opportunities for Metabolic Engineering solicitation listed four areas of particular interest:

1. Instrumentation, tools, and new experimental methods to facilitate the study of metabolic pathways.
2. Quantitative and conceptual models integrated with experimental studies to characterize the regulation and integration of complex, interacting metabolic pathways.

3. Bioinformatics to deduce the structure, function, and regulation of major metabolic pathways.
4. Designing new pathways through engineering of metabolic pathways for a desired product, producing novel substances or otherwise addressing novel problems in the understanding or manipulating such pathways.

Audience members suggested several possible areas for research with the goal of improving and optimizing metabolic engineering for efficient and effective production. These areas are:

1. Cell complexity. Audience members commented that a greater understanding of what is happening inside the cell would benefit metabolic engineering. Techniques, particularly at single cell level, are needed to help scientists visualize intracellular activities. This information can then be used to design models for engineering pathways. It was noted that these two areas are complementary and can be used to inform the other. In other words, research will go back and forth between model and data, informing researchers of approaches to design and modify pathways for efficient production of the desired product.

2. Bioinformatics. Bioinformatic tools were again seen as an important area to aid metabolic engineering. An increased understanding of how to use bioinformatic data is needed. How can bioinformatics help researchers modify pathways to design new desired pathways and products? What is the best way to use bioinformatics in metabolic engineering?

3. Signal transduction. It was noted that development of desirable strains to achieve a specific goal requires an understanding of intricate traits. For instance, how do cells make difficult decisions to deal with toxicity or stress? Greater knowledge of signal transduction, along with tool development, is needed. The resulting conceptualization of signal transduction processes can enable better and more appropriate design of metabolic pathways and new strains for high-yields of desired products.

4. Complex phenotypes. A challenge for metabolic engineering is how do researchers engineer for activities or processes that science doesn't know exactly know how to define such as microbial robustness. Scientists know microbial robustness, if and when they see it, but they can not really define it. However, effective and appropriate design of an organism requires that the complex "indescribable" phenotypes be considered and designed as well. Dr. Heineken remarked that systems biology incorporates metabolic engineering pathway design and also is related to the discussion of complex phenotypes. Systems biology attempts to take into account what the cell is doing and goes into understanding the phenotype the cell expresses and forms. He noted that this area of research is probably beyond the scope of metabolic engineering and really is more systems biology.

5. Enzyme genomics. Audience members commented that there a lot of enzymes, but they can't be used for metabolic engineering, because they haven't been sequenced yet. The question was then asked for what fraction of enzymes having an EC number do we have a

genomic sequence. It was answered that actually less than 50% have genomic sequence. Thus, an effort is needed for a large scale characterization and identification of the genomic sequence of these enzymes. Some of this is due to shortcomings in the databases. Sometimes sequences are known, but simply are not yet available in a database. It was noted that approximately 15% of the enzymes fall into this category. However, it was pointed out that a vast majority of the sequences have been elucidated and it is a matter of matching sequence with enzyme. In this case, researchers could determine the N-terminal sequence of an enzyme and then match that sequence against the genomic sequences in the databases. It was emphasized that proper identification and matching of genomic sequence with specific enzymes or proteins requires both wet laboratory work as well as bioinformatics. This is a big problem. It would need more than single investigator grants and would most likely entail a major initiative using collaborative efforts.

6. Functional plasticity. When considering metabolic engineering, the pathway design and modification can not be looked at in a static way. To the extent to which there is dynamism, organisms when pushed in one direction often will move in another direction. Research needs to consider and address the dynamics of a system or organism, perhaps by developing mutant libraries and looking at function under different conditions.

7. Chemical process. Most of the global chemical processes are continual, while most microbial processes are batch (except ethanol). For many of the microbial processes, the organism is dead after 48 hours. The challenge is to move these microbial processes to continual rather than batch. It was stated that most commercial processes max-out at 4 g/L/hr. Can this be increased and how? What is the maximum currently possible and can that barrier be broken? Research would need to get into sigma factors and what is happening in the organism in log phase. Dr. Heineken noted that this is an important yet challenging area, especially considering the mutations that occur in long term processes.

8. Other issues. Several audience members brought up general matters that should be considered in regard to metabolic engineering.

- One issue is the generation of transgenic molecules and related concerns such as public acceptance.
- Economic metrics was also mentioned as a matter that should be considered. The metrics would include time, energy, and economic environment as factors.
- It was suggested that the program name be broadened to Metabolic and Cellular Engineering. This name would reflect the need to understand overall cellular processes that can affect production. An understanding of metabolism and metabolic pathways is not sufficient; cellular processes such as regulatory mechanisms, signaling, and stress/toxicity tolerance play a significant role in production of a desired product.
- Another issue is whether cell-free, biomimetic studies should be considered in the program. Research is needed on the reconstruction of whole pathways in biomimetic systems for production of new or novel substances.

# 5. Agency Activities in Metabolic Engineering

***Department of Agriculture (USDA)***

The Cooperative State Research, Education and Extension Service (CSREES) is the USDA agency that participates in the Interagency Metabolic Engineering Working Group. In both the CSREES and USDA Strategic Plans, six goals are listed:

- Enhance international competitiveness of American agriculture
- Enhance the competitiveness and sustainability of rural and farm economies
- Support increased economic opportunities and improved quality of life in rural America
- Enhance protection and safety of the nation's agriculture and food supply
- Improve the nation's nutrition and health
- Protect and enhance the nation's natural resource base and environment

Metabolic Engineering (ME) can enhance competitiveness and sustainability of the US agricultural system through the production of commercially useful products such as chemicals, biofuels, and biomolecules from agricultural commodities. Through modification of plants, animals, and microorganisms, ME can also result in new uses for existing crops and animals, added value to traditional agricultural products, and improved quality and yield of agriculturally derived foods and materials. It is also possible through ME to produce plants with enhanced nutritional value or to modify plants and microorganisms for improved sustainable production of biofuels.

The participation in MEWG has allowed CSREES to leverage funding for support of several research projects that address one or more of CSREES' and USDA's goals. Funding is supporting research on metabolic engineering of biofuels that may lead to maximized ethanol production as well as reduced costs. Funded projects in microbes include the development of a strain that converts propane to propanol with high efficiency, with the long term goal of deciphering combinations of genetic modifications and process conditions to improve biocatalytic efficiency and develop novel and practically useful oxygenation biocatalysts and processes, and the production of flavor compounds in microbes that may eventually lead to improvements of metabolic function for processing of agricultural biomass and manufacture of bio-based industrial products. Funded metabolic engineering research projects in plants have the potential to produce fruits and vegetables with increased nutritional value and extended shelf-lives, to increase natural product-based disease and pest resistance, to enhance oil production in oilseeds, to limit effects of stresses on production of specific plant compounds, and to modify plants for production of pharmaceuticals and other economically important compounds. Thus, metabolic engineering, through both basic and applied research, is of vital importance for achieving the strategic goals of CSREES and USDA.

***Department of Commerce (DOC)***

The MEWG supports the DOC mission by advancing research and development of new commercial and industrial processes. As an emerging technology whose scientific basis is developing rapidly, ME is important to DOC'S NIST and especially its Biotechnology Division. NIST is especially interested in ME projects that support the development of biological and metabolic models, measurement methods and standards.

***Department of Defense (DOD)***

The Department of Defense (DoD) currently supports a broad range of research in the area of metabolic engineering through the Army Research Office (ARO) and other Army research activities, the Air Force Office of Scientific Research (AFOSR), the Office of Naval Research (ONR), and the Defense Advanced Research Projects Agency (DARPA). The specific focus of the ARO, AFOSR, ONR, and DARPA efforts will be summarized and future directions in metabolic engineering research and technology development will be addressed.

The broad needs for the DoD that can be served through research efforts in metabolic engineering are summarized below. These science and technology targets will provide enhanced and expanded capabilities for the missions of the services and provide greatly expanded capabilities for the civilian sector.

- Materials
- Processes
- Devices
- Fabrication Schemes
- Information Processing

Current interests in metabolic engineering at ARO are focused on the characterization of biochemical pathways, inter- and intra-cellular signaling, and enzymatic mechanisms, and the genetic basis for manipulation of protein expression, structure and function, and cell fate, in systems with potential relevance to the Army. The goal is to develop a detailed understanding of how macromolecules and cells execute their designated functions and how they interact with other cells and macromolecules. With this information, it will be possible to design and engineer particular sub-cellular elements and metabolic pathways and cell systems to exhibit a set of specific functions and properties, according to Army needs, and to identify and non-invasively correct molecular deficiencies to optimize and maintain cognitive and physical performance under normal and extreme conditions. ARO currently supports research in several areas, including: how molecular transport, subcellular compartmentalization, and reaction sequences are involved in enzymatic regulation and superstructure formation; understanding and manipulating aminoacylation of tRNAs and genetic code expansion to produce new polymeric peptides containing non-natural amino acids; biologically based means for fabrication of functional nanostructures; systems engineering of cell differentiation processes; the role and regulation of classes of proteins differentially expressed in response to environmental or external stimuli;

molecular genetics and genomics of human cognition, performance and function; and the design and implementation of unique biomolecular and cell based strategies for economically and environmentally favorable manufacturing, as well as the biodegradation of environmental pollutants.

AFOSR's metabolic engineering efforts focus on elucidating the fundamental science (a) to advance miniature biofuel cells for sensor and micro UAV applications and (b) to control the metabolic generation of hydrogen and lipid products in photosynthetic microbes for use as novel sources of future fuels. To this end, mechanisms for *in vivo* or *in vitro* metabolism of complex biofuels (mixtures of various sugars, cellulose, etc.) for energy production by biofuel cells are being explored. Characterization of electron and proton transfer in enzymatic redox reactions, and optimization of these reactions at an electrode surface, is also of interest. Knowledge is sought of the interactive photosynthetic pathways and the respective genetic control mechanisms that are able to balance cellular redox state with optimal electron/proton fluxes to ensure efficient and robust hydrogen or lipid production. Identifying and manipulating the molecular switches/sensors involved in re-directing photosynthetic metabolism toward hydrogen, starch or lipid production would be an important part of this knowledge base. Since hydrogenases can catalyze both proton reducing- and hydrogen oxidizing-reactions, they have special significance as key enzymes in the production of hydrogen fuel by photosynthetic microbes and as platinum-electrode substitutes in biofuel cells, respectively. Thus, advancements in both biofuels and biofuel cells may be simultaneously realized when the biophysical mechanisms of hydrogenase catalysis are combined with genetic engineering to improve or modify the enzyme's performance.

One of the metabolic engineering foci at ONR, currently, is the microbial synthesis of energetic materials (EM) and EM precursors for the purposes of cost and environmental impact. Practically all such materials are non-natural products and their biosynthesis therefore requires the re-engineering of existing pathways and/or the assembly of new or hybrid pathways in one or more host organisms. An example of a simple EM precursor now under study is 1,2,4- butanetriol, which as its energetic trinitrate is used as a plasticizer in propellant and explosives formulations. More advanced EM targets, such as RDX, HMX and Cl20, involve high density fused ring cores with multiple nitramino (C-N(NO2)) substituents. While these are very difficult targets, they suggest worthwhile research goals such as the biosynthesis of highly electron withdrawing substituents on carbon (as in C-nitramino) or the assembly of strained heterocyclic rings. Clearly, a theoretical/experimental approach to the prediction of the true scope of enzyme reaction specificity, with energetic boundaries, would be particularly valuable in the design of pathways for EM biosynthesis. Other non-polymeric targets, besides EM, would include novel photonic/electronic/optical materials.

DARPA's metabolic engineering programs are driven by an interest in protecting human assets against biological threats and using biology to maintain human performance. The general concept of this thrust is to understand how nature controls the metabolic rate of cells and organisms (e.g., extremophiles, hibernation) and apply this understanding to problems of interest to DoD. Examples of current investments in metabolic engineering

include efforts to develop technologies for engineering cells, tissues and organisms to survive in the battlefield environment so they can be used as sensors. Related basic research on biochemical circuit engineering in laboratory model organisms is also supported. In addition, DARPA is developing technologies that permit the long-term storage of cells including human blood. More complete descriptions of current DARPA programs and solicitations in these areas can be viewed at http://www.darpa.mil/dso.

***Department of Energy (DOE)***

The Department of Energy is supporting research in metabolic engineering research, largely through the Offices of Science (SC), Energy Efficiency and Renewable Energy (EE), and Environmental Management (EM). The research falls in two main categories: 1) basic research, which involves the advancement of metabolic engineering fundamental knowledge and capabilities, and 2) applied research, which employs metabolic engineering techniques in development of target products. The basic research efforts of the Department reside within SC, whereas most of the applied research in this area is conducted within EE. In general, these research efforts are conducted by universities, national laboratories, and industry.

The Department's goals related to metabolic engineering research are to:

- To expand the level of knowledge and understanding of metabolic pathways and metabolic regulatory mechanisms related to the development of novel bio-based systems for the production, conservation, and conversion of energy.

- Apply metabolic engineering techniques to enhance and develop plants and microorganisms for use in the production of chemicals and fuels or for environmental remediation of waste sites.

***Environmental Protection Agency (EPA)***

The mission of the Environmental Protection Agency is to protect human health and the environment from adverse effects of anthropogenic activity. Included in this mission are various elements for which metabolic engineering can play a useful role.

One prominent concern is the introduction of chemicals to the environment, which may have detrimental effects on humans and other biota. As mandated by statute and implemented by rule, the Agency routinely conducts evaluation of chemicals intended for use, currently in use, or determined to exist at significant levels in the environment. From these evaluations, the Agency may decide to implement management strategies designed to limit the potential for adverse effects.

The application of novel technologies such as the use of biotechnology as a substitute to conventional manufacturing and processing of raw materials into final products is consistent with the mission of the Agency. EPA implements this by supporting development of technologies which 1) use chemical substitutes that are less toxic; 2)

produce more efficient activity resulting in decreased requirement for the chemical or; 3) develop engineering procedures which produce little or no toxic end products. Finally, consistent with the pollution prevention ethic is the reevaluation of chemical stewardship from one of "cradle to grave" to a more multigenerational philosophy in which a chemical may be utilized successively in different forms prior to final disposal. Metabolic engineering has a role to play by enabling the development of biological mechanisms for production or use that meet one or more of these criteria.

While it is generally accepted that chemical-based technologies have evolved to provide a higher standard of living for the general population, it is also recognized that the use of some chemicals, either through the chemical characteristics or the handling, synthesis or disposal, have produced negative effects on human health and/or the environment. Advances in technology allow scientists to better predict the potential for adverse effects from exposure to chemicals as well as mechanisms to diminish the negative effects of chemical production such as production of toxic byproducts and disposal of the chemical. The approach, which strives to identify synthetic pathways that are less polluting than existing pathways and that encourages the development of nontoxic chemical products, is referred to as "Green Chemistry". The use of metabolic engineering to evaluate the potential for increased risk from chemicals, by allowing the study of responsible metabolic pathways and by permitting modification of such pathways to reduce risk, is another way in which metabolic engineering firs within the EPA mission.

Finally, basic research, which utilizes methods of metabolic engineering, can provide longer range approaches to assist EPA in its overall mission of protecting human health and the environment.

***National Aeronautics and Space Administration (NASA)***

One of NASA's strategic goals is to extend the duration & boundaries of human space flight to create new opportunities for exploration & discovery. To prepare for and hasten the journey, the NASA Office of Biological and Physical Research must address the following questions through its research:

- How can we assure the survival of humans traveling far from Earth?
- What technology must we create to enable the next explorers to go beyond where we have been?

NASA's efforts in the area of metabolic engineering are on approaches and applications that will have a significant impact on the reduction of required mass, power, volume, crew time, and on increased safety and reliability, beyond the current baseline technologies. The targeted and purposeful alteration of metabolic pathways found in an organism may play a key role in the development of biological approaches and technologies that enable efficient use of spacecraft resources for long-duration space missions.

### *National Institutes of Health (NIGMS/NIH)*

The National Institute of General Medical Sciences (NIGMS) supports metabolic engineering research, usually in the form of grants to investigators in universities (R01s) or in small businesses (SBIRs). These grants support basic research in two general areas: (1) the development of microbial or plant-based metabolic routes to useful quantities of small molecules such as polyketides; (2) the development of a much better understanding of the control architecture that integrates the genetic and catalytic processes in normal and aberrant cells.

### *National Science Foundation (NSF)*

The mission of NSF is to:

- Promote the Progress of Science
- Advance the National Health, Prosperity, and Welfare
- Secure the National Defense
- Provide for Other Purposes

Support of ME research allows NSF to address specific goals within its mission. These include, but are not limited to; development of technologies integrating theoretical, computational, and experimental approaches to the study of metabolic processes; the targeted and purposeful alteration of metabolic pathways in living organisms in order to better understand and utilize these pathways for chemical transformation, energy transduction, and supramolecular assembly; providing a framework for studying the dynamics of interactions and interconversions of biological molecules in order to understand how organisms regulate specific physiological processes at the cellular and sub-cellular levels and the "cross-talk" between pathways; measurement and control of *in vivo* metabolic fluxes; metabolic control analysis of pathway groups or networks; development of *in vivo* techniques to accomplish these goals.

Metabolic Engineering has been supported in all interagency competitions by three Directorates within NSF. There is recognition at NSF that this activity has been beneficial to NSF and that NSF would like to continue with this Activity.

# Metabolic Engineering

Eighth
Interagency Conference
on
Metabolic Engineering

Fred Heineken
Chair
Interagency Metabolic Engineering Working Group

February 11, 2008

# Metabolic Engineering

## Participating Agencies/Departments

- Department of Agriculture
- Department of Commerce
- Department of Defense
- Department of Energy
- Environmental Protection Agency
- National Aeronautics and Space Administration
- National Institutes of Health
- National Science Foundation

# Metabolic Engineering

## Metabolic Engineering Working Group

## Goals

- Promote the Advancement of Metabolic Engineering
- Coordinate Federal Metabolic Engineering Research
- Identify and Address Gaps in Metabolic Engineering Research

# Metabolic Engineering

## Accomplishments (13 Years)

- Held Three Interagency Workshops on Metabolic Engineering
- Held Eight Interagency Conferences on Metabolic Engineering
- Issued Six Interagency Announcements of Opportunities in Metabolic Engineering
- Generated an Interagency Web Site (www.metabolicengineering.gov)
- Research Results have been Published in High Impact Journals

# Metabolic Engineering

6 Interagency Announcements of Opportunities in Metabolic Engineering

Topic Areas of Interagency Interest

- Instrumentation, Tools, and Methods to facilitate the Study of Metabolic Pathways in Cells
- Quantitative and Conceptual Models
- Bioinformatics
- Engineering of Metabolic Pathways for a Desired Output

# Metabolic Engineering

Responses to Announcements

| | |
|---|---|
| Proposals | 236 |
| Awards | 54 |
| Award Total | $34M |

# Metabolic Engineering

## Future Plans

- Issue New Interagency Announcements of Opportunities in Metabolic Engineering
- Sponsor Annual Interagency Conferences on Metabolic Engineering
- Maintain an Up-To-Date Interagency Web Site for Metabolic Engineering

# Metabolic Engineering

## Conference Theme

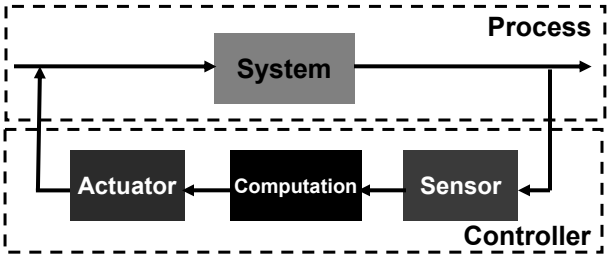
Metabolic Engineering for Biorefineries:
Alternative Fuels and Bioproducts

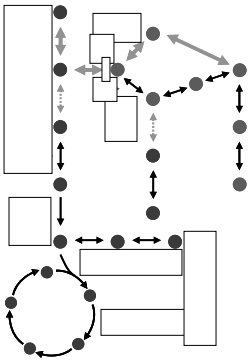
# Foundational advances in RNA engineering applied to control of biosynthesis

Christina D. Smolke
Division of Chemistry and Chemical Engineering
California Institute of Technology

February 11, 2008
MEWG: Interagency Conference on Metabolic Engineering

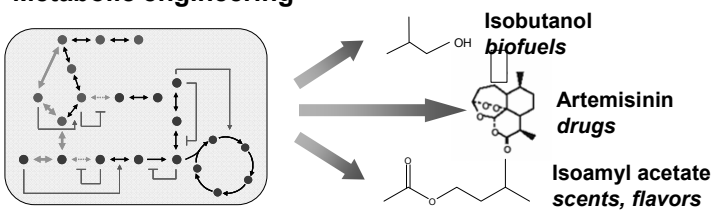
## The role of feedback control in biosynthesis



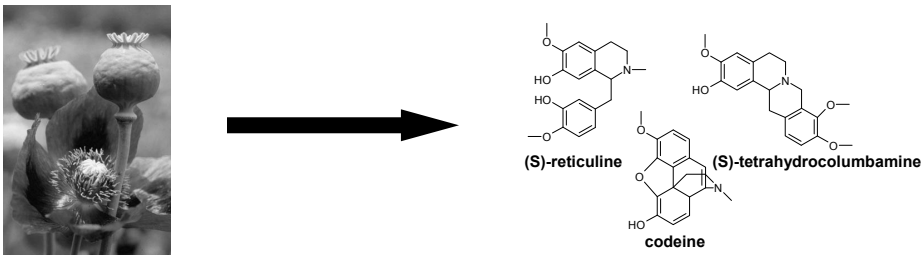



# Applications of engineered biological systems

**Metabolic engineering**




**Benzylisoquinoline alkaloid biosynthesis**

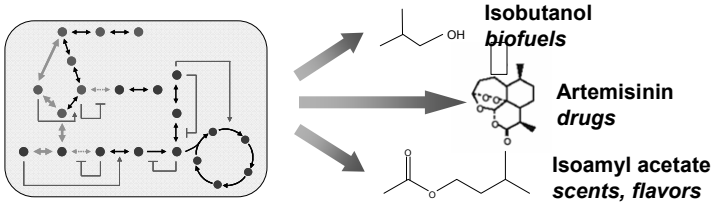



*Papaver somniferum* (opium poppy)

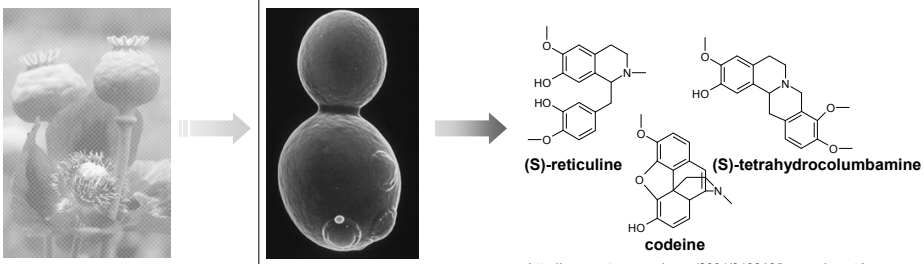
http://www.nature.com/news/2004/040816/images/yeast.jpg
http://www.westwindphotos.ca/images/101%20Opium%20Poppy.jpg

# Applications of engineered biological systems

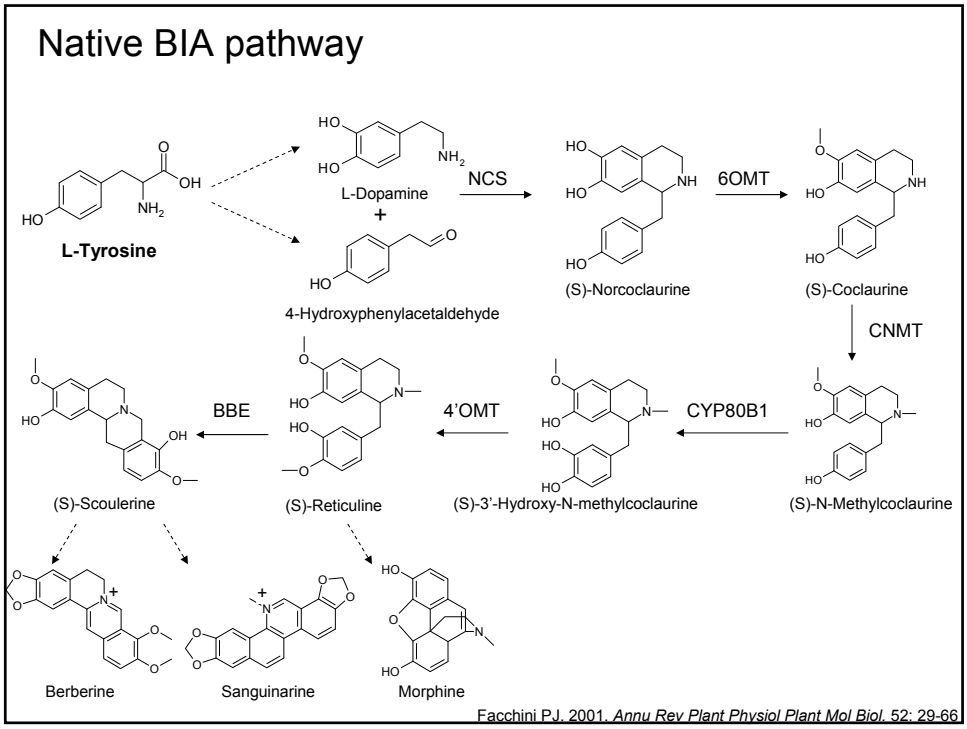
**Metabolic engineering**

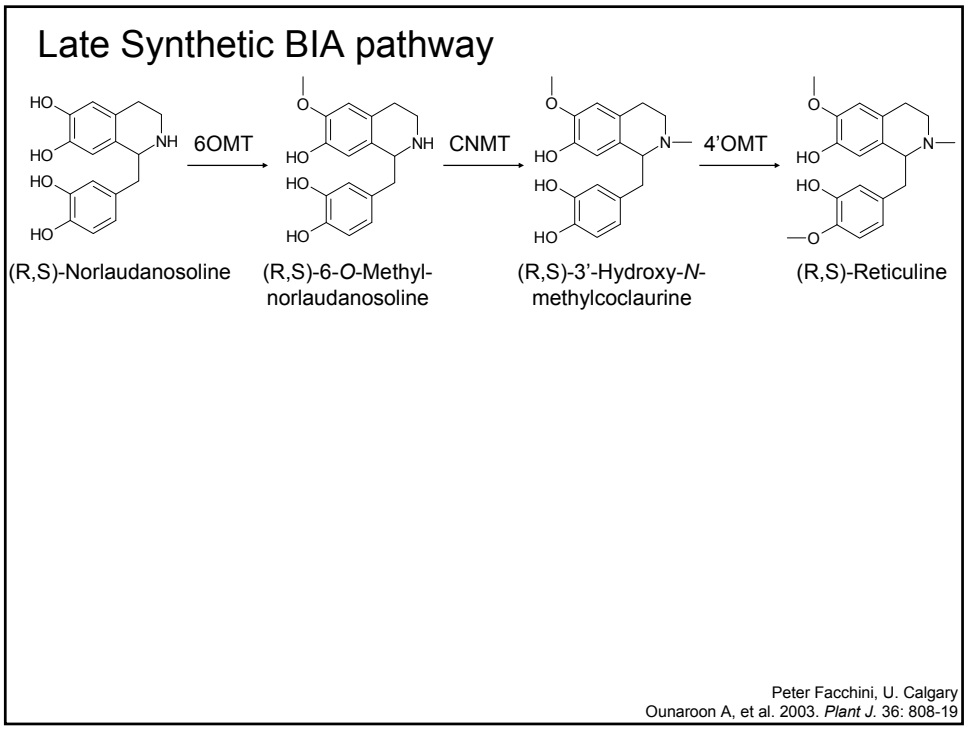


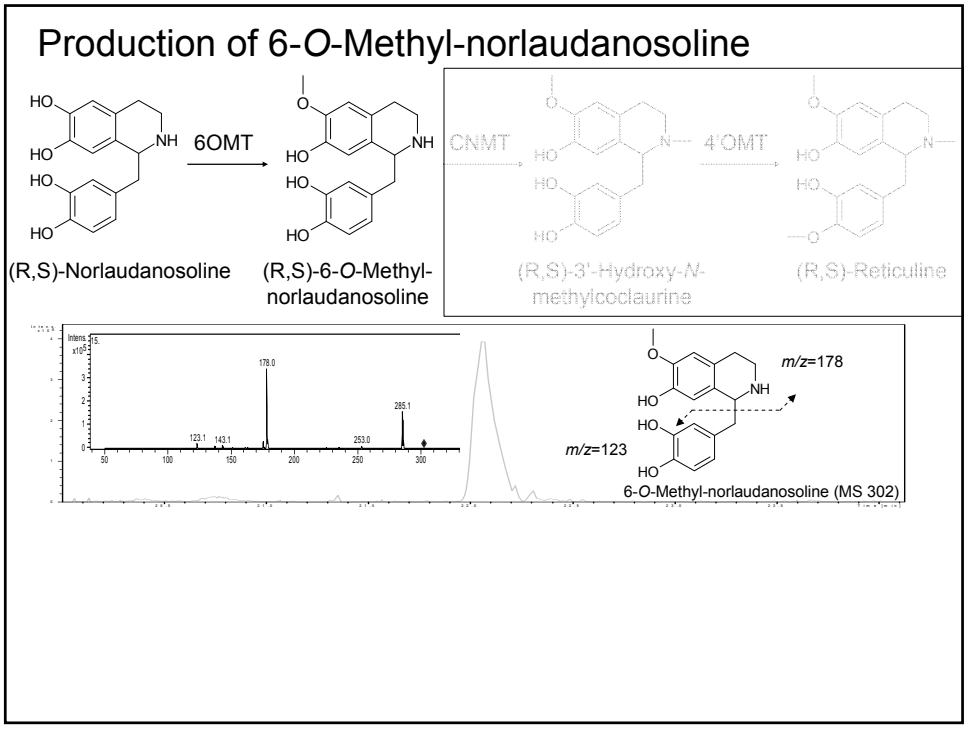

**Benzylisoquinoline alkaloid microbial synthesis**

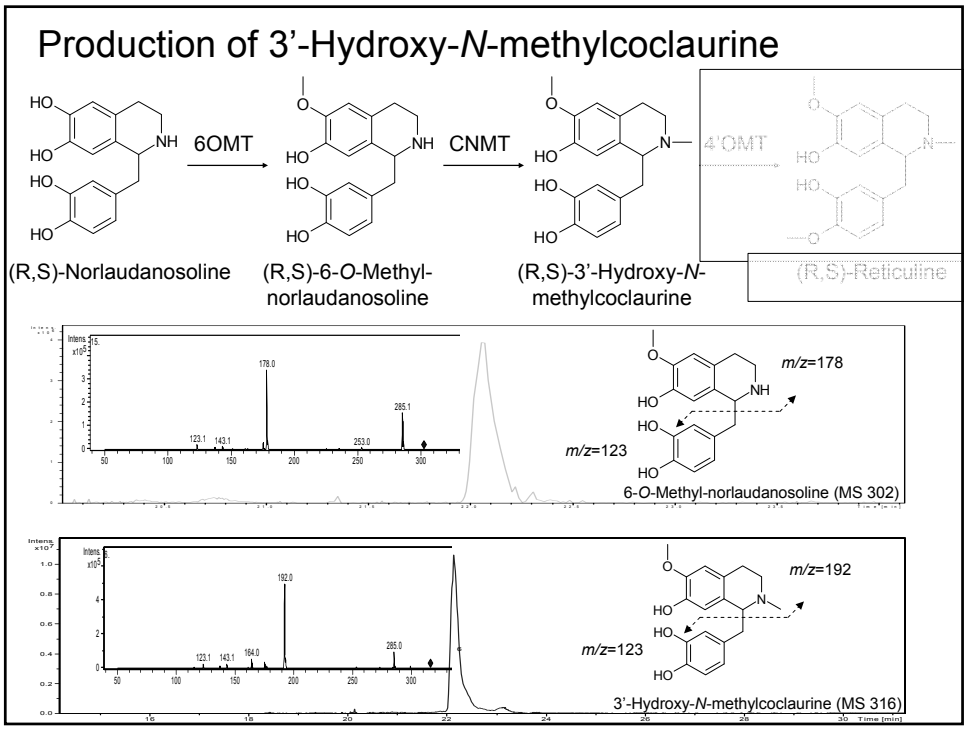


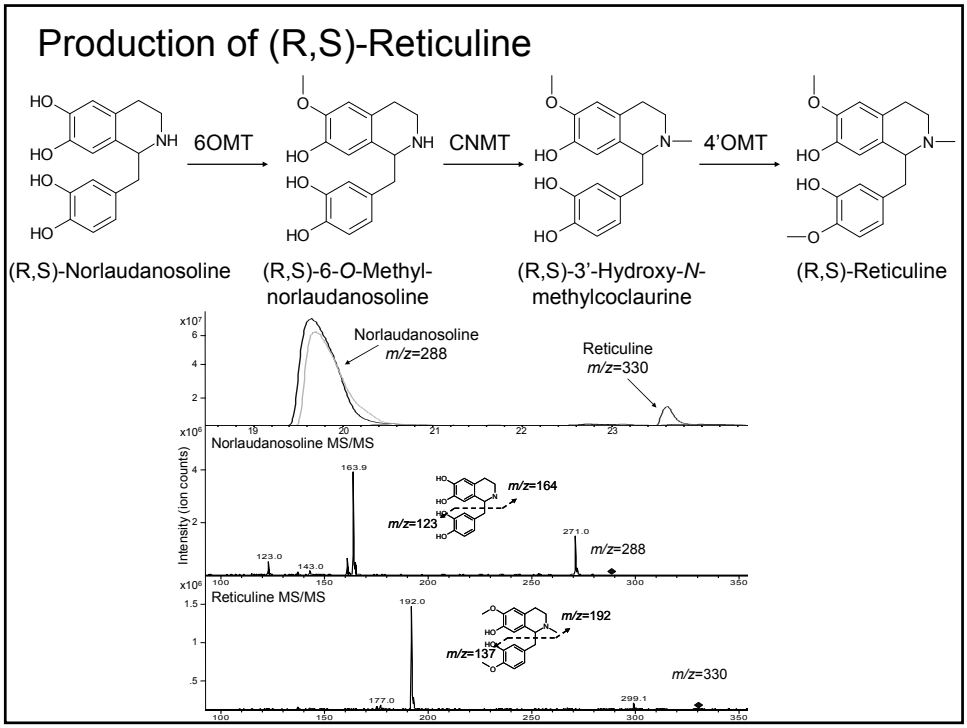

http://www.nature.com/news/2004/040816/images/yeast.jpg
http://www.westwindphotos.ca/images/101%20Opium%20Poppy.jpg

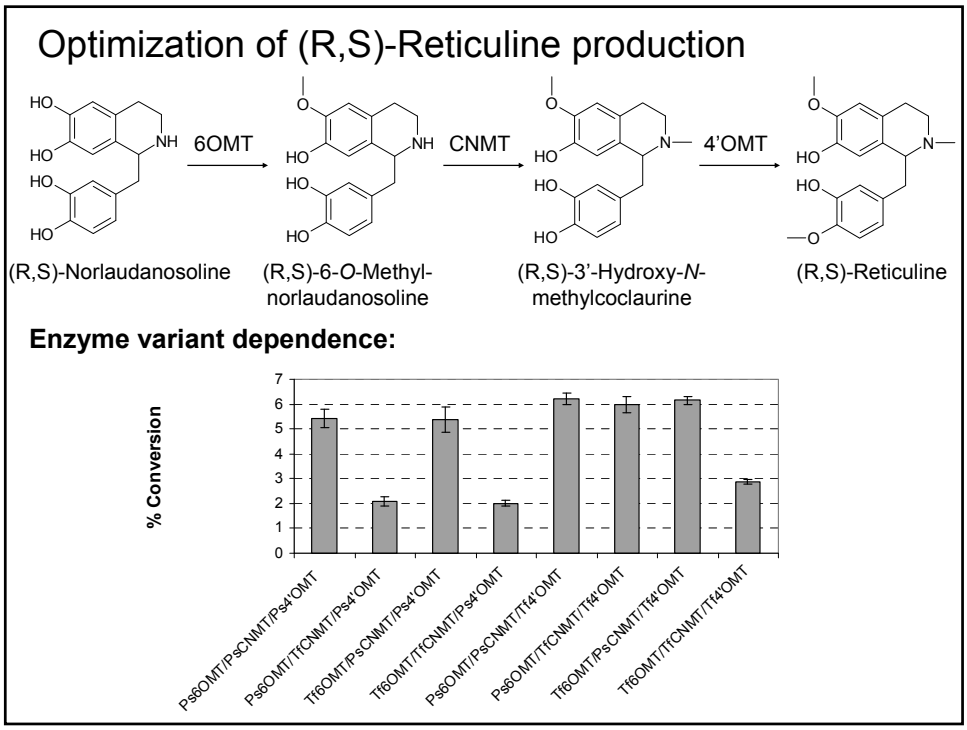

Native BIA pathway
L-Tyrosine
L-Dopamine
+
4-Hydroxyphenylacetaldehyde
NCS
(S)-Norcoclaurine
6OMT
(S)-Coclaurine
CNMT
(S)-N-Methylcoclaurine
CYP80B1
(S)-3'-Hydroxy-N-methylcoclaurine
4'OMT
(S)-Reticuline
BBE
(S)-Scoulerine
Berberine
Sanguinarine
Morphine
Facchini PJ. 2001. Annu Rev Plant Physiol Plant Mol Biol. 52: 29-66

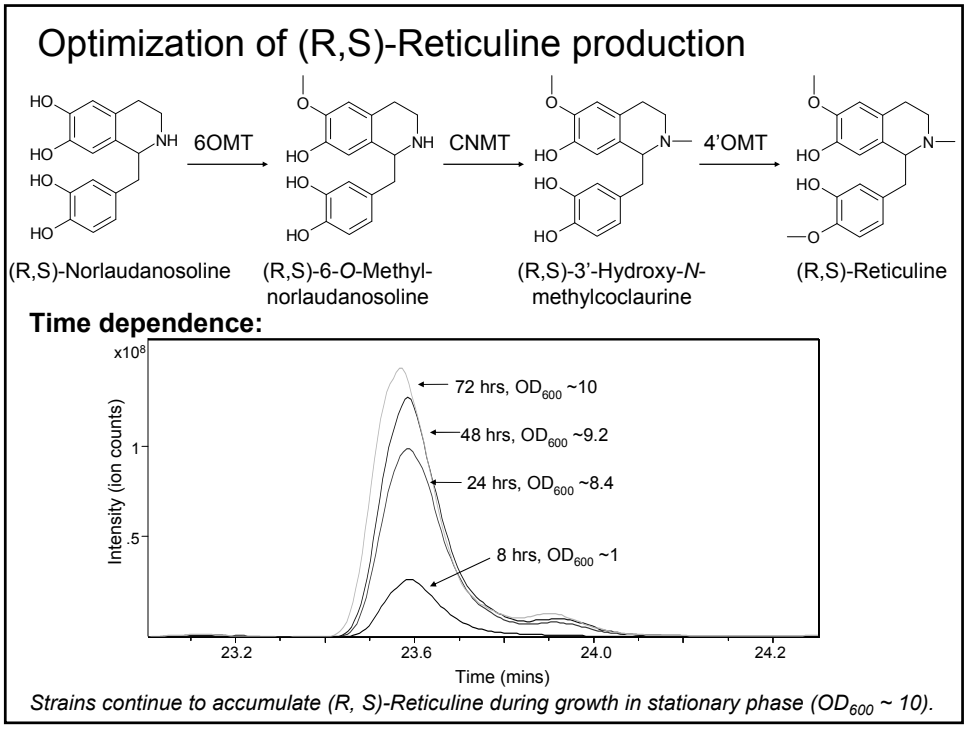

Late Synthetic BIA pathway
(R,S)-Norlaudanosoline
6OMT
(R,S)-6-O-Methyl-norlaudanosoline
CNMT
(R,S)-3'-Hydroxy-N-methylcoclaurine
4'OMT
(R,S)-Reticuline
Peter Facchini, U. Calgary
Ounaroon A, et al. 2003. Plant J. 36: 808-19

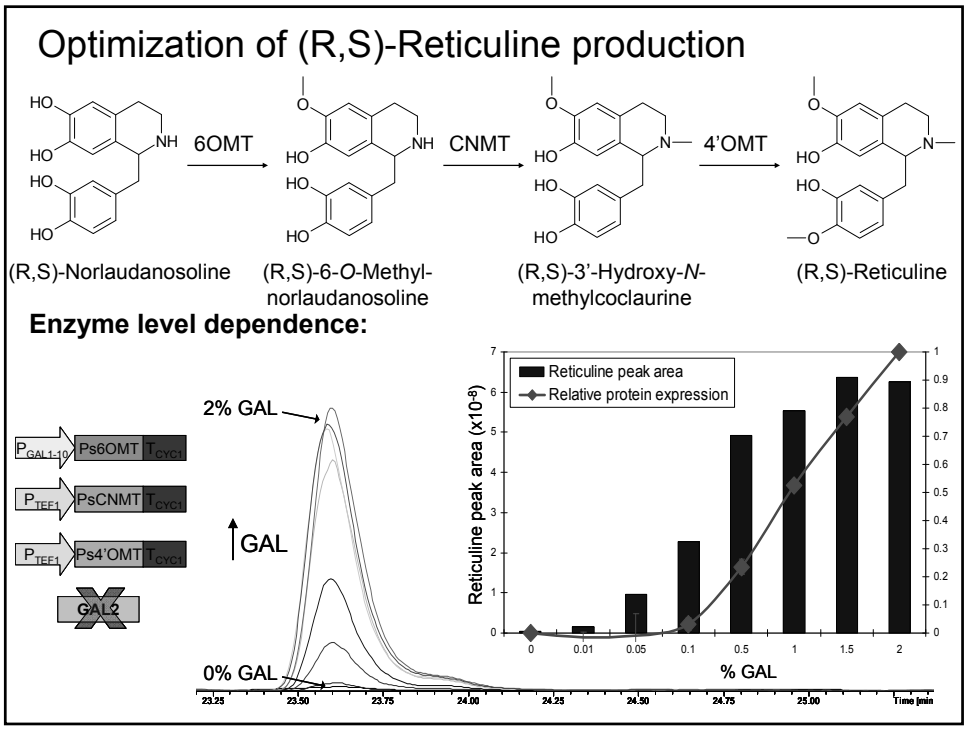

Production of 6-O-Methyl-norlaudanosoline
6OMT
CNMT
4'OMT
(R,S)-Norlaudanosoline
(R,S)-6-O-Methyl-norlaudanosoline
(R,S)-3'-Hydroxy-N-methylcoclaurine
(R,S)-Reticuline
m/z=178
m/z=123
6-O-Methyl-norlaudanosoline (MS 302)


Production of 3'-Hydroxy-N-methylcoclaurine
6OMT
CNMT
4'OMT
(R,S)-Norlaudanosoline
(R,S)-6-O-Methyl-norlaudanosoline
(R,S)-3'-Hydroxy-N-methylcoclaurine
(R,S)-Reticuline
m/z=178
m/z=123
6-O-Methyl-norlaudanosoline (MS 302)
m/z=192
m/z=123
3'-Hydroxy-N-methylcoclaurine (MS 316)

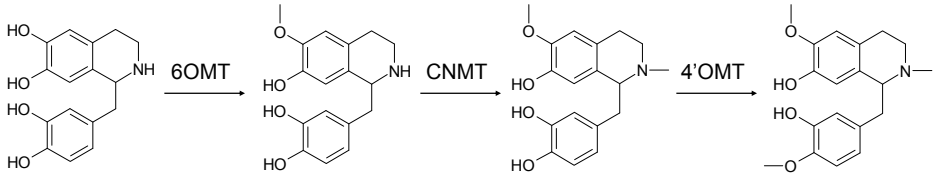

Production of (R,S)-Reticuline
6OMT
CNMT
4'OMT
(R,S)-Norlaudanosoline
(R,S)-6-O-Methyl-norlaudanosoline
(R,S)-3'-Hydroxy-N-methylcoclaurine
(R,S)-Reticuline
Norlaudanosoline m/z=288
Reticuline m/z=330
Norlaudanosoline MS/MS
Intensity (ion counts)
m/z=164
m/z=123
m/z=288
Reticuline MS/MS
m/z=192
m/z=137
m/z=330


Optimization of (R,S)-Reticuline production
6OMT
CNMT
4'OMT
(R,S)-Norlaudanosoline
(R,S)-6-O-Methyl-norlaudanosoline
(R,S)-3'-Hydroxy-N-methylcoclaurine
(R,S)-Reticuline
Enzyme variant dependence:
% Conversion
Ps6OMT/PsCNMT/Ps4'OMT
Ps6OMT/TfCNMT/Ps4'OMT
Tf6OMT/PsCNMT/Ps4'OMT
Tf6OMT/TfCNMT/Ps4'OMT
Ps6OMT/PsCNMT/Tf4'OMT
Ps6OMT/TfCNMT/Tf4'OMT
Tf6OMT/PsCNMT/Tf4'OMT
Tf6OMT/TfCNMT/Tf4'OMT


Optimization of (R,S)-Reticuline production
6OMT
CNMT
4'OMT
(R,S)-Norlaudanosoline
(R,S)-6-O-Methyl-norlaudanosoline
(R,S)-3'-Hydroxy-N-methylcoclaurine
(R,S)-Reticuline
Time dependence:
x10^8
Intensity (ion counts)
72 hrs, OD600 ~10
48 hrs, OD600 ~9.2
24 hrs, OD600 ~8.4
8 hrs, OD600 ~1
23.2
23.6
24.0
24.2
Time (mins)
Strains continue to accumulate (R, S)-Reticuline during growth in stationary phase (OD600 ~ 10).

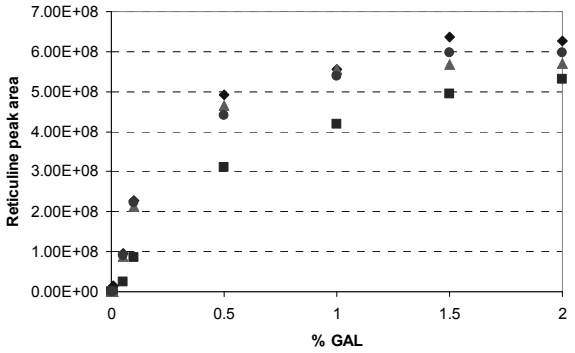

Optimization of (R,S)-Reticuline production
6OMT
CNMT
4'OMT
(R,S)-Norlaudanosoline
(R,S)-6-O-Methyl-norlaudanosoline
(R,S)-3'-Hydroxy-N-methylcoclaurine
(R,S)-Reticuline
Enzyme level dependence:
PGAL1-10
Ps6OMT
TCYC1
PTEF1
PsCNMT
Ps4'OMT
GAL2
2% GAL
GAL
0% GAL
Reticuline peak area
Relative protein expression
Reticuline peak area (x10^-8)
% GAL
Time [min]

## Optimization of (R,S)-Reticuline production




(R,S)-Norlaudanosoline (R,S)-6-*O*-Methyl-norlaudanosoline (R,S)-3'-Hydroxy-*N*-methylcoclaurine (R,S)-Reticuline

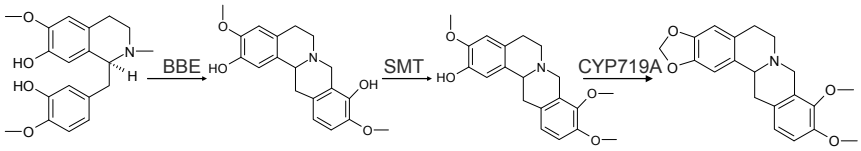
**Enzyme level dependence:**




- $P_{GAL}$Ps6OMT
- $P_{GAL}$Ps4O'MT
- $P_{GAL}$PsCNMT
- $P_{GAL}$PsCNMT $P_{TEF}$Tf4'OMT

***~15% full expression levels from $P_{TET}$ sufficient for maximum conversion***

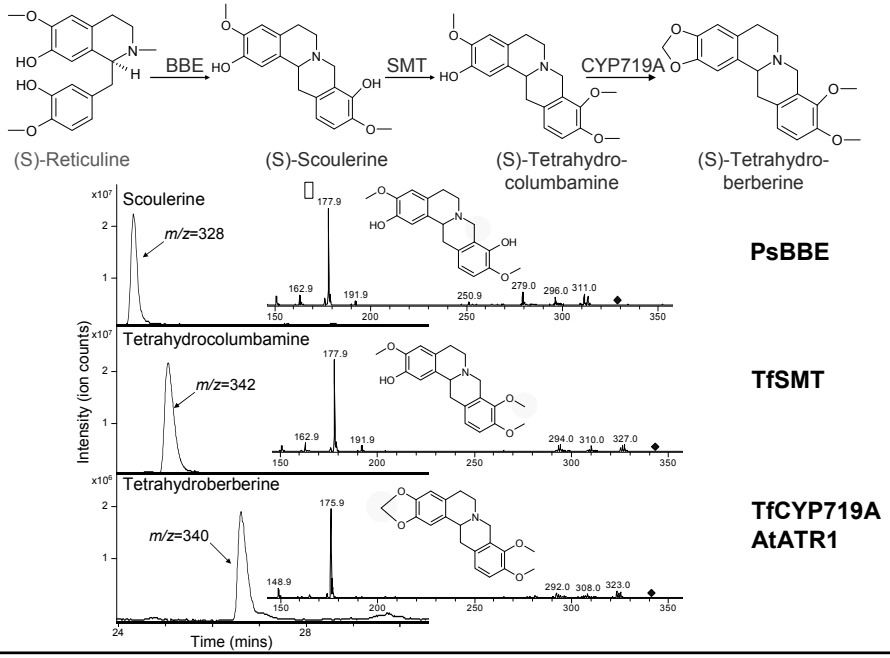
## BIA synthesis beyond reticuline – berberine branch



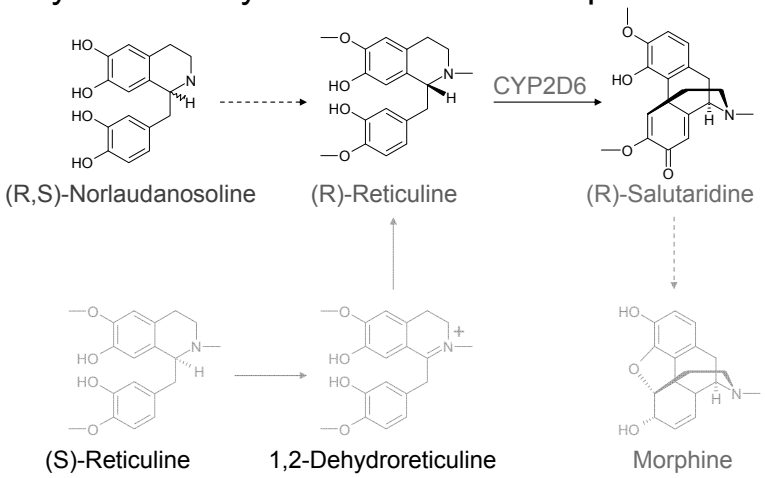

(S)-Reticuline (S)-Scoulerine (S)-Tetrahydro-columbamine (S)-Tetrahydro-berberine

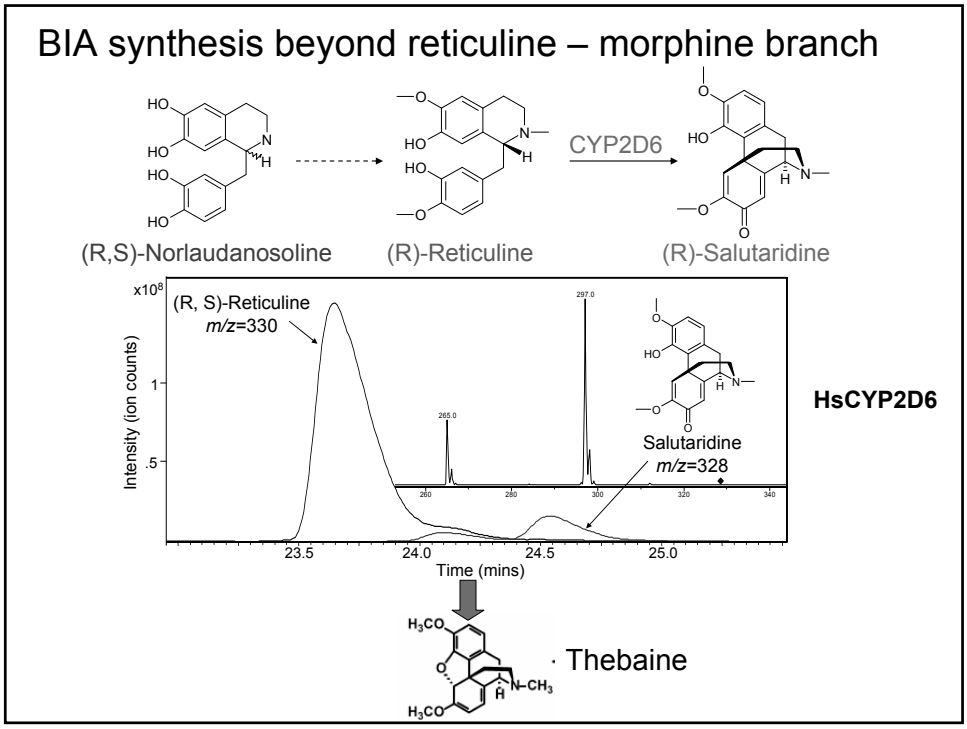
Sanguinarine Berberine

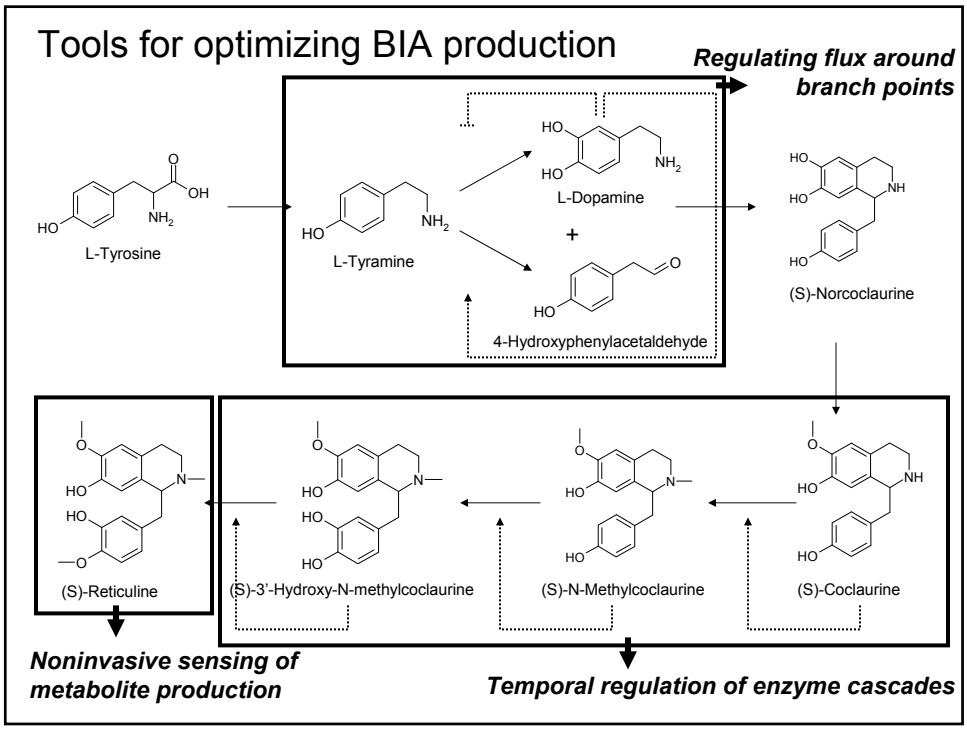
## BIA synthesis beyond reticuline – berberine branch



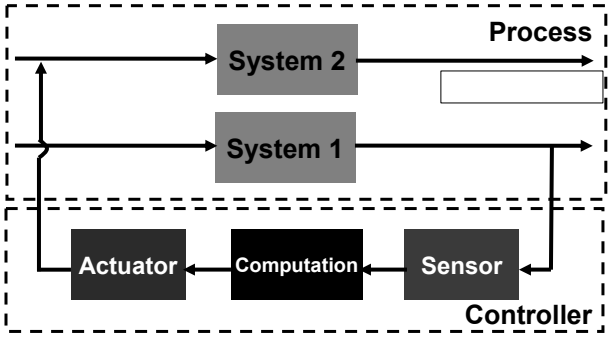

## BIA synthesis beyond reticuline – morphine branch



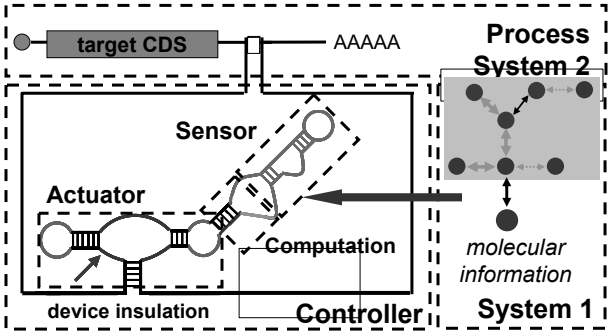

BIA synthesis beyond reticuline – morphine branch
CYP2D6
(R,S)-Norlaudanosoline
(R)-Reticuline
(R)-Salutaridine
(R, S)-Reticuline
m/z=330
Salutaridine
m/z=328
Intensity (ion counts)
Time (mins)
HsCYP2D6
Thebaine


Tools for optimizing BIA production
Regulating flux around branch points
L-Tyrosine
L-Tyramine
L-Dopamine
4-Hydroxyphenylacetaldehyde
(S)-Norcoclaurine
(S)-Reticuline
(S)-3'-Hydroxy-N-methylcoclaurine
(S)-N-Methylcoclaurine
(S)-Coclaurine
Noninvasive sensing of metabolite production
Temporal regulation of enzyme cascades

## General control system

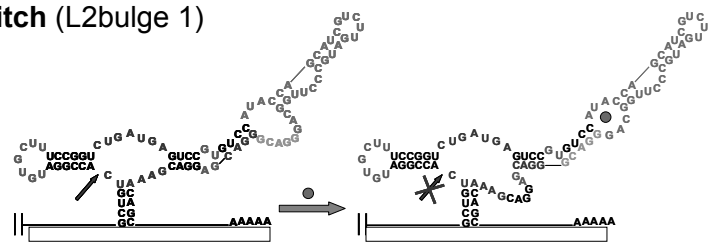


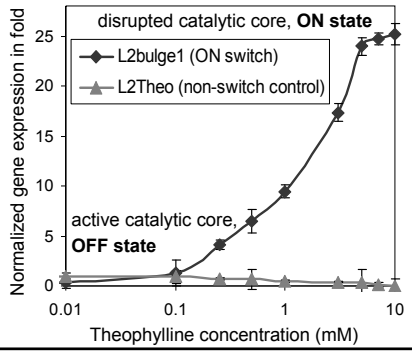

## General biological control system



## A ribozyme switch platform for up-regulating expression

**ON switch** (L2bulge 1)

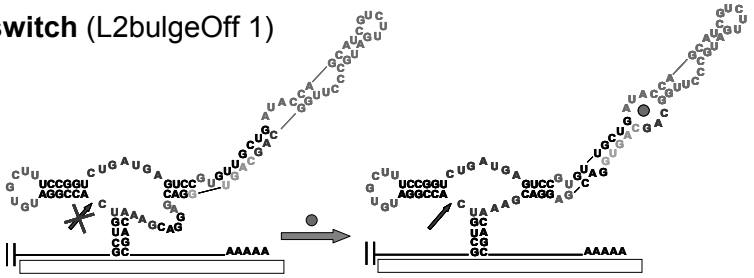


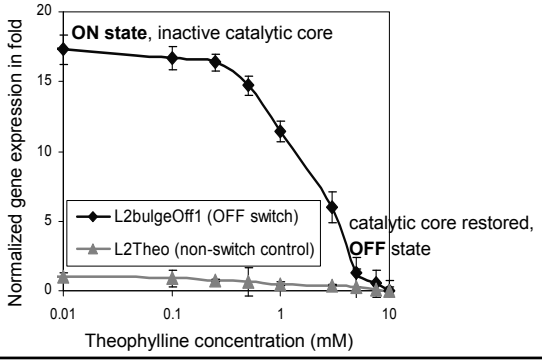




Win MN, Smolke CD. 2007. *PNAS*. 104: 14283-8

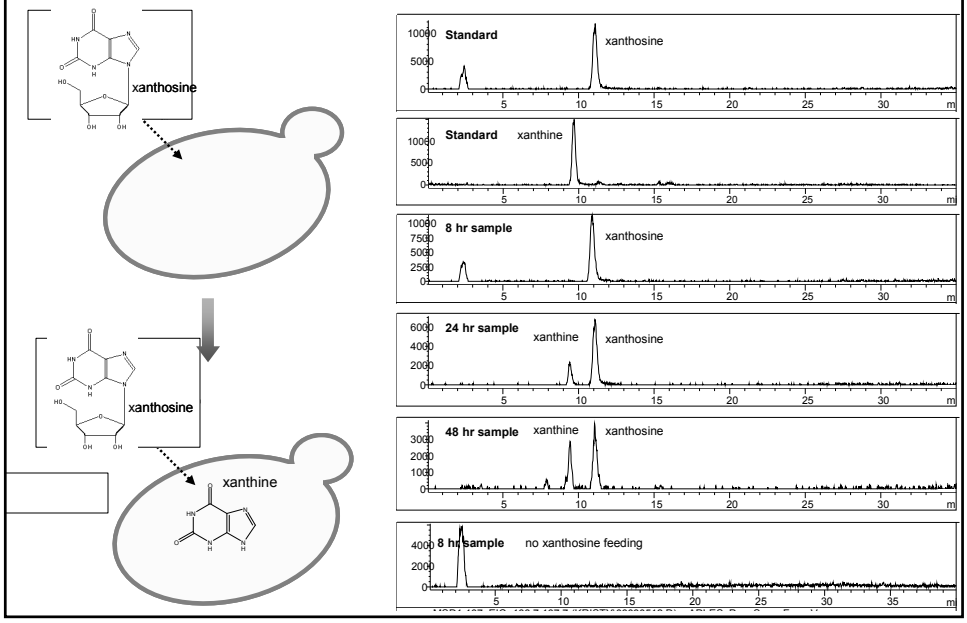
## A ribozyme switch platform for down-regulating expression

**OFF switch** (L2bulgeOff 1)

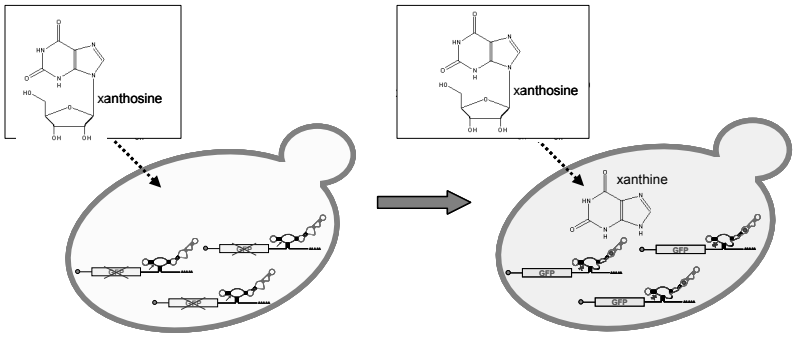


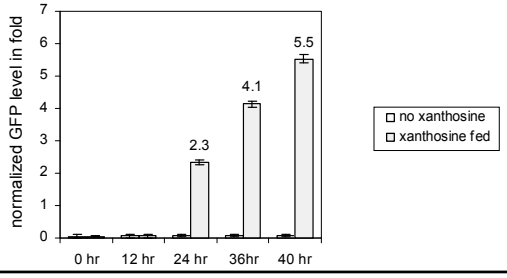
# Integrating RNA devices as noninvasive sensors of metabolite concentration

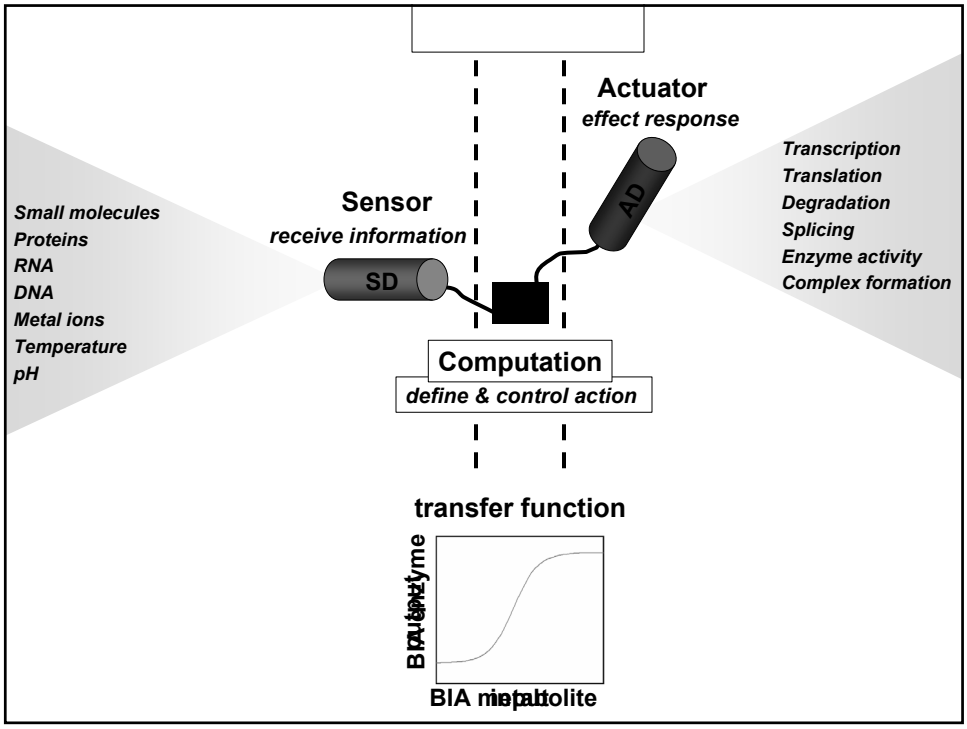



# Integrating RNA devices as noninvasive sensors of metabolite concentration




***Implementation of ON ribozyme switch as an real-time noninvasive xanthine sensor***

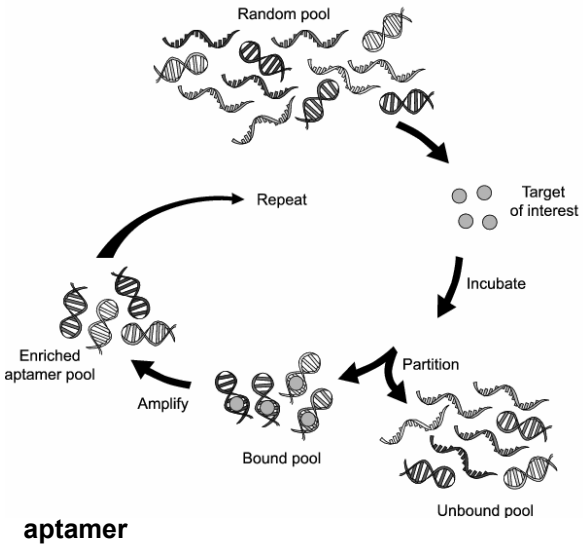





## Scalability challenge: libraries of nucleic acid sensors

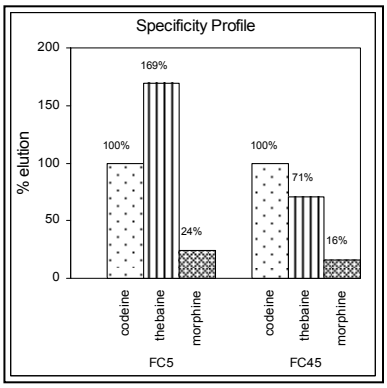
**aptamer**

- bind wide range of ligands
- high specificity and affinity
- generated through *in vitro* selection process

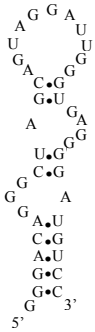



**aptamer selection scheme**

## Specificities of BIA-binding aptamers

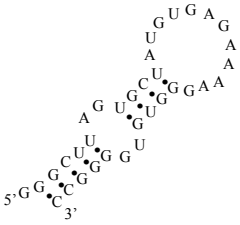



**FC5**

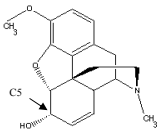



$K_d$ (FC5L) = 4.55±0.14 μM

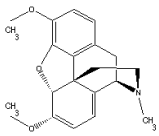
**FC45**



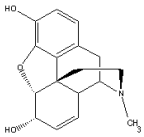

Kd (FC45L) = 2.59 ± 0.09 μM




codeine



thebaine



morphine

***High-throughput SPR-based characterization strategy enables rapid screening for mini-aptamer sequences and structural analysis***

Win MN, Klein JS, Smolke CD. 2006. *Nuc Acids Res.* 34: 5670-82.

## Integrating synthetic metabolic networks and RNA-based control systems

- Metabolic pathway engineering requires a host of tools for optimizing flux and product accumulation
- User-programmed feedback control systems are useful for dynamically controlling flux through pathways
- Developing new genetically encoded tools for receiving, processing, and transmitting molecular information
- Response properties can be programmed to fit the performance specifications of a given application
- These technologies will advance the engineering of more robust cellular systems

# Acknowledgements

**The Smolke Lab**

**Postdoctoral Researchers**

Kevin Hoff

**Graduate Researchers**

Andrew Babiskin
Chase Beisel
**Arwen Brown**
Yvonne Chen
Stephanie Culler
Leo d'Espaux
Katie Galloway
**Kristy Hawkins**
**Joe Liang**
**Josh Michener**
**Michael Siddiqui**
Jay Vowles
**Maung Nyan Win**

**Technicians**

**Midori Greenwood-Goodwin**

**Collaborators**

The Facchini Lab (U. Calgary)
The Jensen Lab (COH)

**Funding Sources**

Arnold and Mabel Beckman Foundation
Caltech
Center for Biological Circuit Design (Caltech)
City of Hope Cancer Center
Defense Advanced Research Projects Agency
Department of Defense (BCRP)
Grubstake Program (OTT Caltech)
Joseph Jacobs Institute for Molecular Engineering for Medicine (Caltech)
National Institutes of Health (NCI)
National Institutes of Health (NIGMS)
National Science Foundation (BES)

# Large-scale Modeling of The Non-linear Enzymatic Reaction Kinetics to Optimize Engineered Pentose Fermentation in *Zymomonas mobilis*

**M. Mete Altintas[1], Christina Eddy[2], Min Zhang[2], James D. McMillan[2] and Dhinakar S. Kompala[1]**

1. Chemical and Biological Engineering Dept, University of Colorado, Boulder
2. Biotechnology Division for Fuels and Chemicals, NREL, Golden, CO

## Background

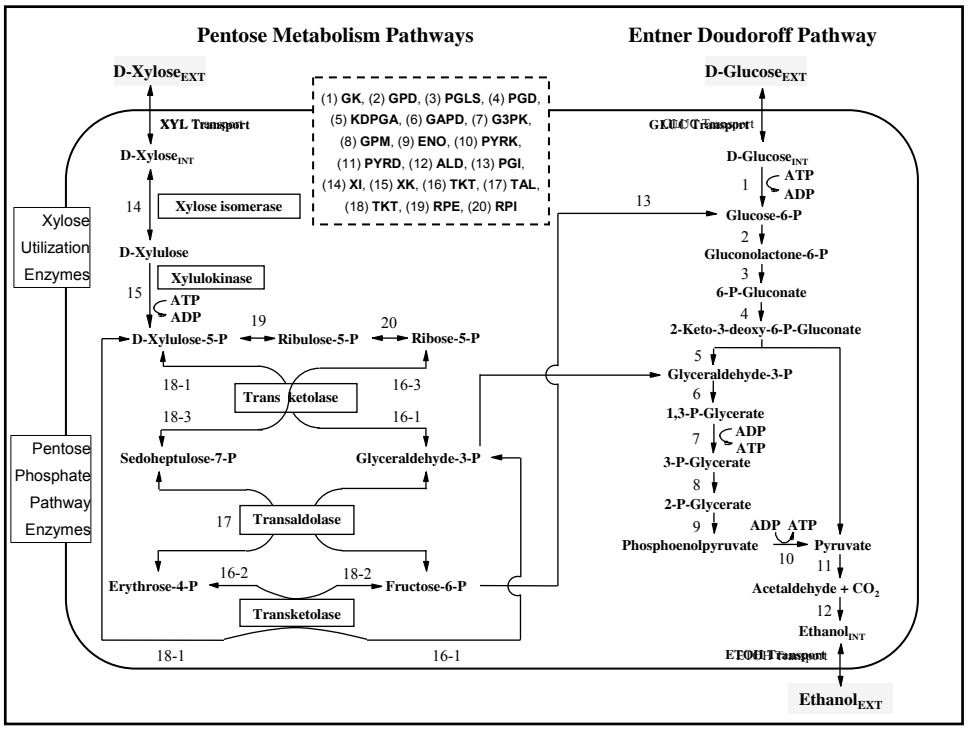
*Zymomonas mobilis* has been engineered with 4 new enzymes to ferment xylose along with glucose and a network of pentose pathway enzymatic reactions interacting with the native glycolytic Entner Doudoroff pathway has been hypothesized.

We investigated this proposed reaction network by developing a kinetic model for all the enzymatic reactions of the pentose phosphate and glycolytic pathways.

Kinetic data on different sugar metabolism rates and enzymatic activity data was used to refine the model parameters available in the literature and validate the proposed reaction network.

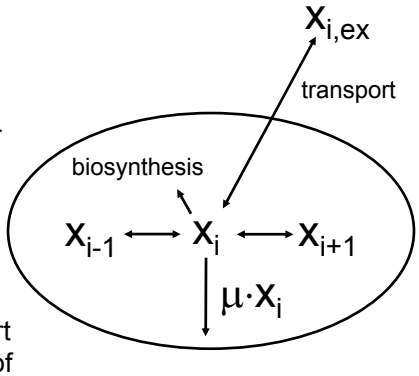
# Objectives

- To investigate the assumed network of enzymatic reactions linking the pentose metabolism and glycolysis pathways.
- To incorporate the non-linear rate expressions for the feedback regulation of enzymatic reactions.
- To find an optimum combination of enzymes needed for maximizing ethanol concentration.



# Features of Kinetic Modeling

- The mechanistic rate equations for each of the enzymatic reactions occurring inside the cell mass.
- The rate equations for the transport of major substrates into cells and of major products out of the cells.




# The *Zymomonas* System

| # | ENZYME | REACTANTS | PRODUCTS | INHIBITORS | MECHANISM |
|---|---|---|---|---|---|
| 1 (→) | Glucokinase | GLUC + ATP | GLUC6P + ADP | GLUC6P | Michaelis-Menten |
| 2 (→) | Glucose-6-P dehydrogenase | GLUC6P + NAD(P) | PGL + NAD(P)H | ATP | Michaelis-Menten |
| 3 (→) | 6-phosphogluconolactonase | PGL | PG | GLUC6P | Michaelis-Menten |
| 4 (→) | 6-phosphogluconate dehydratase | PG | KDPG | G3P | Michaelis-Menten |
| 5 (→) | 2-keto-3-deoxy-6-P-gluconate aldolase | KDPG | GAP + PYR | | Michaelis-Menten |
| 6 (↔) | Glyceraldehyde-3-P dehydrogenase | GAP + NAD | DPG + NADH | | M.-Menten & Rev. |
| 7 (↔) | 3-phosphoglycerate kinase | DPG + ADP | G3P + ATP | | M.-Menten & Rev. |
| 8 (→) | Phosphoglycerate mutase | G3P | G2P | | Michaelis-Menten |
| 9 (→) | Enolase | G2P | PEP | | Michaelis-Menten |
| 10 (→) | Pyruvate kinase | PEP + ADP | PYR + ATP | | Michaelis-Menten |
| 11 (→) | Pyruvate decarboxylase | PYR | ACET | | Michaelis-Menten |
| 12 (↔) | Alcohol dehydrogenase | ACET + NADH | ETOH + NAD | | M.-Menten & Rev. |
| 13 (→) | Phosphoglucose isomerase | FRUC6P | GLUC6P | PG | Michaelis-Menten |
| 14 (→) | Xylose isomerase | XYL | XYLU | | Michaelis-Menten |
| 15 (→) | Xylulokinase | XYLU + ATP | XYLU5P + ADP | | Michaelis-Menten |
| 16 (↔) | Transketolase | XYLU5P + RIB5P | SED7P + GAP | | M.-Menten & Rev. |
| 17 (↔) | Transaldolase | SED7P + GAP | FRUC6P + E4P | | Ping-Pong Bi Bi |
| 18 (↔) | Transketolase | XYLU5P + E4P | FRUC6P + GAP | | M.-Menten & Rev. |
| 19 (↔) | Ribulose 5-phosphate epimerase | XYLU5P | RIBU5P | | M.-Menten & Rev. |
| 20 (↔) | Ribose 5-phosphate isomerase | RIBU5P | RIB5P | | M.-Menten & Rev. |

## Approach

- The kinetic model developed to describe the *Zymomonas* system is comprised of 24 rate expressions and 24 balance equations.
- All the derivatives of balance equations were set to zero to calculate a steady state. The resulting system of non-linear equations was solved numerically for the steady state metabolite concentrations.
- The amounts of 5 enzymes (PGI, XI, XK, TKT and TAL) were varied such that the total of ED and PP pathway enzymes ranges from 42% to 66% of the total cellular protein.
- The $k_{cat}$ values reported at varying temperatures are normalized by using the 'glucose consumption vs. temperature' table presented by Scopes and Griffiths-Smith (1986) to estimate the $k_{cat}$ values at 30°C.

## Assumptions and Conditions

- The $k_{cat}$ and $K_m$ values used in the model were collected from the literature at varying pH's.
- The concentrations of ATP, ADP, NAD, NADH and NADP were assumed to be constant and equal to 2, 1, 1.5, 1 and 0.5 mM, respectively.
- A constant level of gene expression was assumed, i.e. the enzyme levels that were measured in the wild-type strain were used.
- For the heterologous enzymes, $k_{cat}$ and $K_m$ values from *E. coli* were used.

✯ ✯ ✯

- Xylose and glucose are constant at 5.0 and 0.875 g/L, respectively.
- The chemostat is operated at a constant dilution rate of 0.05 $h^{-1}$.
- The cell mass concentration in the bioreactor is constant at 1.0 g/L.
- The cytoplasmic volume of *Zymomonas* cells growing in the presence of glucose and xylose were taken to be 2.2 ml/g dcm.

# Sample Rate Expression

$$\upsilon_1 = \frac{k_{cat} \cdot [E]_T \times \left(\frac{GLUC}{K_{m,GLUC}}\right) \times \left[\frac{ATP}{K_{m,ATP} \times \left(1 + \frac{GLUC6P}{K_{i,GLUC6P}}\right)}\right]}{1 + \frac{GLUC}{K_{m,GLUC}} + \frac{ATP}{K_{m,ATP}} + \frac{GLUC \times ATP}{K_{m,GLUC} \times K_{m,ATP} \times \underbrace{\left(1 + \frac{GLUC6P}{K_{i,GLUC6P}}\right)}_{\text{Competitive product inhibiton term}}} + \frac{GLUC6P}{K_{i,GLUC6P}}}$$

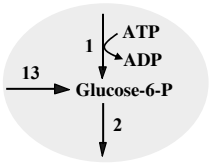
*2 Sub.s + 2 Prod.s*

*Michaelis-Menten Mech.*

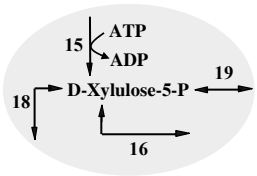
*Irreversible Rxn.*

- $K_{cat}$, $K_m$ and $K_i$ values are reported in the literature
- $\upsilon_{max} = k_{cat} \cdot [E]_T$
- $[E_T]$ = (g enzyme/g total protein) x (g protein/g dry cell)

# Representative Balance Equations




$$\frac{d[GLUC6P]}{dt} = \underbrace{(\upsilon_1 + \upsilon_{13} - \upsilon_2)}_{\text{rate expressions}} - \underbrace{\mu \cdot [GLUC6P]}_{\text{dilution due to growth}}$$



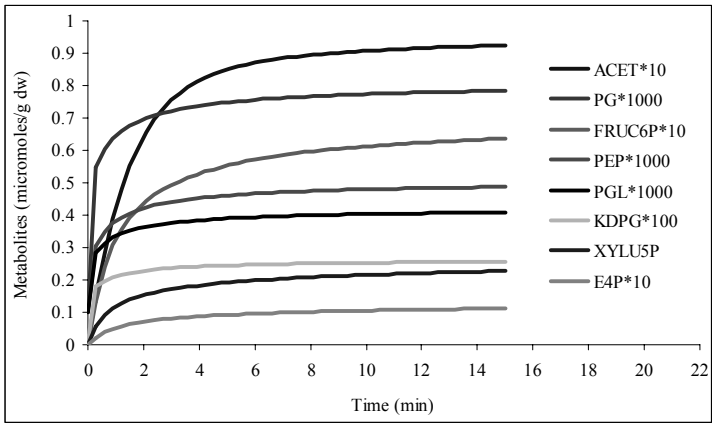

$$\frac{d[XYLU5P]}{dt} = (\upsilon_{15} - \upsilon_{16} - \upsilon_{18} - \upsilon_{19}) - \mu \cdot [XYLU5P]$$

METABOLITES

| Metabolites | Dynamic Balance Equations (d[metabolite]/dt = …) | Steady State Intracellular Concentrations (mmol/g dry weight) |
|---|---|---|
| GLUC | $(\upsilon_{GLUC} - \upsilon_1) - (\mu \cdot GLUC)$ | $4.48 \times 10^{-3}$ |
| GLUC6P | $(\upsilon_1 + \upsilon_{13} - \upsilon_2) - (\mu \cdot GLUC6P)$ | $4.70 \times 10^{-3}$ |
| PGL | $(\upsilon_2 - \upsilon_3) - (\mu \cdot PGL)$ | $4.15 \times 10^{-4}$ |
| PG | $(\upsilon_3 - \upsilon_4) - (\mu \cdot PG)$ | $7.97 \times 10^{-4}$ |
| KDPG | $(\upsilon_4 - \upsilon_5) - (\mu \cdot KDPG)$ | $2.61 \times 10^{-3}$ |
| GAP | $(\upsilon_5 + \upsilon_{16,1} - \upsilon_6 - \upsilon_{17} - \upsilon_{18,1}) - (\mu \cdot GAP)$ | 0.26 |
| BPG | $(\upsilon_6 - \upsilon_7) - (\mu \cdot BPG)$ | 0.11 |
| G3P | $(\upsilon_7 - \upsilon_8) - (\mu \cdot G3P)$ | $2.15 \times 10^{-3}$ |
| G2P | $(\upsilon_8 - \upsilon_9) - (\mu \cdot G2P)$ | $4.68 \times 10^{-3}$ |
| PEP | $(\upsilon_9 - \upsilon_{10}) - (\mu \cdot PEP)$ | $4.95 \times 10^{-4}$ |
| PYR | $(\upsilon_5 + \upsilon_{10} - \upsilon_{11}) - (\mu \cdot PYR)$ | $1.28 \times 10^{-2}$ |
| ACET | $(\upsilon_{11} - \upsilon_{12}) - (\mu \cdot ACET)$ | 0.14 |
| ETOH | $(\upsilon_{12} - \upsilon_{ETOH}) - (\mu \cdot ETOH)$ | 192.10 |
| XYL | $(\upsilon_{XYL} - v_{14}) - (\mu \cdot XYL)$ | 64.92 |
| XYLU | $(\upsilon_{14} - \upsilon_{15}) - (\mu \cdot XYLU)$ | 7.92 |
| XYLU5P | $(\upsilon_{15} + \upsilon_{18,1} - \upsilon_{16,1} - \upsilon_{19}) - (\mu \cdot XYLU5P)$ | 0.25 |
| RIBU5P | $(\upsilon_{19} - \upsilon_{20}) - (\mu \cdot RIBU5P)$ | 0.27 |
| RIB5P | $(\upsilon_{16,3} + \upsilon_{20} - \upsilon_{18,3}) - (\mu \cdot RIB5P)$ | 0.52 |
| SED7P | $(\upsilon_{18,3} - \upsilon_{16,3} - \upsilon_{17}) - (\mu \cdot SED7P)$ | 7.94 |
| E4P | $(\upsilon_{16,2} + \upsilon_{17} - \upsilon_{18,2}) - (\mu \cdot E4P)$ | $1.25 \times 10^{-2}$ |
| FRUC6P | $(\upsilon_{17} + \upsilon_{18,2} - \upsilon_{13} - \upsilon_{16,2}) - (\mu \cdot FRUC6P)$ | $6.67 \times 10^{-2}$ |

# Concentration Trajectories

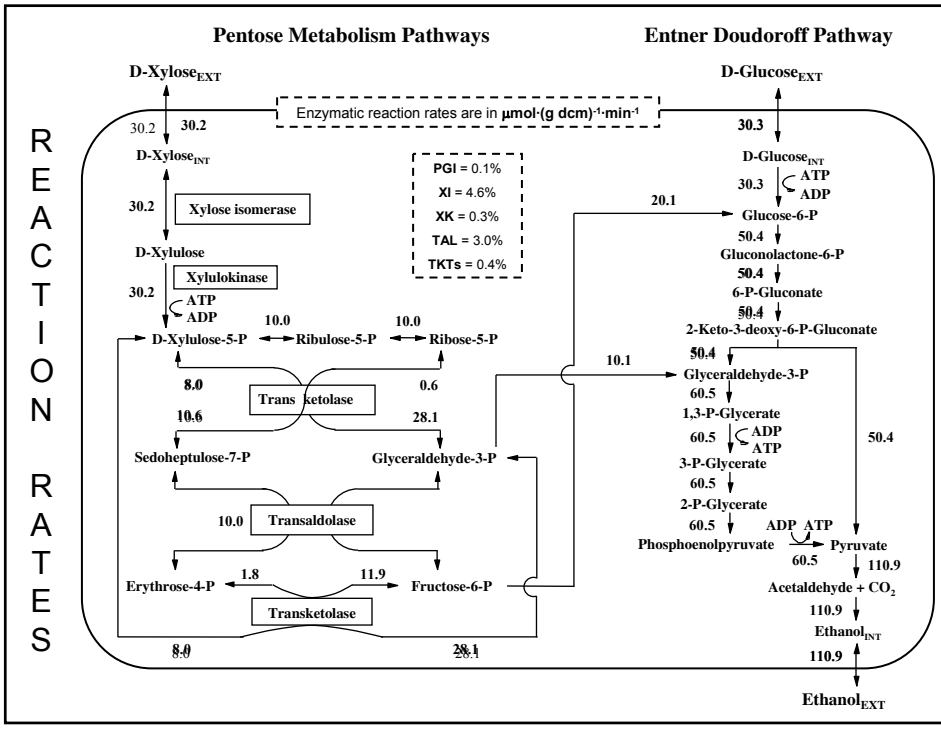
**PGI** = 0.1%
**XI** = 4.6%
**XK** = 0.3%
**TAL** = 3.0%
**TKTs** = 0.4%




**ACET: acetaldehyde**
**PG: 6-phosphogluconate**
**FRUC6P: fructose-6-phosphate**
**PEP: phosphoenolpyruvate**
**PGL: 6-phosphogluconolactone**
**KDPG: 2-keto-3-deoxy-6-phosphogluconate**
**XYLU5P: xylulose-5-phosphate**
**E4P: erythrose-4-phosphate**

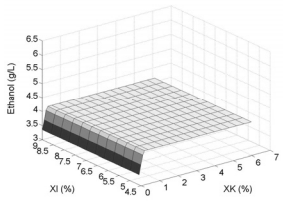
# Maximum Ethanol Concentration & Optimal Ethanol Production Efficiency

| PGI (%) | XI (%) | XK (%) | TAL (%) | TKTs (%) | | Total Key Enzyme Expression Level (%) | ETOH (g/L) | ETOH / Key Enzyme Expression Level |
|---|---|---|---|---|---|---|---|---|
| | | | | TKT | TKT-$C_2$ | | | |
| ***Objective: Maximize ethanol concentration*** | | | | | | | | |
| **1.1** | **9.0** | **3.7** | **6.1** | **0.8** | **4.1** | 24.8 | **6.1258** | 0.2470 |
| 1.3 | 9.0 | 3.7 | 6.1 | 0.9 | 4.2 | 25.2 | **6.1258** | 0.2431 |
| 1.1 | 9.0 | 3.7 | 6.0 | 0.9 | 4.3 | 25.0 | **6.1258** | 0.2450 |
| 1.4 | 9.0 | 4.1 | 5.7 | 0.8 | 3.7 | 24.7 | **6.1258** | 0.2480 |
| 1.3 | 9.0 | 3.9 | 5.9 | 0.8 | 3.9 | 24.8 | **6.1258** | 0.2470 |
| ***Objective: Optimize ethanol production efficiency*** | | | | | | | | |
| **0.1** | **4.6** | **0.3** | **3.0** | **0.1** | **0.3** | 8.4 | 6.1202 | **0.7286** |
| 0.1 | 4.6 | 0.3 | 3.1 | 0.1 | 0.2 | 8.4 | 6.1202 | **0.7286** |
| 0.1 | 4.6 | 0.3 | 2.9 | 0.2 | 0.3 | 8.4 | 6.1191 | **0.7285** |
| 0.1 | 4.6 | 0.3 | 3.1 | 0.1 | 0.3 | 8.5 | 6.1207 | **0.7201** |
| 0.1 | 4.6 | 0.3 | 3.2 | 0.1 | 0.2 | 8.5 | 6.1207 | **0.7201** |

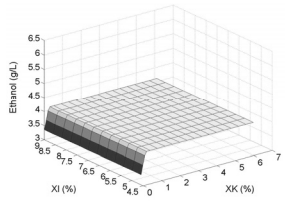


# Impact of PGI, XI and XK Concentrations

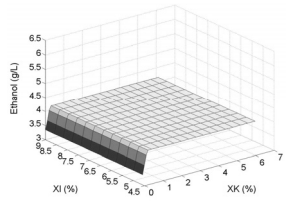
PGI=0%, TAL=0.1%, TKTs=0.1%



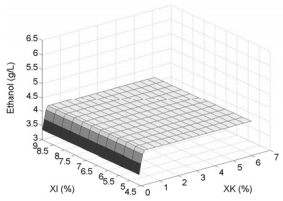

PGI=0%, TAL=6.1%, TKTs=0.1%



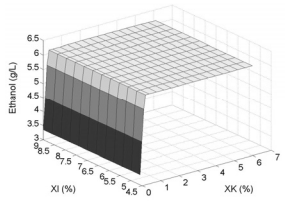

PGI=0%, TAL=6.1%, TKTs=6.1%



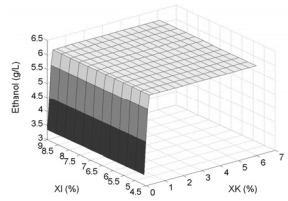

PGI=0.1-1.5%, TAL=0.1%, TKTs=0.1%




PGI=0.1-1.5%, TAL=6.1%, TKTs=0.1%



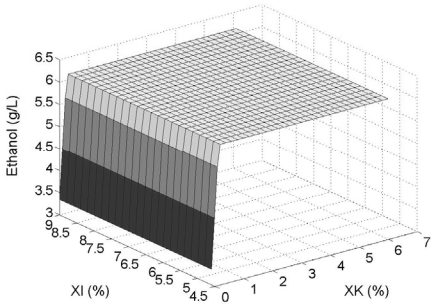

PGI=0.1-1.5%, TAL=6.1%, TKTs=6.1%

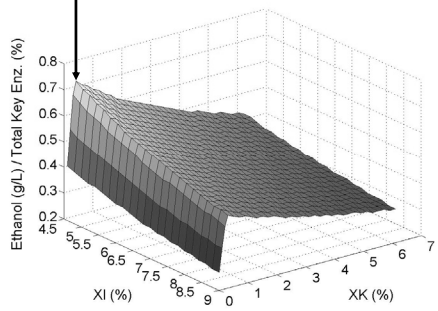



# Impact of XI and XK Concentrations

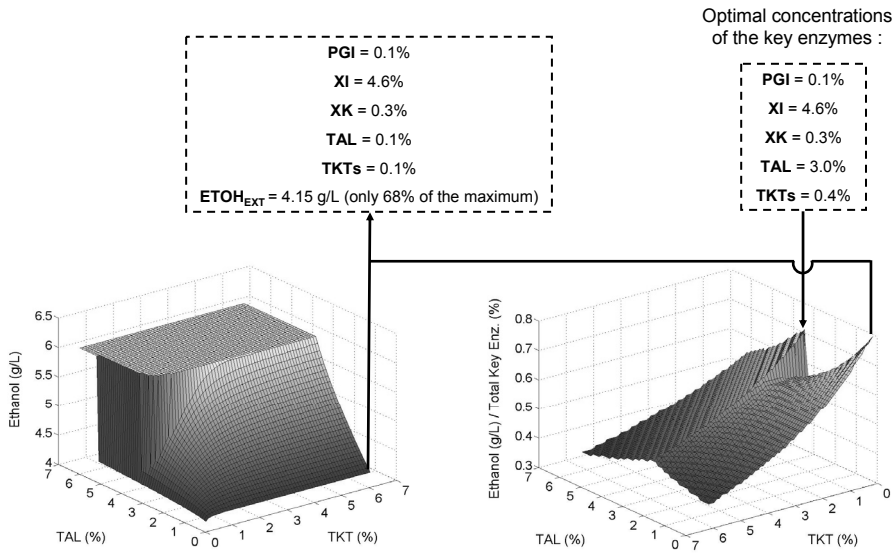
Optimal concentrations of the key enzymes :

**PGI** = 0.1%
**XI** = 4.6%
**XK** = 0.3%
**TAL** = 3.0%
**TKTs** = 0.4%



# Impact of TAL and TKT Concentrations




# Results

- The kinetic model simulations at each enzyme combination produce enormous amount of information, including the steady state metabolite concentrations as well as the dynamic variations in the individual reaction rates for each enzymatic reaction and metabolite.
- Amongst the 5 enzymes whose amounts were tested in terms of the ethanol production, it was found that TAL, XI and XK were the most important enzymes in terms of their effect on the extracellular ethanol concentration. Their presence at sufficient levels eliminates pathway bottlenecks and significant accumulation of intermediate metabolites.
- The relatively low amounts of native PGI and heterologous TKT are sufficient to enable maximal ethanol production.

## Results

- Since the reaction rates in the glycolysis pathway are higher than the rates in PP pathway, ‘glyceraldehyde 3-P’ is channeled quickly into the glycolysis pathway.
- Following the flow of ‘glyceraldehyde 3-P’ into the glycolysis pathway, the amounts of ‘erythrose 4-P’ and ‘fructose 6-P’ drops dramatically when TAL and TKT enzymes are available.
- The relatively slow rate of reactions in the PP pathway makes XI and XK enzymes important in terms of their ability to control the reactions immediately following xylose transport into the cell.

## Conclusion

- The kinetic modeling strategy allows us to incorporate known information about non-linear rate equations and regulation by other metabolites in determining the optimum combination of heterologous enzymes to maximize pathway efficiency.
- The model enables us to compare the enzymatic reaction rates between the pentose utilization (PP) and glycolytic (ED) pathways and calculate the metabolic flux for each enzymatic reaction at all the assumed levels of heterologous enzymes.

## Acknowledgements

This work was supported by the Office of the Biomass Program of the U.S. Department of Energy (DOE) and U.S. Department of Agriculture (USDA).

# Development of tolerant and other complex phenotypes for biofuel production

**E. Terry Papoutsakis, Univ. Delaware**

**& Keith Alsaker, Jacob Borden, Carles Paredes, Ryan Bruskiewicz**

**Dept. of Chemical & Biological Engineering**

**Northwestern University, Evanston, IL**

**Acknowledgements**

**Funding: NSF, DOE (& JB by NIH Biotechnology Training Grant)**

## Complex phenotypes

- **Involve many genes, which are frequently or mostly not known in their entirety (class I)**
- **Even if we know the genes involved, we do not understand their dynamic behavior to allow us to alter or develop the phenotype (Class II)**

## Examples?

- **Cells that can tolerate harsh bio-processing conditions beyond what may currently exist in nature? (Class I)**
- **Like ethanol toxicity: Can we create cells that can tolerate 30% ethanol? Why not? What would it take? What do we know? What is the problem? Biophysical?**

## More examples…

- **Plants with dramatically improved photosynthetic machinery that produce cellulosic biomass twice or thrice as fast as what is currently possible? (Class II)**
- **Cells that utilize quickly and effectively *(5-10x better than currently known)* cellulose or xylans to produce one major product (such as ethanol or butanol or other biofuels or commodity chemicals) (Class II or I?)**

## Which way to solve such problems?

- An *ab initio* cellular design?
  - How far can we go based on what we know?
  - What is holding us back? Synthetic biochemistry/biology or knowledge?
- A hybrid first approach: some knowledge based, some empirical or semi- empirical?
  - What is the evidence that it will work? What hypotheses must be tested to that effect?
- Luck, and trial & error empiricism like mostly currently practiced?

## Our Research Goals

- Identify genes and cellular programs in clostridia affected by solvent (e.g., butanol) and carboxylic acid (butyrate & acetate) stress in order to identify:
  - Specific and general stress regulons
  - Genes which may impart solvent/acid tolerance

## Applications: Bioprocessing for:

- Solvent-production: fermentations
- Biocatalysis
- Bioremediation

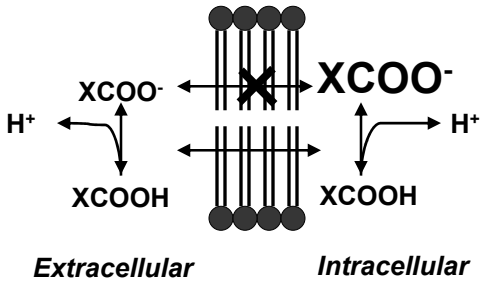
## Production of solvents via fermentation provides a green alternative to petrochemicals

- 1910s-1950s: butanol produced by anaerobic fermentation: *C. acetobutylicum* (acetone, butanol, ethanol)
  - Fermentation yields 1.5% butanol (very low)
  - A limiting step in final titers: product inhibition (toxicity)
- Two phase fermentations
  - Exponential growth: production of butyrate, acetate
  - Stationary phase: uptake of butyrate and acetate, production of acetone, butanol, and ethanol
- TOXICITY: butanol (but also butyrate & acetate & interactions)

## THEORY: Mechanisms of Solvent Toxicity

- Inhibition of growth and glucose uptake
- Disruption of membrane integrity
  - Loss of membrane $\Delta pH$ and $\Delta\Psi$
  - Loss of ATP production efficiency
  - Cells adapt slowly by altering membrane fluidity
    - Adaptation may inhibit membrane function

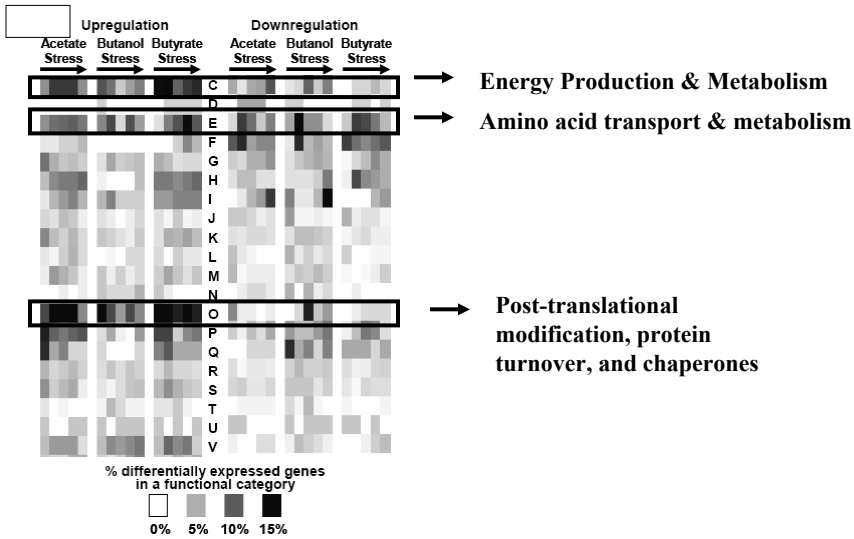
## Uncoupling THEORY of Carboxylic Acid Toxicity: Undissociated acids cross cellular membrane and acidify the cytoplasm



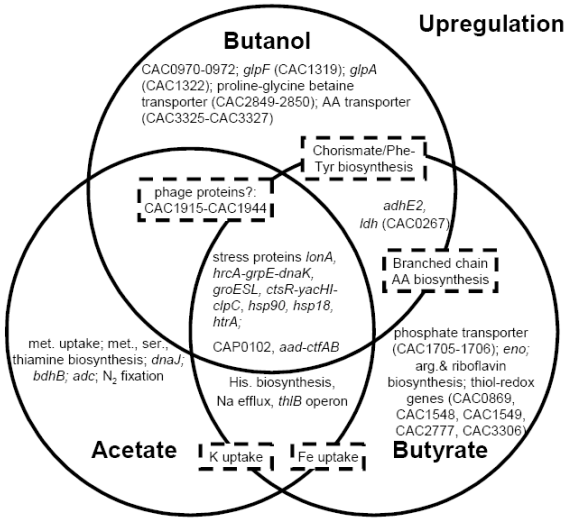

Russell and Diez-Gonzalez, *Adv. Microb. Phys.*, 1998, 208

- **Is it that simple?**

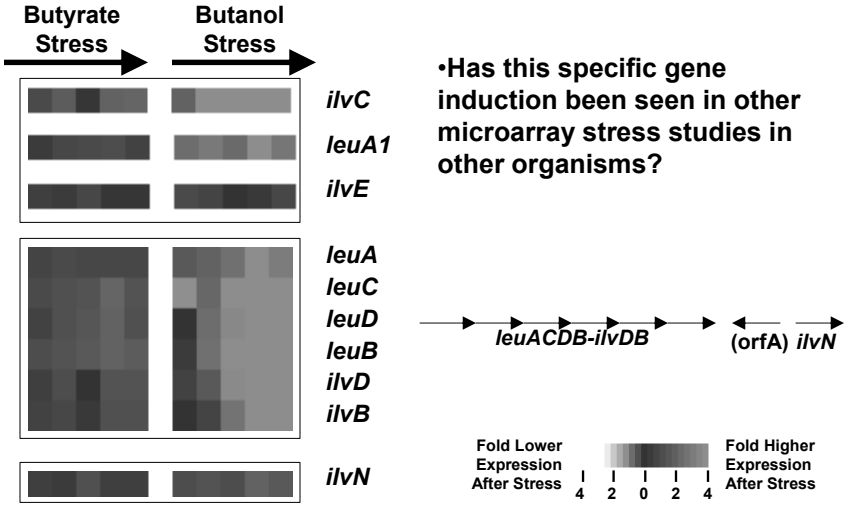
## Ontological Analysis of multi-stress responses (based on Genome-scale μ-array analysis)



# Common & Specific Stress Responses
## (based on Genome-scale μ-array analysis)



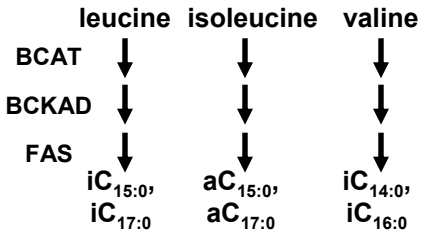

# Differential Expression of Branched Chain Amino Acid Synthesis Genes




- Has this specific gene induction been seen in other microarray stress studies in other organisms?

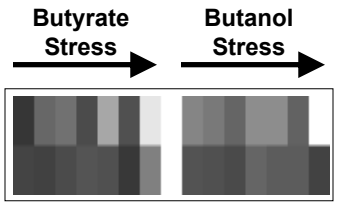
- Adaptation due to cold-shock of *B. subtilis* includes:
  - decreased membrane fluidity
  - downregulation of branched-chain amino acids (Kaan et al. *Microbiol.*, 2002)

- *Bacillus subtilis* can convert branched-chain amino acids into branched-chain fatty acids




- *C. acetobutylicum* increase membrane fluidity in response to butanol (Vollherst-Schneck et al. *J. Bacteriol.*, 1984)
- Synthesis of branched-chain fatty acids in *C. acetobutylicum* may help cells adapt to metabolite stress

# Genes Specifically Induced by Butanol Stress



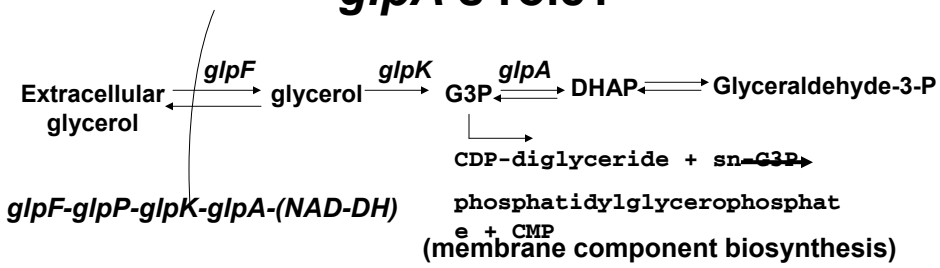

*glpF-glpP-glpK-glpA-(NAD-DH)*

Glycerol-3-phosphate dehydrogenase, *glpA*

NAD-dependent dehydrogenase

- In *E. coli*, homolog *glpC* is associated with solvent tolerance (Shimizu, *AEM*, 2005), although cannot act alone…

# *glpA*'s role?




- **What is the common role of glycerol-3-phosphate dehydrogenase in solvent tolerance?**
  - **First step in P-lipid biosynthesis as the need for an altered membrane composition develops (H. Goldfine et al.)**
  - **Yeast: *glp* genes play role in osmotolerance (Brisson, *BioEssays*, 2001)**

# Universally Upregulated Genes by Acetate, Butyrate, and Butanol

- **Stress proteins**
  - **Chaperones: *dnaKJ*, *groESL*, *clpC, hsp90, hsp18***
  - **Protease: *lonA***
  - **Benefits of *groESL* overexpression in enhancing solvent tolerance has been established by our group[1,2]**
- ***Establishing protein stability and functionality seems to be key in responding to toxic stress***

- ***THUS, stress response can (sometimes) predict stress tolerant phenotypes***

1. Tomas, CT, NE Welker, ET Papoutsakis. 2003. *Appl. Environ. Microbiol.* 69: 4951-4965.
2. Tomas, CT, J Beamish, ET Papoutsakis. 2004. *J. Bacteriol.* 186: 2006-2018.

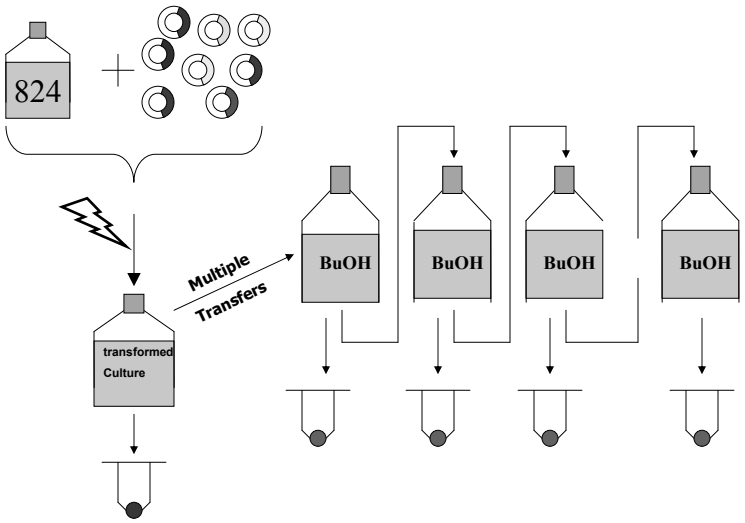
## A lot of new information…

- **But no obvious high-throughput way to test and use such information for achieving strong tolerant phenotypes…**
- **Need then for such (empirical for now) tools, that can be more broadly used for developing other complex phenotypes**
  - **Genomic or expression libraries have been used extensively to identify genes or loci (or their variants) imparting a selectable phenotype, incl. tolerant ones**
  - **What is NOT known is how well and completely that works for either SINGLE loci/genes (let alone for INTERACTING MULTIPLE LOCI) and if these genes/loci are related to stress regulons**

## Library approach

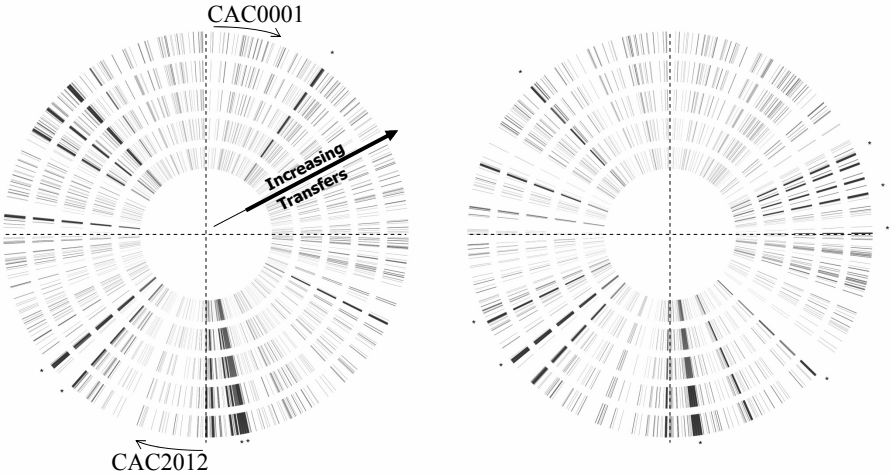
- Several types of libraries
- Have applied them for identifying tolerance to butanol and butyrate and have tested some of the identified genes

## Library Selection




## Assessment by high resolution HT (μ-array or deep sequencing analysis): stochasticity vs determinism

**Biological Replicate 1** **Biological Replicate 2**




Genes color coded according to percentile rank in a given transfer: the top 5% are red, 5-33% are green, and 33-100% are gray

## Several of the identified genes

- Impart solvent tolerance alone…
- But better as a group in a mixed population rather than in pure recombinant strains (extracellular molecules? Cell-to-cell communication?)
- Novel transcriptional regulators…

## Other interesting findings

- **Butyrate-tolerance studies lead consistently to enrichment of IR (non-coding) DNA of the rRNA locus. Regulon has been determined by μ-array analysis but mechanism not known yet: a srRNA or what?**
- **Based on these identified DNAs, tolerance to butyrate imparts also tolerance to all tested carboxylic acids and leads to higher butanol formation (and tolerance?)**

## To summarize, Library approach

- **Useful for identifying many loci, but not all**
  - **Little overlap with BuOH stress regulon**
  - **Significantly, it missed chaperons (why?) and likely more tolerance genes…**
  - **Appears to identify more transcriptional regulators (solo players)**
- **Misses large, multigenic programs…**
- **It is not strictly deterministic, and depends on**
  - **Selection assay (select for growth only?)**
  - **Application mode of selection assay**
  - **Insert size and regulation of gene expression**

## Stress-regulon approach works best…

- When there is a dormant adaptive response such as in, e.g., clostridial adaptation to butanol and butyrate accumulation
- For identifying potential target programs and pathways

## Current & future work in my lab: methods

- **HT methods to capture and combine distant multi-locus effects and allogeneic traits (for tolerant and other phenotypes, …**
- **Coupled with "enhanced" , "temporarily-induced" recombination, …**
- **to develop strains that quickly and simultaneously directly and efficiently ferment cellulose and xylans without enzymatic pre-treatment**

## …and exploration of the relationship between differentiation (sporulation) and tolerance?

- Mature or even immature spores are extremely resistance to chemicals, radiation, everything,…
- Can we use "Differentiation Engineering" to freeze "differentiation" stage at cellular type which is tolerant and a good producer?

DOE GENOMICS:GTL SYSTEMS BIOLOGY FOR ENERGY AND ENVIRONMENT
OFFICE OF SCIENCE U.S. DEPARTMENT OF ENERGY

Oregon State UNIVERSITY OSU

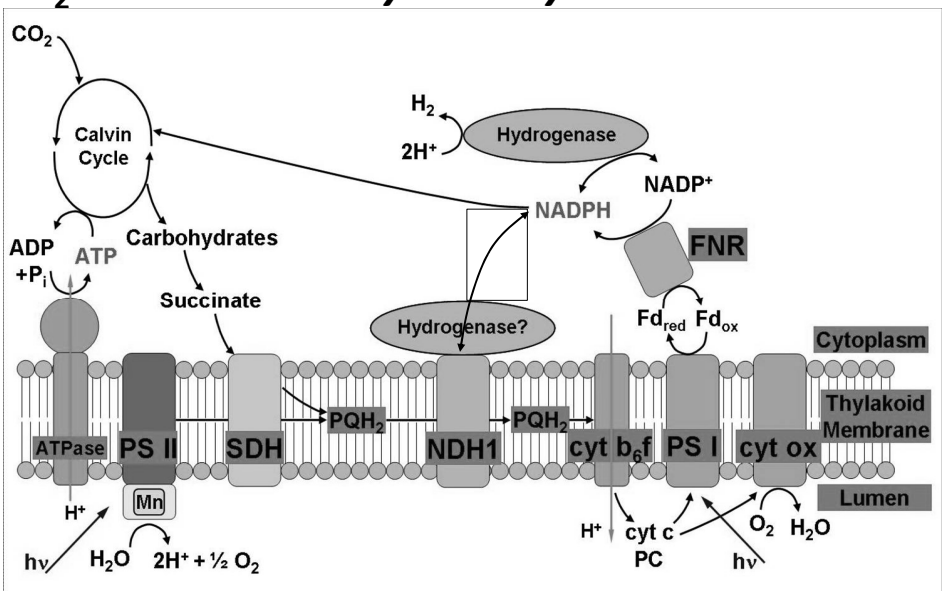
# Metabolic Engineering of Light and Dark Biochemical Pathways in Wild-Type and Mutant *Synechocystis* PCC 6803 Strains for Maximal, 24-Hour Production of Hydrogen Gas

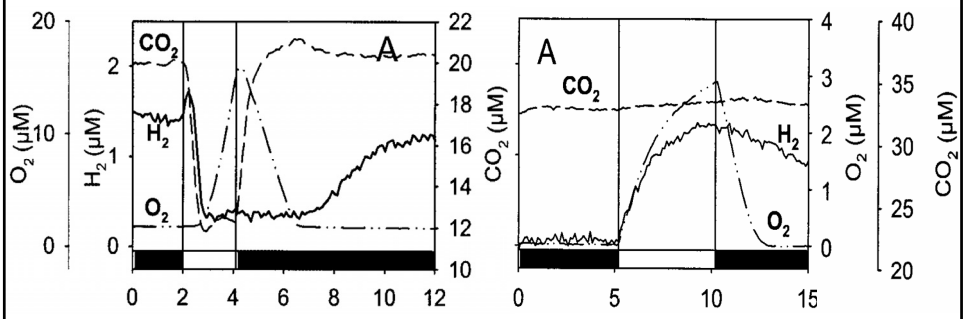
**Roger L. Ely**
**Frank W.R. Chaplen**

**Oregon State University**

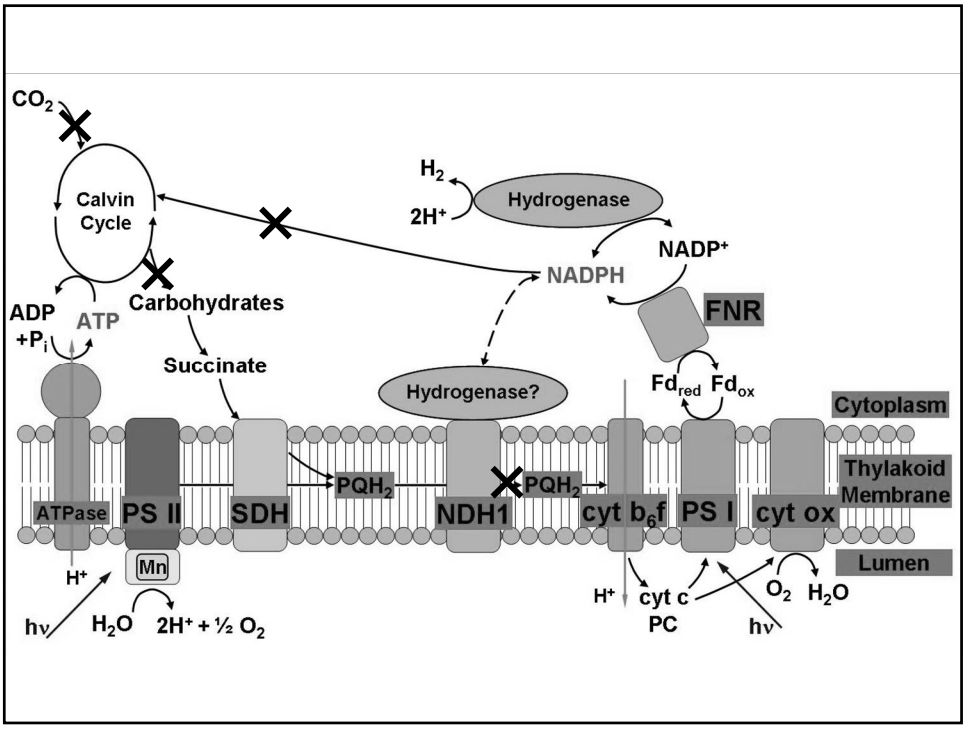
**DOE-BER GTL-MEWG Annual Awardee Workshop**
**Feb 11, 2008**

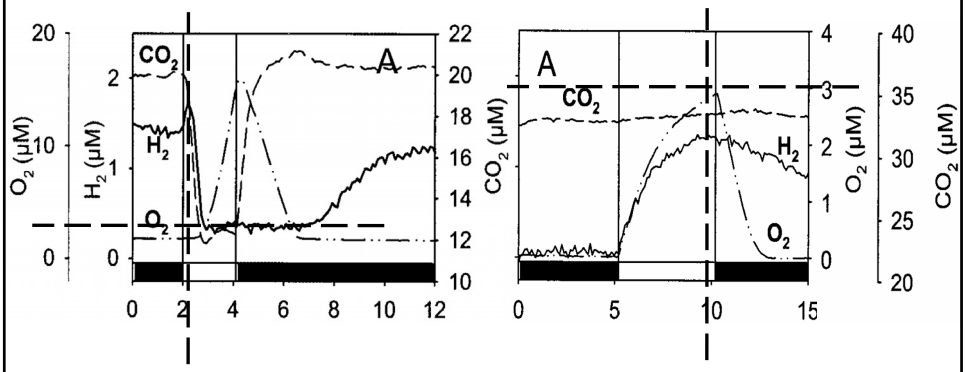
## $H_2$ Production in *Synechocystis* PCC 6803






Source: Cournac, L., G. Guedeney, G. Peltier, and P. M. Vignais. 2004. J. Bacteriol. 186(6):1737–1746

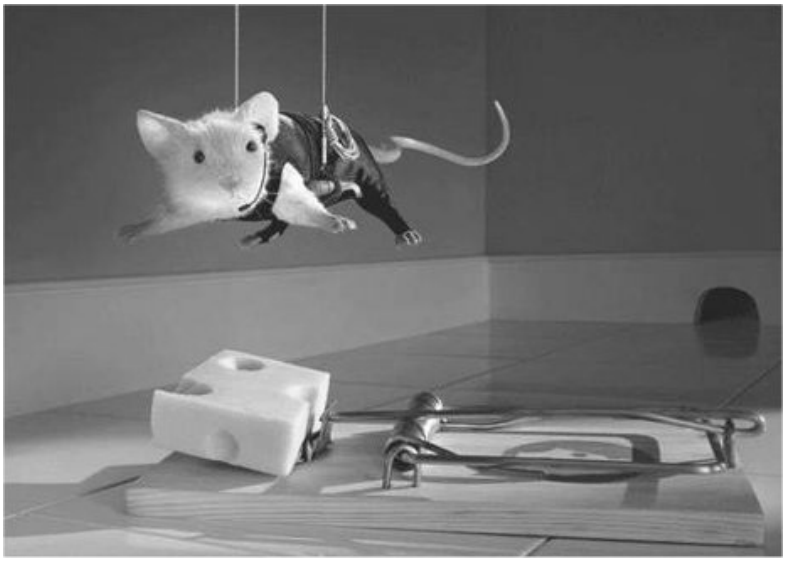


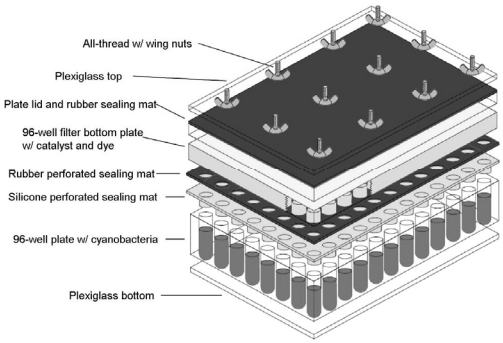


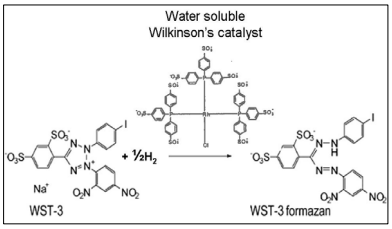

Source: Cournac, L., G. Guedeney, G. Peltier, and P. M. Vignais. 2004. J. Bacteriol. 186(6):1737–1746

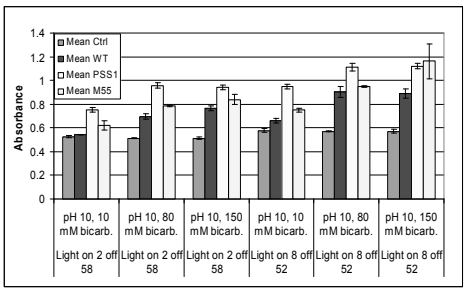
## Approach

- **Reductant Supply**
  - **Initial Screening**
  - **$H_2$ Production Profiling**
  - **Metabolic Flux Analyses**
  - **Specific Inhibitors**
- **Oxygen Sensitivity**
  - **Rational Enzyme Modifications**
  - **Directed Evolution**

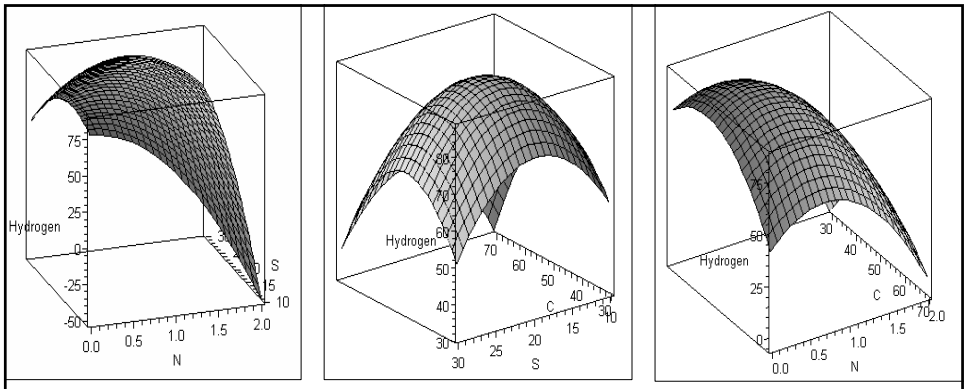



All-thread w/ wing nuts
Plexiglass top
Plate lid and rubber sealing mat
96-well filter bottom plate w/ catalyst and dye
Rubber perforated sealing mat
Silicone perforated sealing mat
96-well plate w/ cyanobacteria
Plexiglass bottom


Water soluble
Wilkinson's catalyst
+ ½H2
WST-3
WST-3 formazan

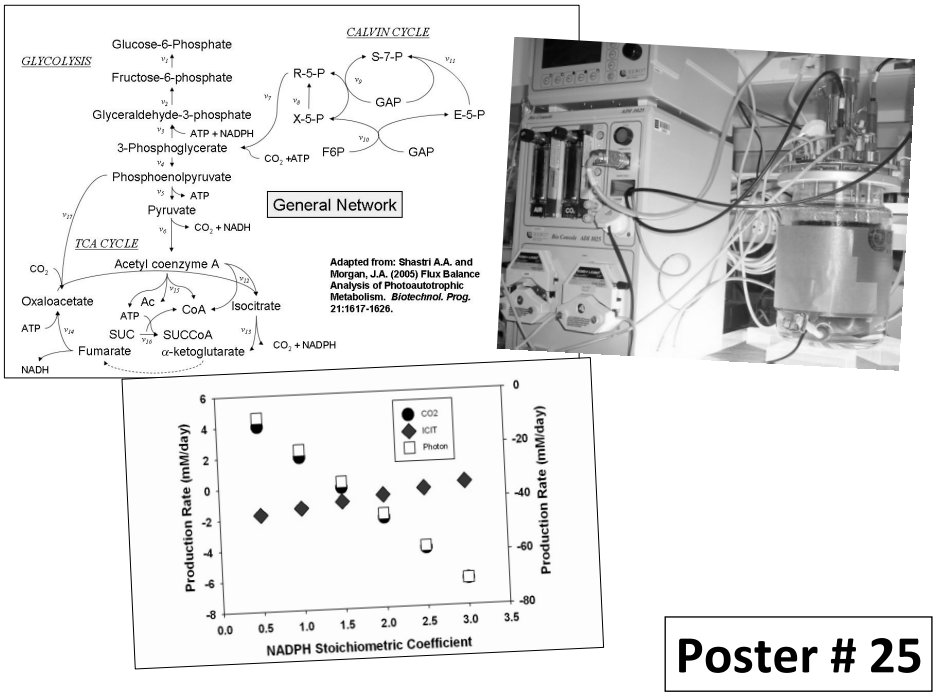

Mean Ctrl
Mean WT
Mean PSS1
Mean M55
Absorbance
1.4
1.2
1
0.8
0.6
0.4
0.2
0
pH 10, 10 mM bicarb.
pH 10, 80 mM bicarb.
pH 10, 150 mM bicarb.
pH 10, 10 mM bicarb.
pH 10, 80 mM bicarb.
pH 10, 150 mM bicarb.
Light on 2 off 58
Light on 2 off 58
Light on 2 off 58
Light on 8 off 52
Light on 8 off 52
Light on 8 off 52

**Poster #**

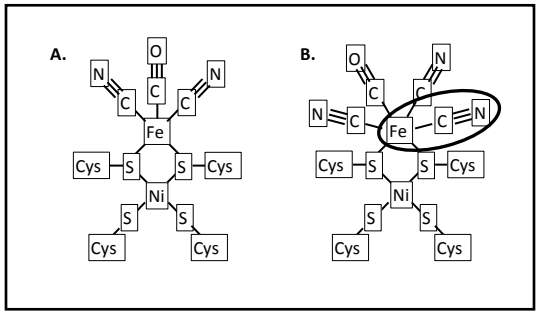


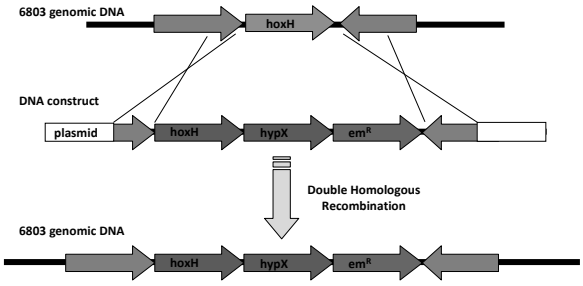

| | $H_2$ accumulation in headspace ($\mu$M) | |
|---|---|---|
| | 14 hours | 40 hours |
| Predicted maximum | 85 ( 23 ) | 210 ( 54 ) |
| S-deprived cultures | 0 ( 0 ) | 0.16 ( 0.3 ) |

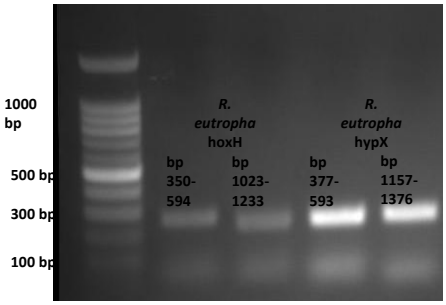
**Poster # 26**

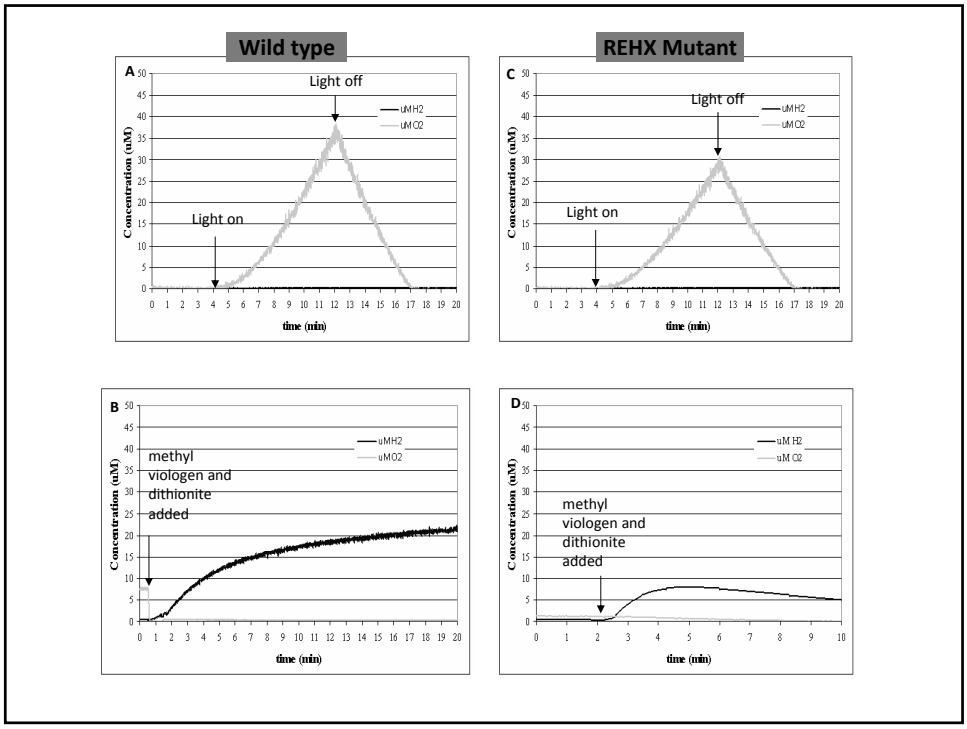


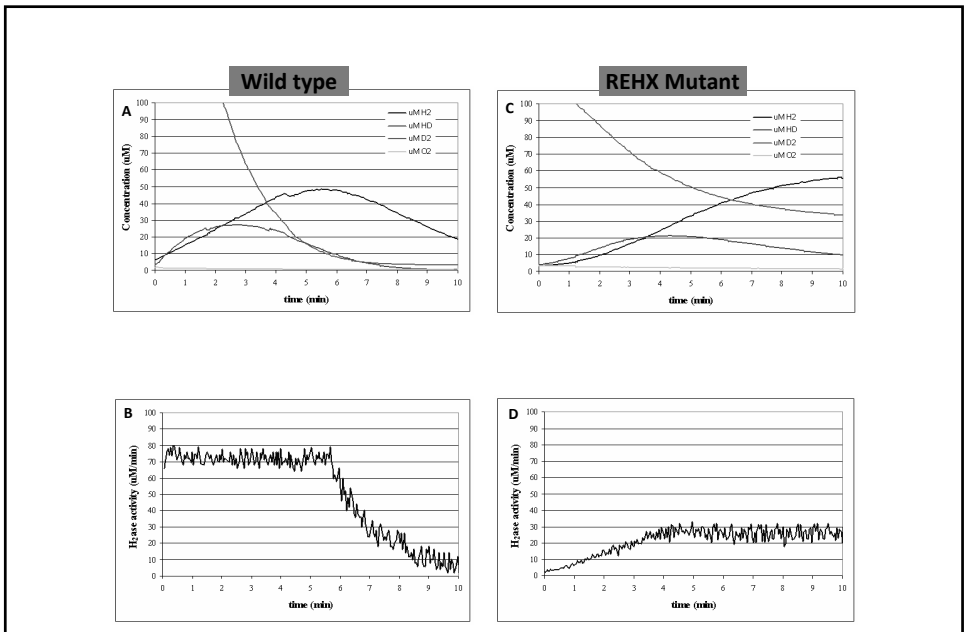

**Poster # 25**


A.
B.
O
N
C
Fe
Cys
S
Ni

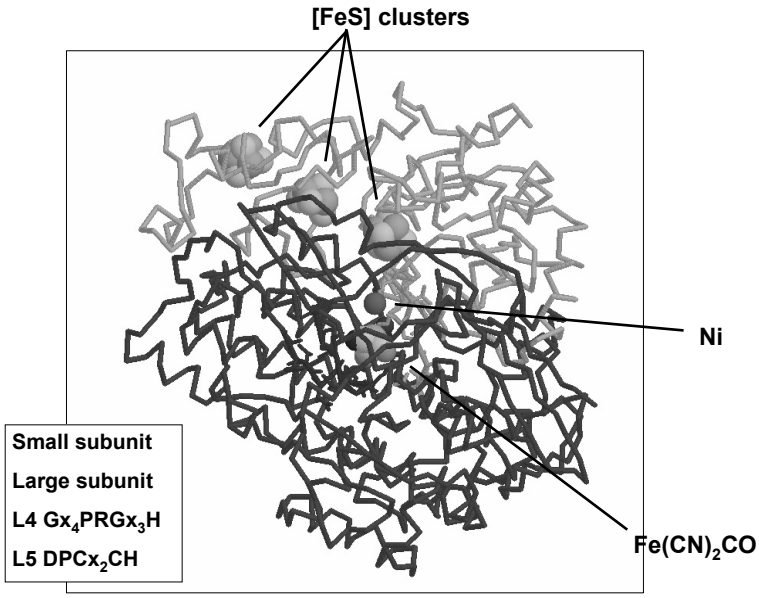

6803 genomic DNA
hoxH
DNA construct
plasmid
hypX
emR
Double Homologous Recombination
6803 genomic DNA

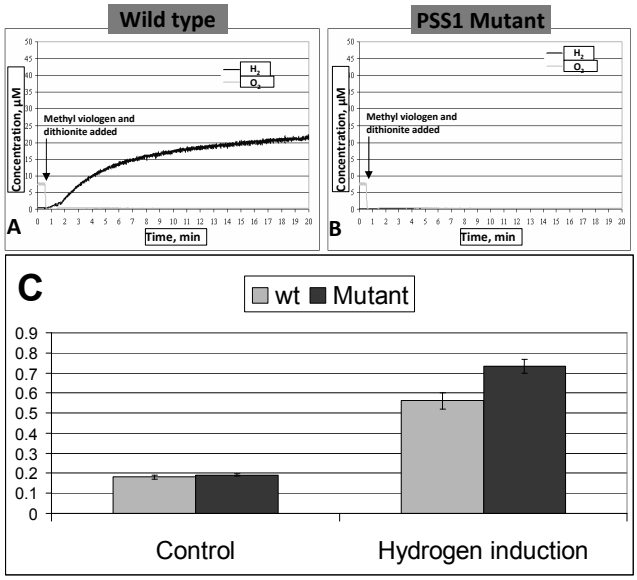

1000 bp
500 bp
300 bp
100 bp
R. eutropha hoxH
bp 350-594
bp 1023-1233
R. eutropha hypX
bp 377-593
bp 1157-1376

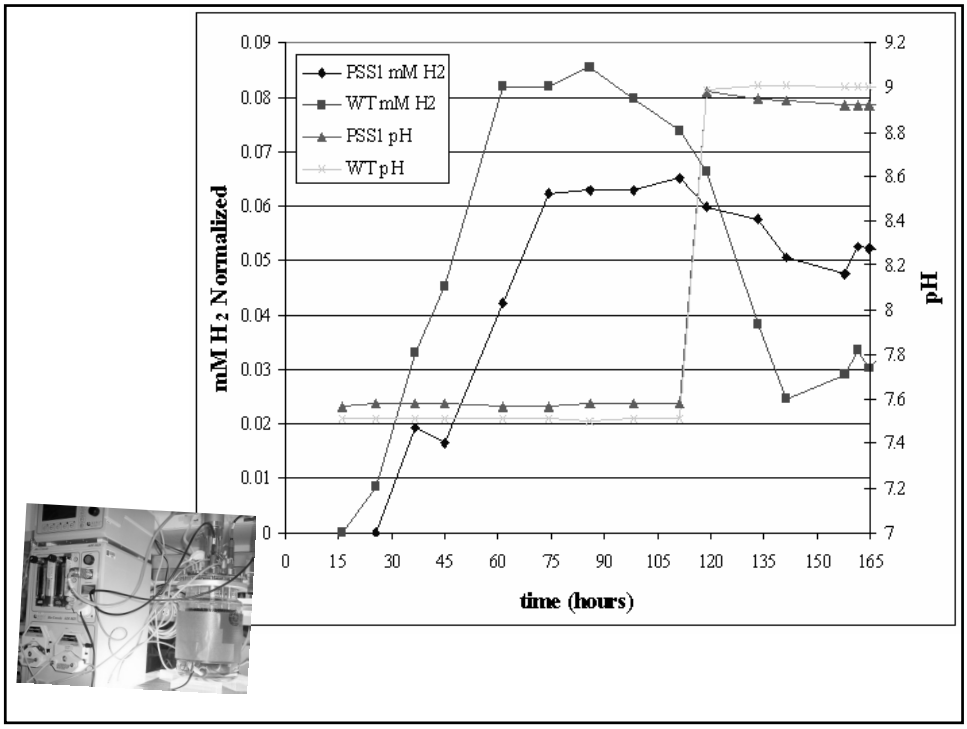

Wild type
REHX Mutant
A
B
C
D
Light off
Light on
uM H2
uM O2
methyl viologen and dithionite added
Concentration (uM)
time (min)

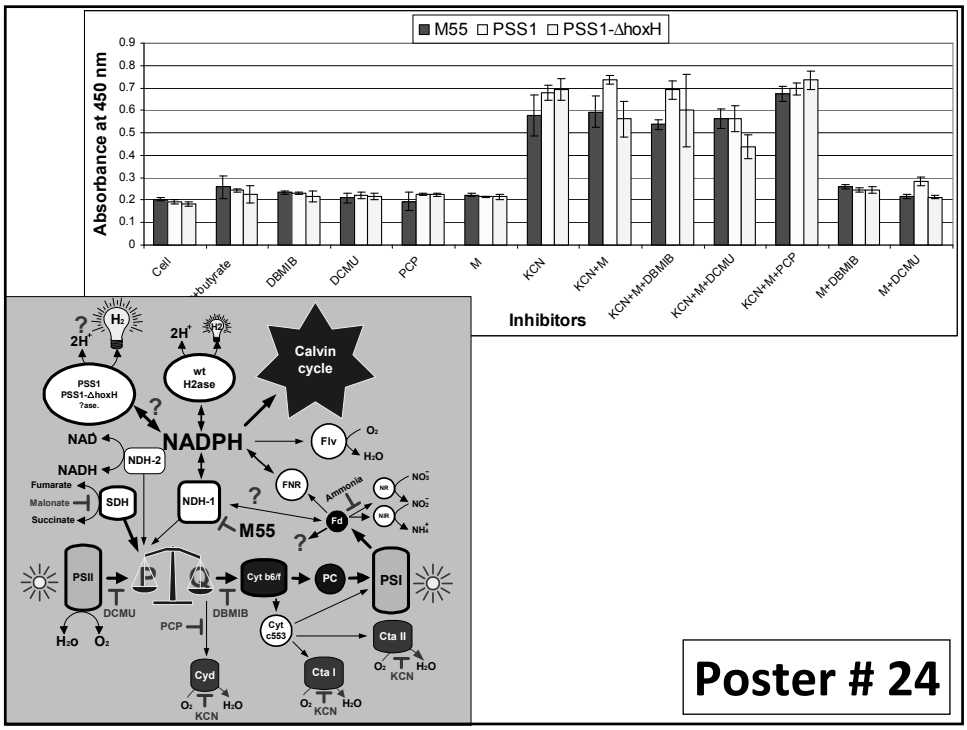

Wild type
REHX Mutant
A
B
C
D
uM H2
uM HD
uM D2
uM O2
Concentration (uM)
H2ase activity (uM/min)
time (min)

**Poster # 27**


[FeS] clusters
Ni
Fe(CN)2CO
Small subunit
Large subunit
L4 Gx4PRGx3H
L5 DPCx2CH


Wild type
PSS1 Mutant
H2
O2
Methyl viologen and dithionite added
Concentration, µM
Time, min
A
B
C
wt
Mutant
Control
Hydrogen induction


PSS1 mM H2
WT mM H2
PSS1 pH
WT pH
mM H2 Normalized
pH
time (hours)
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
9.2
9
8.8
8.6
8.4
8.2
8
7.8
7.6
7.4
7.2
7
0 15 30 45 60 75 90 105 120 135 150 165


M55 PSS1 PSS1-ΔhoxH
Absorbance at 450 nm
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Cell
+butyrate
DBMIB
DCMU
PCP
M
KCN
KCN+M
KCN+M+DBMIB
KCN+M+DCMU
KCN+M+PCP
M+DBMIB
M+DCMU
Inhibitors
H2
2H+
PSS1
PSS1-ΔhoxH
?ase.
wt
H2ase
Calvin
cycle
NAD
NADH
NADPH
NDH-2
NDH-1
Flv
O2
H2O
FNR
Ammonia
NR
NIR
NO3-
NO2-
NH4+
Fd
Fumarate
Malonate
Succinate
SDH
M55
PSII
DCMU
H2O
O2
PCP
DBMIB
Cyt b6/f
PC
PSI
Cyt
c553
Cta II
Cyd
Cta I
KCN

**Poster # 24**

# Recap

- High-Throughput Screening Assay
- Optimization
- Metabolic Flux Analysis
- REHX Mutant
- MV-DT Assay and Performance of Mutants

## Acknowledgments

**DOE-BER:**
**John Houghton**
**Sharlene Weatherwax**
**David Thomassen**

**NSF-MEWG:**
**Fred Heineken**

**Post-doctoral Associate:**
**Hatem Mohamed**

**PhD Students:**
**Paul Schrader**
**Liz Burrows**

**This research was supported by the Office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-05ER64130.**

THANK YOU!

# Creating a Pathway for the Biosynthesis of 1,2,4-Butanetriol

John Frost
Michigan State University

Collaborators:

Wei Niu
Man Kit Lau
Mapitso Molefe

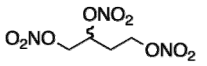
**February 11, 2008 MEWG: Interagency Conference on Metabolic Engineering**

Draths

Okemos, Michigan
Minneapolis, Minnesota

Research Areas:

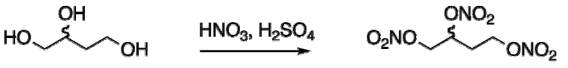
- synthetic chemistry
- catalysis
- materials science
- biocatalysis
- metabolic engineering




1,2,4-butanetriol trinitrate

BTTN

- less volatile than nitroglycerine (NG)
- more thermally stable than NG
- lower shock sensitivity relative to NG



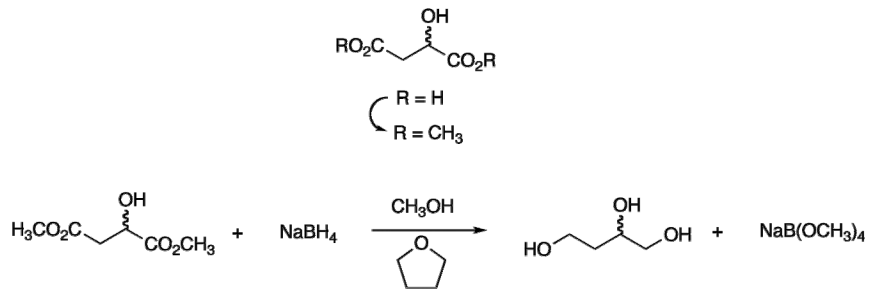

1,2,4-butanetriol

BT

1,2,4-butanetriol trinitrate

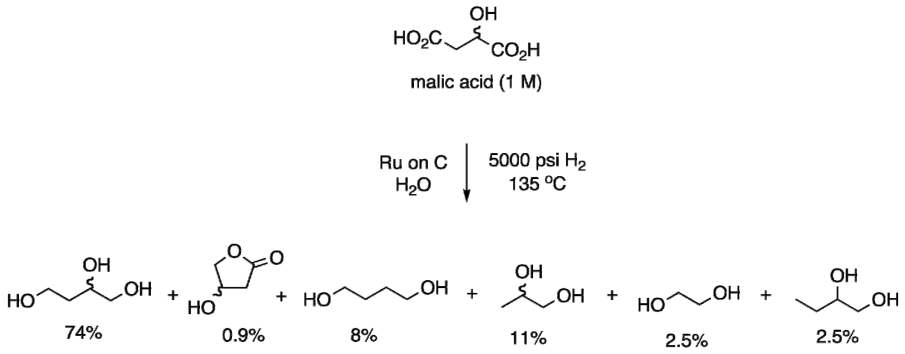
BTTN

## BT Manufacture:Stoichiometric Reduction

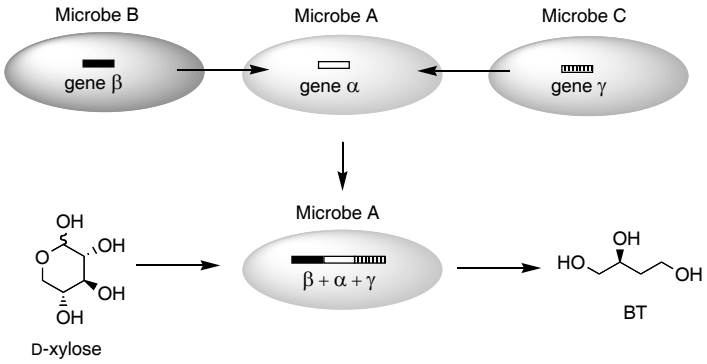



For every ton of BT produced, 2-6 tons of byproduct borate salts are generated. (WO 98/08793).

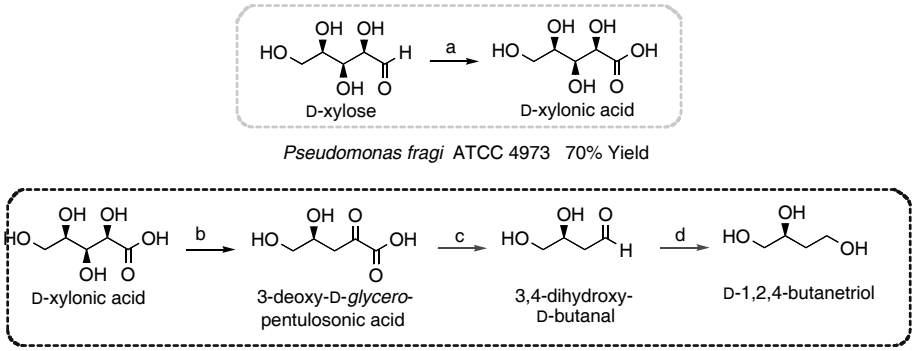
## BT Synthesis: Catalytic Reduction



## Creation of a Microbial Catalyst




## Microbial Synthesis of D-BT




*Pseudomonas fragi* ATCC 4973 70% Yield

*E. coli* DH5α/pWN6.186A 25% Yield

a. D-xylonate dehydrogenase (*P. fragi*); b. D-xylonate dehydratase (*E. coli*);
c. benzoylformate decarboxylase (*P. putida*); d. alcohol dehydrogenase (*E. coli*)

→ enzymes catalyzing reactions on native substrates
→ enzymes catalyzing reactions on non-native substrates

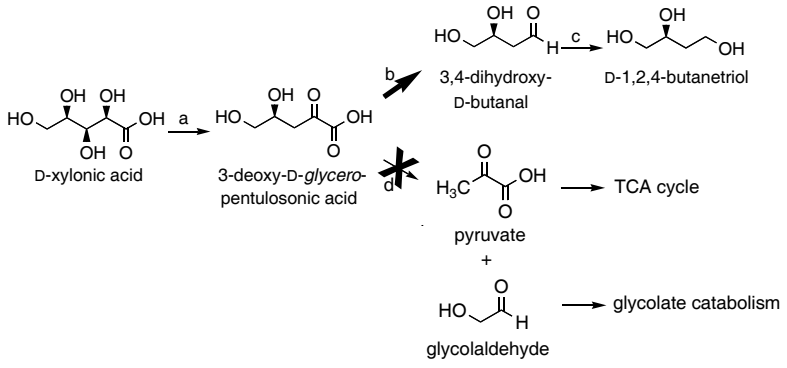
## *Escherichia coli* Synthesis of D-1,2,4-Butanetriol from D-Xylonic Acid in Rich Medium and Minimal Salts Medium

| host strain | plasmid and encoding genes | 1,2,4-BT (g/L) | yield (mol%) | culture medium and carbon source |
|---|---|---|---|---|
| DH5α | pWN6.186A<br>$P_{tac}$-*mdlC*, $Kan^R$, *lacI*$^Q$ | 1.6 | 25 | rich medium[a] and glucose[c] |
| W3110 serA | pWN7.126B<br>$P_{tac}$-*mdlC*, *serA*, *lacI*$^Q$ | 0.020 | 0.3 | minimal medium[b] and glucose[c] |

[a] tryptone (20 g/L), yeast extract (10 g/L), NaCl (5 g/L), $K_2HPO_4$ (3.75 g/L); [b] $K_2HPO_4$ (7.5 g/L), ammonium iron citrate (III) (0.3 g/L), citric acid monohydrate (2.1 g/L); [c] when cell growth entered mid-log phase, IPTG and D-xylonic acid were added into the culture medium.

Niu, W.; Molefe, M. N.; Frost, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 12998.

## *Escherichia coli* Biosynthesis of D-1,2,4-Butanetriol from D-Xylonic Acid




a. D-xylonate dehydratase (*E. coli*); b. benzoylformate decarboxylase *(P. putida*); c. alcohol dehydrogenase (*E. coli*); d. aldolase (*E. coli*).

## Purification of *Pseudomonas fragi* D-Xylonate Dehydratase

205
116
97
66
45
29

← ~ 60 kDa

QuickTime™ and a TIFF (Uncompressed) decompressor are needed to see this picture.

| | total protein (mg) | total activity (units) | specific activity (unit/mg) | yield (%) | purification fold |
|---|---|---|---|---|---|
| cell lysate | 7274 | 215 | 0.03 | 100 | 1.0 |
| DE-52 | 527 | 76 | 0.14 | 35 | 4.8 |
| hydroxyapatite | 219 | 66 | 0.30 | 31 | 10 |
| phenylsepharose | 24 | 50 | 2.10 | 23 | 70 |
| resource Q | 11 | 30 | 2.91 | 14 | 97 |

## N-Terminal Sequences of Trypsin-Digested *Pseudomonas fragi* D-Xylonate Dehydratase

| estimated position [a] | amino acid sequence | degenerate primers |
|---|---|---|
| 1 | TDSTPKRGRAQL[b] | JWF1006-JWF1009 |
| 70 | DAGGIPMEFPVHPIAEQSR | N/A |
| 150 | IGSGTVLWHAR | N/A |
| 300 | HSGVELSLEDWQRVG[b] | JWF1057-JWF1059 |
| 400 | MSVVGEAFR | N/A |
| 430 | AXVFEGPEDYTAR[b] | JWF1004-JWF1005 |

[a.] Positions of peptide fragments were assigned using the positions of the highest scored homologs derived from blasting the NCBI database;

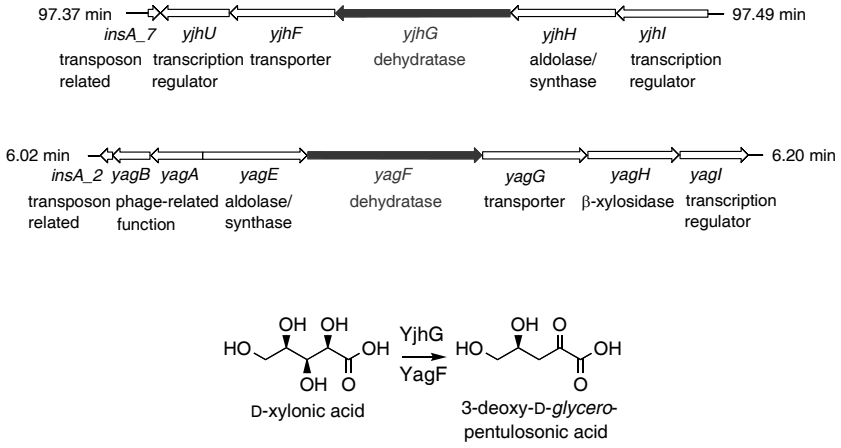
[b.] Degenerate primers were designed based on the underlined sequences.

## Partial DNA Sequence of *P. fragi* (ATCC 4973) D-Xylonate Dehydratase

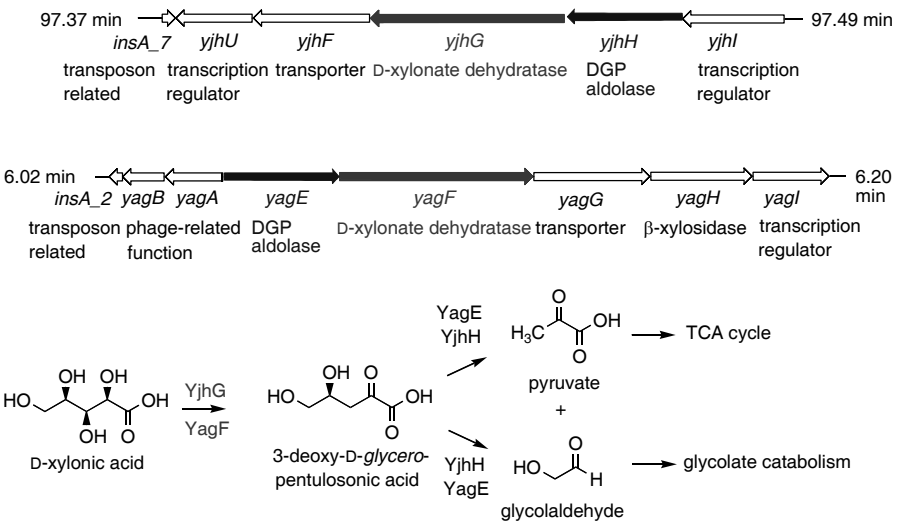
```
  1  CTC GAG GAT TGG CAG CGC GTG GGT GAA GAC GTG CCC TTG CTG GTC AAC TGC ATG CCT GCC
  1▶  L   E   D   W   Q   R   V   G   E   D   V   P   L   L   V   N   C   M   P   A
 61  GGC GAG TAC CTG GGC GAA AGC TTC CAC CGC GCC GGT GGC GTA CCG GCG GTG ATG CAT GAG
 21▶  G   E   Y   L   G   E   S   F   H   R   A   G   G   V   P   A   V   M   H   E
121  CTG GAC AAA GTG GGC CGC CTG CAC CGC GAT TGC CTC ACG GTC AGT GGC CGC AAC ATG GGT
 41▶  L   D   K   V   G   R   L   H   R   D   C   L   T   V   S   G   R   N   M   G
181  GAA GTG GTC GCC GAC TGC GTC ACC GGC GAC CGC GAC GTG ATC CGC TCC TAC GAA GAC CCG
 61▶  E   V   V   A   D   C   V   T   G   D   R   D   V   I   R   S   Y   E   D   P
241  CTG ATG CAC CGC GCT GGT TTT ATT GTG CTC AGC GGC AAC TTC TTC GAC AGC GCG ATC ATG
 81▶  L   M   H   R   A   G   F   I   V   L   S   G   N   F   F   D   S   A   I   M
301  AAA ATG TCG GTG GTG GGC GAA GCC TTC CGC AAG ACC TAC CTC AGC GAC CCG CTG CAA CCC
101▶  K   M   S   V   V   G   E   A   F   R   K   T   Y   L   S   D   P   L   Q   P
361  AAC AGC TTC GAG GCG CGG GCC ATT GTG TTC GAA GGC CCC GAA GAC TAC AC
121▶  N   S   F   E   A   R   A   I   V   F   E   G   P   E   D   Y
```

- The DNA sequence was isolated using PCR primer JWF1058 and JWF1004.
- DNA encoding the internal N-terminal sequence at position 400 is colored in red.
- The peptide has homology to hypothetical proteins from *Bradyrhizobium japonicum* USDA 110 and *Burkholderia fungorum* LB400.

## Elucidation of the *Escherichia coli* D-Xylonic Acid Catabolic Pathway



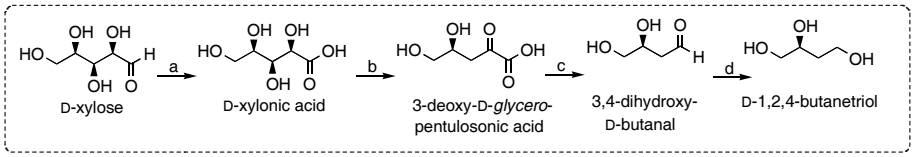
## Elucidation of the *Escherichia coli* D-Xylonic Acid Catabolic Pathway




## Microbial Synthesis of D-1,2,4-Butanetriol from D-Xylonic Acid Under Fermentor-Controlled Conditions

| entry | *E. coli* construct | plasmid inserts | xylonic acid | BT titer | BT yield |
|---|---|---|---|---|---|
| 1 | W3110 *serA* pWN7.126 | $P_{tac}$*mdlC* *serA* *lacI*$^{Q}$ | 18 g | 0.08 g/L | 1 % |
| 2 | W3110 *serA* Δ*yjhH* Δ*yagE* pWN7.126 | $P_{tac}$*mdlC* *serA* *lacI*$^{Q}$ | 28 g | 8 g/L | 45 % |

## Synthesis of D-BT from D-Xylose




a. D-xylonate dehydrogenase

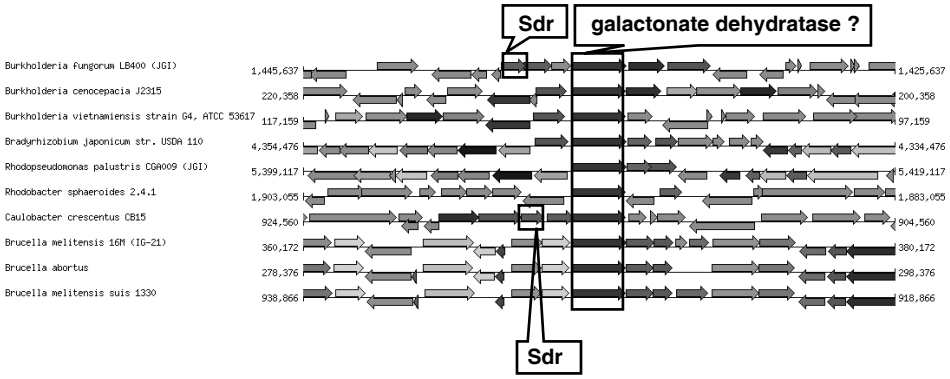
b. D-xylonate dehydratase (*E. coli*)

c. benzoylformate decarboxylase (*P. putida*)

d. alcohol dehydrogenase (*E. coli)*

## Identification of a D-Xylose Dehydrogenase

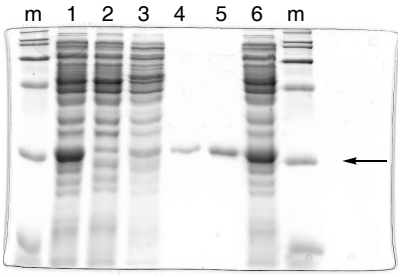
• BLAST search the ERGO database using the partial amino acid sequence of the previously isolated *P. fragi* D-xylonate dehydratase revealed Orfs with 50-70% of sequence identity.



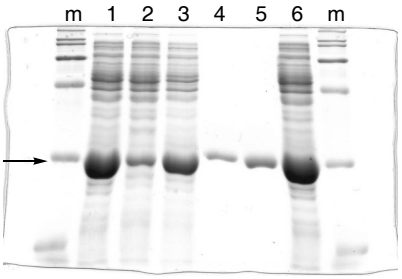
## Characterization of Putative D-Xylose Dehydrogenases

• The *xdh* gene from *Burkhoderia fungorum* LB400 and the *xdh* gene from *Caulobacter crescentus* CB15 were respectively cloned into vector pQE-30. The two D-xylose dehydrogenases were expressed in *E. coli* and purified as 6×His-tagged fusion proteins.

• SDS-PAGE for 6xHis-Xdh-LB400




• SDS-PAGE for 6xHis-Xdh-CB15




29 KDa

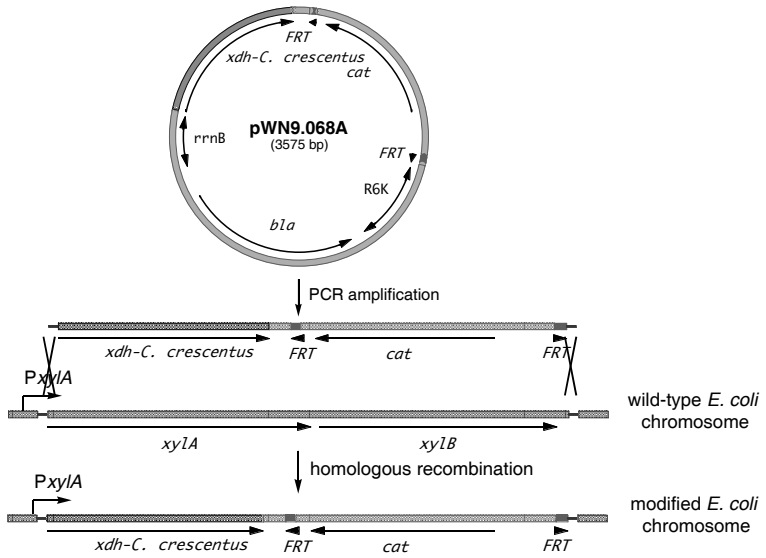
m, molecular weight markers; 1, cell lysate; 2, flow-through; 3, wash; 4, elution 1; 5, elution 2; 6, cell lysate.

## Characterization of D-Xylose Dehydrogenases

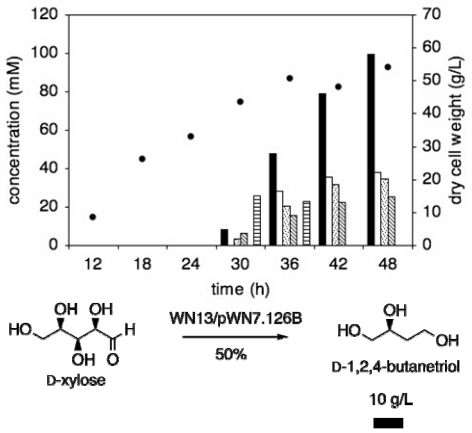
| | NAD | D-xylose | | D-glucose | | L-arabinose | |
|---|---|---|---|---|---|---|---|
| | $K_m$ (mM) | $K_m$ (mM) | $k_{cat}$[c] ($s^{-1}$) | $K_m$ (mM) | $k_{cat}$ ($s^{-1}$) | $K_m$ (mM) | $k_{cat}$ ($s^{-1}$) |
| Xdh-*B. fungorum* | 0.26[a] | 0.97 | 29 | 176 | 12 | 43 | 13 |
| Xdh-*C. crescentus* | 0.13[a] | 0.099 | 41 | 538 | 24 | 34 | 40 |

D-fructose, D-galactose, D-mannose and D-ribose were not substrates for either *B. fungorum* or *C. crescentus* xylose dehydrogenase

## Insertion of *C. crescentus xdh* into the *E. coli* Chromosome

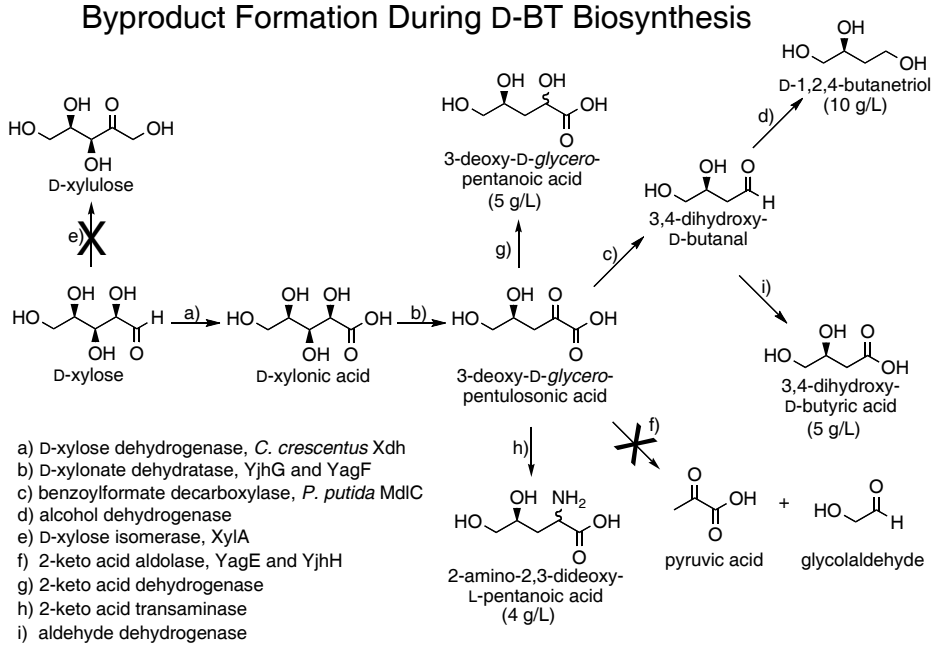



## Fermentor-Controlled Synthesis of BT from Xylose



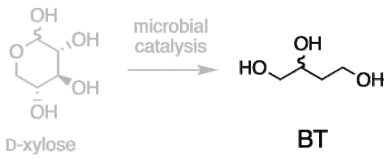

*WN13: E. coli* W3110*serA*Δ*yjhH*Δ*yagExylAB*::*xdh*-Cm$^R$ pWN7.126B: *serA*, *lacI*$^Q$, $P_{tac}$ *mdlC*

## Byproduct Formation During D-BT Biosynthesis




a) D-xylose dehydrogenase, *C. crescentus* Xdh
b) D-xylonate dehydratase, YjhG and YagF
c) benzoylformate decarboxylase, *P. putida* MdlC
d) alcohol dehydrogenase
e) D-xylose isomerase, XylA
f) 2-keto acid aldolase, YagE and YjhH
g) 2-keto acid dehydrogenase
h) 2-keto acid transaminase
i) aldehyde dehydrogenase

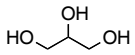
## Green Synthesis of BT



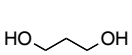

- nontoxic, renewable xylose
- reduction of salt waste streams
- avoidance of elevated temperatures and pressures
- single step
- domestic manufacture

# Dual-Use Markets

## C-3 Biobased



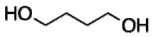

glycerol



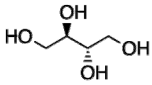

1,3-propanediol

1,2-propanediol

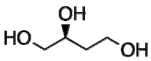
## C-4 Biobased




1,4-butanediol




erythritol




1,2,4-butanediol