# FINAL REPORT

## Applying Synthetic Biology Principles to Increase Biocellulose (BC) Production

## SERDP Project WP-2333

AUGUST 2017

Dr. Mark Fuller
**CB&I Federal Services**



SERDP
DOD ▪ EPA ▪ DOE

*Page Intentionally Left Blank*

*Page Intentionally Left Blank*

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| 01-31-2017 | Final | September 2013 - January 2017 |

**4. TITLE AND SUBTITLE**
Applying Synthetic Biology Principles To Increase Biocellulose (BC) Production

**5a. CONTRACT NUMBER**
W912HQ-13-C-0065

**5b. GRANT NUMBER**
NA

**5c. PROGRAM ELEMENT NUMBER**
NA

**6. AUTHOR(S)**
Fuller, Mark E., Ph.D.

**5d. PROJECT NUMBER**
WP-2333

**5e. TASK NUMBER**
NA

**5f. WORK UNIT NUMBER**
NA

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
CB&I Federal Services, LLC
17 Princess Road
Lawrenceville, NJ 08648

**8. PERFORMING ORGANIZATION REPORT NUMBER**
NA

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Strategic Environmental Research and Development Program
4800 Mark Center Drive, Suite 17D08
Alexandria, VA 22350-3600

**10. SPONSOR/MONITOR'S ACRONYM(S)**
SERDP

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**
NA

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Distribution Statement A: Approved for Public Release, Distribution is Unlimited

**13. SUPPLEMENTARY NOTES**
None

**14. ABSTRACT**

The key objective of this project was to apply the tools and techniques of synthetic biology to create a biocatalyst for the production of high purity biocellulose (BC), and to develop a scalable process to produce this material for use in critical DoD applications. The biocellulose was subsequently to be formulated into military grade nitrocellulose (NC) to evaluate its characteristics and conformity to military specifications. The main results of this project were as follows: 1) Inducible promoter systems for Gluconacetobacter spp. were identified which should allow control of biocellulose production. 2) A method to use attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy to detect bacterial impurities (e.g., nucleic acids, lipids, protein) in biocellulose was developed. 3) Detailed metabolic models of the cellulose producing strains used in this research were developed and used to identify key genes/pathways that could be modified to increase biocellulose production. 4) Proof-of-concept synthesis of nitrocellulose from biocellulose indicated that the material was comparable to military grade nitrocellulose.

**15. SUBJECT TERMS**
nitrocellulose, biocellulose, synthetic biology, Gluconacetobacter xylinus, propellant, green chemistry

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Dr. Mark E. Fuller |
| U | U | U | UU | 108 | 19b. TELEPHONE NUMBER *(Include area code)* 609-895-5348 |

*Page Intentionally Left Blank*

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ACRONYM LIST

| | |
|---|---|
| ATCC | American Type Culture Collection |
| ATR | attenuated total reflection |
| BC / BNC | biocellulose or bacterial cellulose / bionitrocellulose |
| °C | degree Celsius |
| CDG | c-di-GMP |
| cm | centimeter |
| CS | cellulose synthase |
| d | day |
| DMSO | dimethylsulfoxide |
| DNA | deoxyribonucleic acid |
| DoD | Department of Defense |
| FMN/$FMNH_2$ | flavin mononucleotide |
| FTIR | Fourier-transform infrared |
| g | gram |
| *g* | gravity |
| GMP | guanosine monophosphate |
| GFP | green fluorescent protein |
| Gx | *Gluconacetobacter xylinus* |
| h | hour |
| IPTG | isopropyl β-D-1-thiogalactopyranoside |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| L | liter |
| µ | micro |
| m | milli or meter |
| M | molar |
| min | minute |
| mL | milliliter |
| mm | millimeter |
| mol | mole |
| n | nano |
| NADP/NADPH | nicotinamide adenine dinucleotide phosphate |
| NAVAIR | Naval Air Warfare Weapons Division, China Lake, CA |
| NC | nitrocellulose |
| OB/OD | open burn/open detonation |
| $OD_{600}$ | optical density at 600 nm |
| ONPG | ortho-Nitrophenyl-β-galactoside |
| qPCR | quantitative polymerase chain reaction |
| RFP | red fluorescent protein |
| rpm | revolutions per minute |
| RT-qPCR | reverse transcriptase quantitative polymerase chain reaction |
| SERDP | Strategic Environmental Research and Development Program |
| SSDB | sequence similarity database |
| T4MO | toluene-4-monooxygenase |
| TCA | tricarboxylic acid |
| THF | tetrahydrofuran |
| UV | ultraviolet |
| wt | weight |

# ACKNOWLEDGEMENTS

This SERDP Project was a collaborative effort among Dr. Jenifer Reed at the University of Wisconsin, Madison, Dr. Steve Fallis at the Naval Air Warfare Center Weapons Division, China Lake and Dr. Mark E. Fuller of CB&I Federal Services, LLC. We wish to acknowledge all of the collaborators for their significant contributions to this research project. We would also like to acknowledge the excellent effort of Prashant Kumar (UW), Dr. Thomas Groshens (Chinal Lake), Kevin McClay (CB&I), and Christina Andaya (CB&I).

We gratefully acknowledge SERDP for their financial support, and Dr. Robin Nissan, the Weapon Platforms Program Manager at SERDP, for his guidance and support during this project.

# EXECUTIVE SUMMARY

Objective: This SERDP project was a collaborative effort among scientists at CB&I Federal Services, the University of Wisconsin-Madison, and the Naval Air Warfare Center Weapons Division-China Lake. The key objective of this project was to apply the tools and techniques of synthetic biology to create a biocatalyst for the production of high purity biocellulose (BC), and to develop a scalable process to produce this material for use in critical DoD applications. The biocellulose was subsequently to be formulated into military grade nitrocellulose (NC) to evaluate its characteristics and conformity to military specifications.

Technical Approach: This project used a multi-pronged approach to increase the production of high purity biocellulose. Synthetic biology was used to construct a bacterial strain that had controllable cellulose producing machinery. The process involved genome sequencing and manipulation of the key cellulose synthesis genes. Metabolic modeling was used to identify pathways in the strain that could be manipulated to optimize carbon and energy flow to enhance biocellulose yields, including both anabolic and catabolic pathways to allow the strain to utilize low cost feedstocks (e.g., waste glycerol from biodiesel production) as carbon and/or energy sources to support cellulose synthesis. Batches of biocellulose produced at various stages during the project were analyzed for key parameters (e.g., crystallinity, purity), and used to make and nitrocellulose.

Results: The key results of this project were as follows:

- One of the first reported inducible promoter systems for *Gluconacetobacter* spp. was identified and validated.
- Isoforms of the key biocellulose synthesis enzyme AcsA that were insensitive to the metabolic control molecule CDG were engineered and should to result in greater cellulose production than the native AcsA.
- A method to use attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy to detect bacterial impurities (e.g., nucleic acids, lipids, protein) in biocellulose was developed.
- Detailed metabolic models of the cellulose producing strains used in this research were developed and used to identify key genes/pathways that could be modified to increase biocellulose production.
- Proof-of-concept synthesis of nitrocellulose from biocellulose was successfully performed, and preliminary characterization indicated that the material was comparable to military grade nitrocellulose with respect to FTIR spectra and nitrogen content.

Benefits: This research has added to the scientific knowledge base regarding cellulose producing bacteria and the cellulose synthesis control. With further development, these findings could lead to the ability to produce biocellulose at an industrial scale, for use in other DoD and non-DoD purposes, including functional papers and textiles, medical implants, and food additives.

## 1.0 PROJECT OBJECTIVES AND INTRODUCTION

### 1.1 Project Objectives

The U.S. Department of Defense is continuing to make its operations more sustainable while assuring that the resources required to maintain military readiness are secure and consistent. Recent advances in biotechnology, particularly in the area of synthetic biology, can contribute to these efforts by enabling sustainable production of the propellant and explosive compounds used in munitions.

The key objective of this project was to apply synthetic biology principles to create a biocatalyst for the production of high purity biocellulose for use in DoD applications. Metabolic modeling was used to identify beneficial genetic changes that could be made, at both the enzymatic and metabolic pathway levels, to optimize carbon and energy flow to enhance cellulose production. Initial efforts to synthesize and characterize nitrocellulose from biocellulose were performed. An overview of the project is illustrated in Figure 1-1 below. A flowchart of project activities is presented in Figure 1-2.

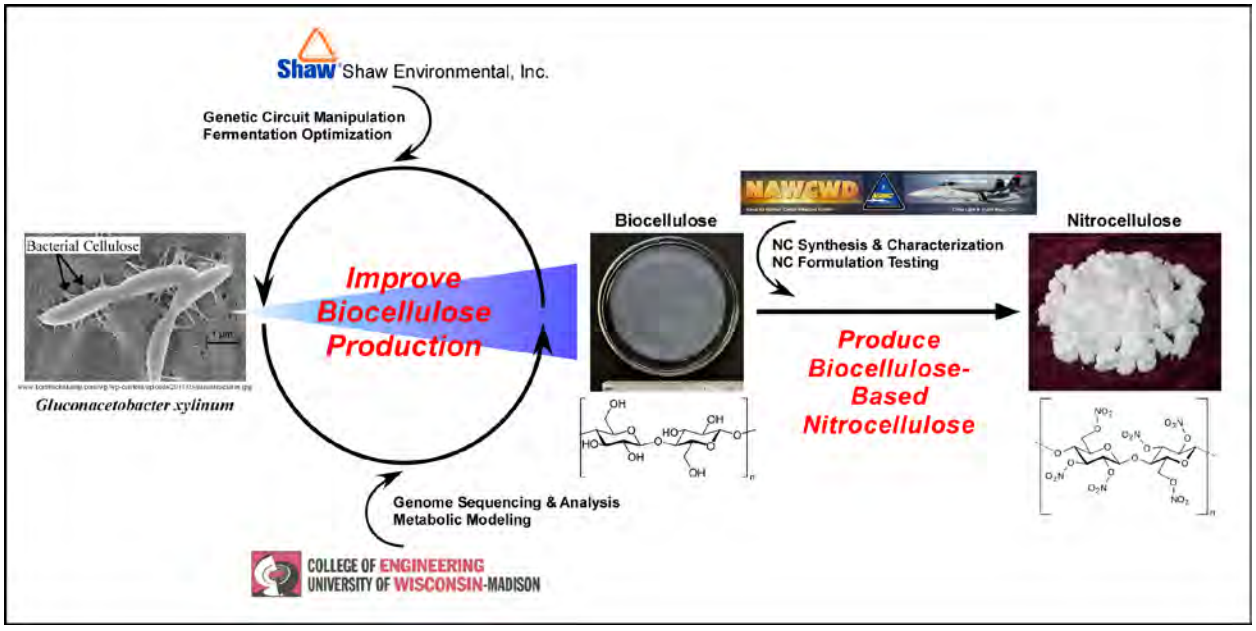



**Figure 1-1. Conceptual overview of the biocellulose project.**

The two key technical questions addressed were:

1) Can cellulose producing bacterial strains be engineered to efficiently produce biocellulose?
2) Can biocellulose-derived nitrocellulose meet or exceed military specifications?

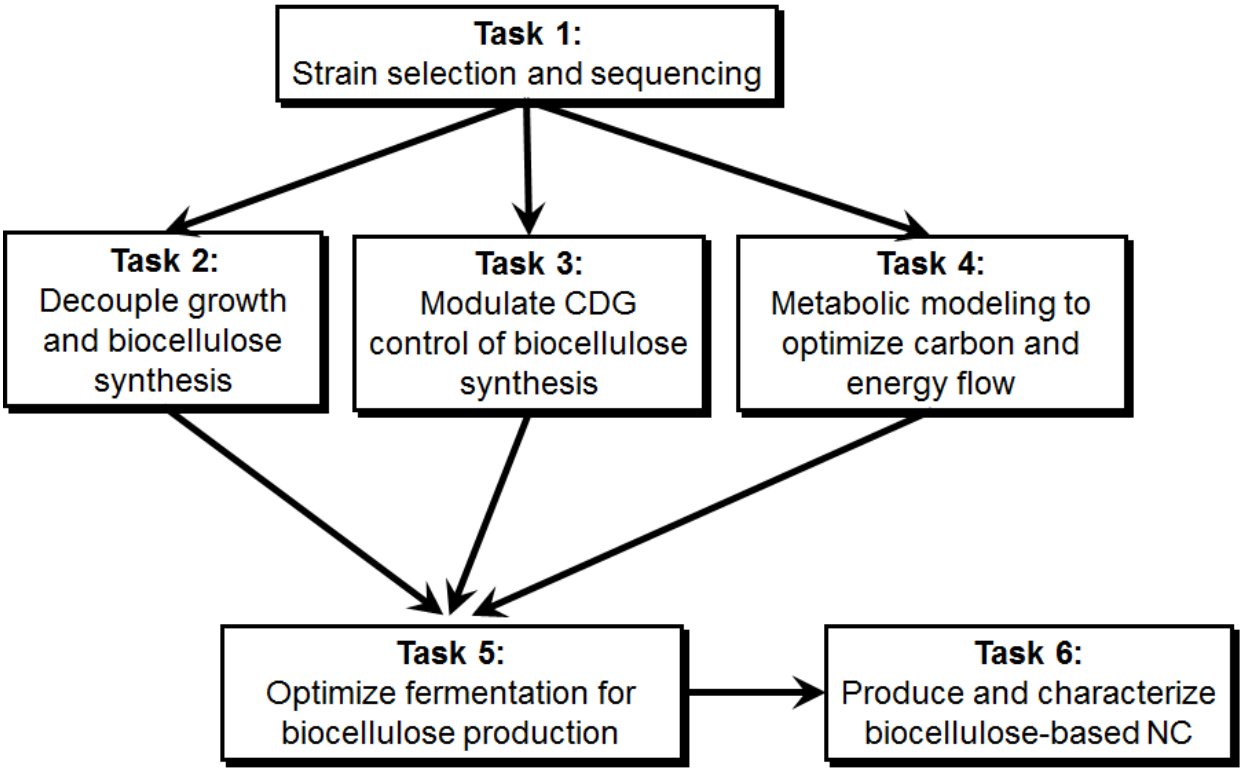



**Figure 1-2. Project flowchart.**

## 1.2 Introduction

***1.2.1 Nitrocellulose production.*** The base material for many of the Dodd's propellant formulations is nitrocellulose (NC), and many other high-performance propellants under development contain varying levels of NC. NC is produced in bulk by reacting cellulose with a combination of nitric and sulfuric acid. Nitration occurs via ester linkage to the cellulose backbone, replacing the hydroxyl groups and resulting in nitrogen levels between 10 and 14% (volume NO evolved per g NC during a controlled denitration process). The acidic NC is then neutralized and washed extensively.

The current sources of cellulose for NC production are cotton linters or wood pulp. One of the key issues that arise during the preparation of NC is the production of impurities. Impurities lead to deterioration of propellant formulations during storage (17), resulting in reduced performance or total failure during use. The generation of waste or off-spec nitrocellulose due to deterioration leads to NC disposal issues. The current practice for NC disposal is open burn/open detonation (OB/OD). However, OB/OD is coming under increased scrutiny by regulators, especially with respect to air emissions, which may lead to much higher costs if more advanced disposal methods are required in the future.

The production and use of higher purity NC would extend the useful life of existing propellant systems, as unwanted side reactions associated with impurities that lead to propellant degradation would be minimized or eliminated. This could lead to significant reductions in Total Ownership Costs. Issues have arisen with procuring the starting cellulosic materials with the required purity specifications and batch-to-batch consistency (Mr. Phillip Hui, Army Materiel Command, personal communication). Thus, the DoD is seeking alternative sources of cellulose, including non-plant derived materials such as bacterial cellulose.

***1.2.2 Bacterial cellulose – Current understanding.*** Cellulose is produced by a number of bacterial strains, and cellulose production has been studied in depth in terms of genetics and physiology in the obligate aerobe *Gluconacetobacter xylinus* (formerly *Acetobacter xylinus*) (55). *G. xylinus* has been shown to have a high degree of metabolic flexibility, producing cellulose when grown on a wide variety of substrates (24, 31, 32, 38, 44, 53). A traditional application of this ability takes place in East Asia to produce the edible dessert "nata de coco" (67), which has been scaled up to a limited industrial scale. Because of its very high purity and desirable physical characteristics, bacterial cellulose (or biocellulose) is being investigated as a specialty chemical for use in wound dressings, functionalized papers, and food additives (55). On a microscopic level, biocellulose produced during static incubation is formed as sheets or ribbons.

Several *G. xylinus* strains are available for use as potential biocatalysts for biocellulose production, and previous studies have demonstrated genetic enhancement of *G. xylinus* through the introduction of exogenous DNA, via either bacterial conjugation or electroporation (5, 11, 27, 46, 64). Our own experience with *G. xylinus* (ATCC 53582), and that of others (16, 22), indicated that *G. xylinus* does not readily accept DNA via bacterial conjugation, but experiments evaluating electroporation to introduce DNA showed initial promise.

***1.2.3 Increasing biocellulose production – Genetic circuit modification.*** The key to improving biocellulose production to a scale that would provide material for industrial-scale process such as NC production is to increase cellulose synthesis efficiency (rate, substrate conversion, and volumetric yield), which requires targeted genetic modification of the cellulose synthesizing organism. During growth on glucose, wildtype *G. xylinus* directs and balances carbon flow into both cell biomass and cellulose production (Figure 1-3) resulting in ~3 molecules of glucose being incorporated into cellulose for every molecule of glucose that is used for growth and energy (55). The diversion of 75% of the glucose into cellulose during the initial growth phase is not ideal, as it delays the buildup of productive biomass, thereby limiting the amount of cellulose that will eventually be synthesized. The presence of the cellulose matrix itself during the early stage of fermentation also slows growth by physically decreasing mass transfer of oxygen and other nutrients to cells (67). Uncoupling cell growth and cellulose production by using synthetic biology to create a controllable on/off switch for the biocellulose synthesis machinery was the first step to increase efficiency and enhance cellulose production rates. With such a switch, the cells could initially be grown to high density (i.e., switch off), then triggered (i.e., switch on) to produce biocellulose. This alone was expected to greatly increase the amount of biocellulose that can be produced per unit time and volume.

Another factor that can increase yield is increasing the efficiency of carbon flow into cellulose after triggering cellulose production. That is, under normal conditions only three glucose molecules are converted to cellulose for every glucose molecule metabolized, far short of the maximum theoretical incorporation rate of 18 glucose molecules to cellulose per glucose molecule metabolized (based on ATP yield vs. ATP required for cellulose production). The control of glucose use in *G. xylinus* is dependent primarily (but not solely) on the fact that the cellulose synthase complex requires a positive activator, c-di-GMP (CDG), to function. Positive activators enhance or turn on gene transcription. The level of CDG present in the cell is regulated by two enzymes, one is a cyclase that synthesizes CDG, and the other is an esterase that degrades the CDG. These enzymes monitor both the oxygen concentration and cellular energy level, which allows the cells to produce biomass/energy and produce cellulose at a consistent and sustainable level in their natural environment. To shift the system in favor of cellulose production in an industrial setting, the controlling effects of CDG must be alleviated. During this project we made changes to the wildtype cellulose synthase genes to alter their sensitivity to, and control by, CDG, leading to enhanced biocellulose production.

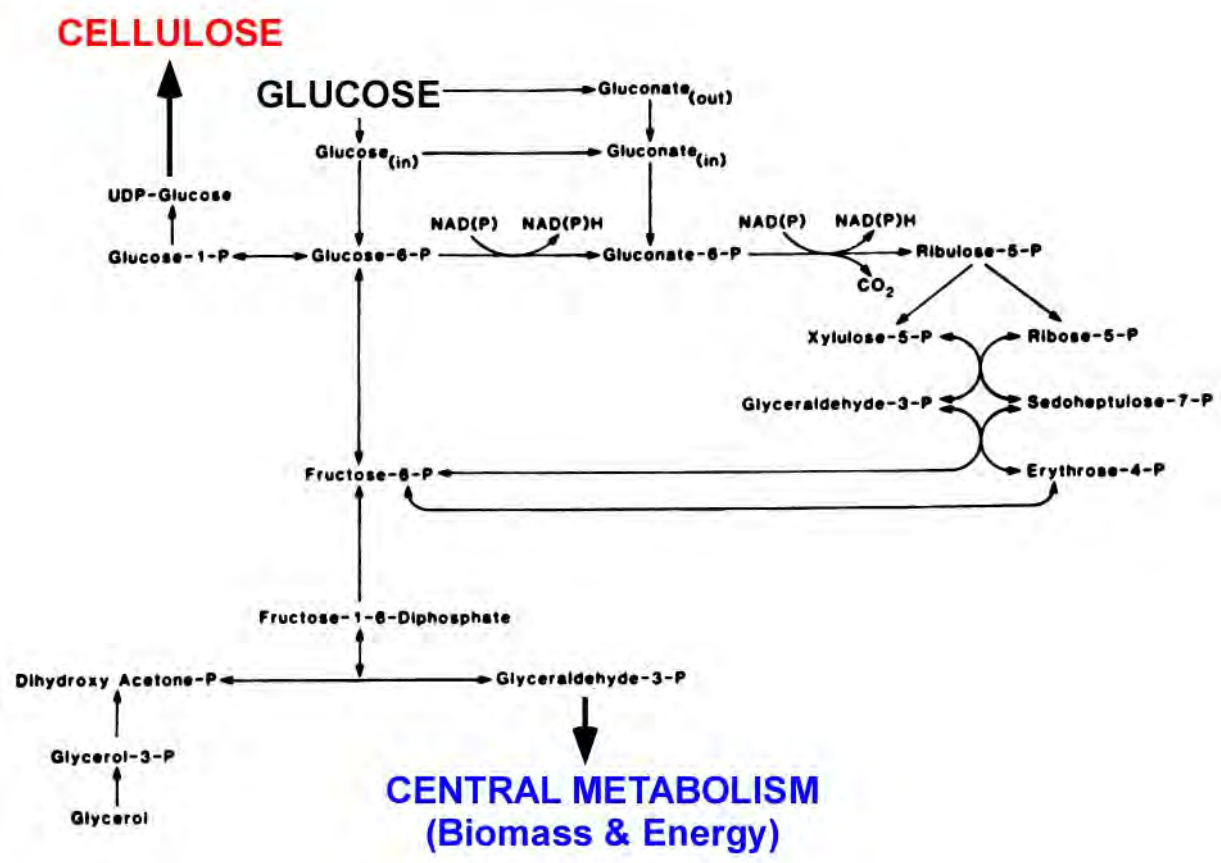



**Figure 1-3. Pathway depicting metabolism of glucose and cellulose synthesis (adapted from Ross, 1991)**

***1.2.4 Increasing biocellulose production – Metabolic modeling.*** In addition to modulating cellulose synthesis regulation directly, shutting down non-productive metabolic pathways and addressing metabolic inefficiencies that limit the amount of carbon and energy that can be directed towards cellulose production would result in higher biocellulose yields per unit of substrate. For example, Shigematsu et al. (59) found that deleting the glucose dehydrogenase gene in *G. xylinus* reduced the formation of gluconic acid from 23 g/L (nearly half of the glucose added) to zero. Taking additional steps, including over-expressing key enzymes such as nicotinomide nucleotide transhydrogenase (which replenishes NADPH) would complement the deletion of glucose dehydrogenase and result in faster and more efficient cellulose synthesis. During this project, metabolic modeling was used to identify metabolic network changes (e.g., addition or removal of enzymes) and design biocellulose synthesis circuits that would allow for more efficient use of cellular energy stores and production of biocellulose from lower cost feedstocks (e.g., glycerol).

***1.2.5 Increasing biocellulose production – Fermentation.*** Previous findings indicated that the mode of culturing cellulose-producing bacteria has moderate to dramatic effects on biocellulose production. For instance, while most aerobic bacteria are cultured in stirred or agitated cultures (to facilitate mass transfer of nutrients and oxygen), investigators reported that otherwise identical strains of *G. xylinus* grown in agitated cultures produce less biocellulose than when grown in static culture (66). Similarly, during initial experiments we had observed that cellulose production (at least during the initial growth phases) was inversely proportional to mixing speed in stirred flask reactors. Various other growth configurations have been tested with incremental gains in biocellulose production, including biofilm reactors (10), two-stage processes (50), and rotating disk reactors (35). Oxygen saturation and pH play a major role in cellulose production (25), so these parameters need to be carefully controlled to optimize biocellulose output.

Additionally, once biocellulose is produced, it must be cleaned to remove cellular biomass, and processed in a manner to produce a final dry product that is amenable to nitration reactions to produce nitrocellulose. Procedures to routinely assess the purity of the final biocellulose must also be developed.

Therefore, work performed during this project focused on scalable fermentation and processing technologies, combined with standardized analytical characterization methods, to produce biocellulose capable of being converted into military grade nitrocellulose.

## 2.0 PROJECT TASKS

### 2.1 Task 1 – Strain selection

#### *2.1.1 Goal and Introduction*

The goal of this task was to acquire several cellulose producing bacterial strains, and then perform a downselect process based on several criteria to choose one or two strains for further study and development.

#### *2.1.2 Methods, Results, and Discussion*

**2.1.2.1 Bacteria strains.**

The initial group of biocellulose producing strains used for this project are listed in Table 2.1-1. Several biocellulose producers were obtained from the American Type Culture Collection (ATCC).

In addition, new strains isolated from several samples of kombucha starter cultures used to make traditional fermented teas. Small samples were aseptically removed from the kombucha "symbiotic 'colony' of bacteria and yeast" (SCOBY) using a flame sterilized razor blade and forceps and used to inoculate R2A agar plates. Single colonies growing on the R2A plates were then inoculated into the media described below to check for cellulose production.

**2.1.2.2 Growth media.**

Routine culturing of biocellulose producing strains was performed using ATCC 1765 *Acetobacter xylinus* medium which contains (g per L distilled water): glucose, 20; peptone, 5; yeast extract, 5; $Na_2HPO_4$, 2.7; citric acid, 1.5. Growth was also performed in a defined basal salts medium (BSM) (23) amended with glucose. BSM with other carbon sources like glycerol, and with varying pH levels, were also used depending on the experiment.

**2.1.2.3 Downselection criteria.**

The following criteria were used to select the best strains for additional characterization and manipulation.

1) Relative biocellulose production.
Each strain was evaluated qualitatively by visually assessing their production of biocellulose in standard medium.

2) Ability to transfer DNA into the cells.
Using the standard electroporation protocol described in section 2.2.2.2 below, the ability to transfer DNA into the strains was evaluated. Successful DNA transfer was scored by production of ampicillin resistant transconjugates.

3) Availability of existing genome sequence data.
Genome databases were evaluated to determine if genomic sequence data was available for each strain (e.g., for the selected ATCC strains), or for closely related strains, so as to maximally leverage the existing published data.

The results of the downselect process are presented in Table 2.1-2.

Known Gluconacetobacter strains:
ATCC 53582 (*G. xylinus*, strain used for preliminary experiments)
ATCC 23769 (*G. hansenii*, full genome sequenced, previous reports of genetic manipulation)
ATCC 10245 (*G. xylinus*, previous reports of genetic manipulation)

In-house BC strains:
Gx 399 (originally EM2, isolated from kombucha)
Gx 400 (originally STH, isolated from kombucha)
Gx 401 (originally GBH, isolated from kombucha)

**Table 2.1-1. Identification and source of biocellulose producing strains.**

| | **Native BC production** | **Genetic amenability** |
|---|---|---|
| ATCC 53582 | Excellent | Very Poor |
| ATCC 23769 | Poor | ND |
| ATCC 10245 | ND | ND |
| Gx 399 | Good | Moderate |
| Gx 400 | Good ± | Moderate |
| Gx 401 | Moderate | Moderate |

**Table 2.1-2. Results of downselect screening of biocellulose producing strains.**

### *2.1.3 Conclusions*

Based on the outcome of the downselect process, the decision was made to perform most of the subsequent work with strains ATCC 53582 and Gx 399, with sporadic use of ATCC 23769 for some experiments, based on the number of literature citations pertaining to this strain and availability of an annotated genome sequence near the beginning of the project.

## 2.2 Task 2 – Genetic manipulation to decouple growth from biocellulose production

The results summarized in this section will be published in the following manuscript:

> Fuller, M.E., McClay, K.R., Andaya, C., Kumar P., and Reed, J.L. Development of an inducible genetic expression system for biocellulose producing *Gluconacetobacter xylinus* strain 399 (in preparation).

### *2.2.1 Goal and Introduction*

The goal of this task was to identify and evaluate a controllable genetic expression system for *G. xylinus* which would allow growth and biocellulose production to be decoupled. This was hypothesized to be the most direct approach to increase the yield of biocellulose by minimizing the competition for the available carbon that exist in wildtype cellulose producing strains between cell growth and biocellulose synthesis.

### *2.2.2 Methods, Results, and Discussion*

#### 2.2.2.1 Proof-of-concept effects of decoupling of growth and biocellulose synthesis.

The first step towards the goal of this task was to validate the hypothesis that cellulose synthesis negatively impacts the growth rate of the cellulose producing cells. The initial proof of concept experiment was conducted with wildtype *G. xylinus* 399 (or Gx 399). Gx 399 was grown in the presence and absence of *Tricoderma reesei* and *T. longbrachiatum* (Sigma-Aldrich, St Louis, MO) in 50-ml conical tubes with 5 ml of liquid culture, shaking horizontally at 100 rpm. At the end of the growth phase, additional cellulase was added to both cultures to fully dissolve the extra-cellular matrix. The optical density measured at 600 nm ($OD_{600}$) of the cultures was measured, which is directly proportional to the concentration of cells present.

The culture that was grown in the presence of cellulase had an optical density ~2.5-fold greater than the untreated culture (Figure 2.2-1). This experiment provided a level of validation for the hypothesis that cellulose production and accumulation negatively affects biomass growth. The culture grown in the presence of cellulase was still synthesizing cellulose, but the formation of the extracellular cellulose matrix was being inhibited by the presence of the cellulase. This indicates that cellulose synthesis negatively impacts cell growth due to decreases in the mass transfer of essential nutrients (oxygen, glucose, etc.) by the extracellular cellulose matrix.

This result supported our approach to temporally separate growth and cellulose synthesis by genetically modifying the cellulose producing strains.

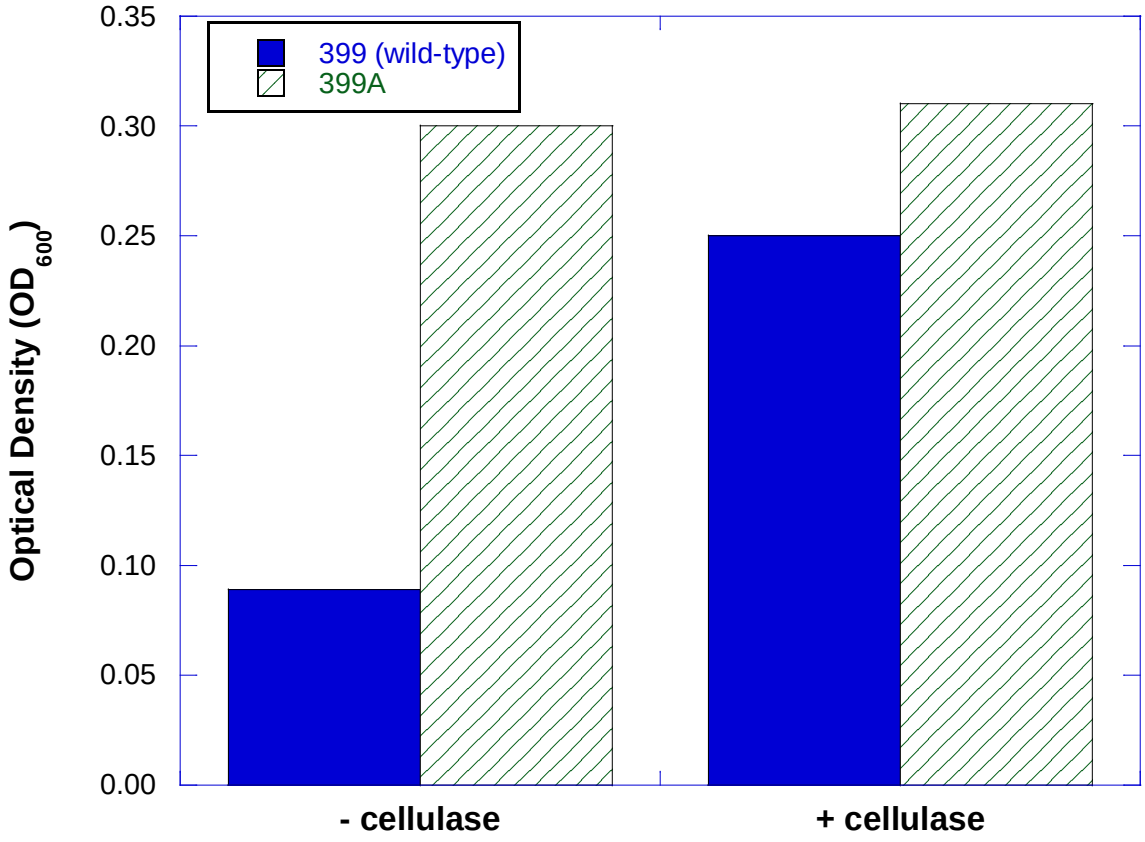



**Figure 2.2-1. Growth of wildtype Gx 399 with and without added cellulase compared to cellulose negative mutant Gx 399A.**

**2.2.2.2 DNA transformation.**

It was necessary to be able to deliver DNA into our strains (termed transformation) in order to genetically engineer the organisms. However, the development and evaluation of a protocol for effective transformation is not always a straight forward process based on a number of factors that must work in concert with each other for successful DNA transformation to occur, and for the transfer of DNA to be detected (detailed below in section 2.2.2.4).

Numerous methods to introduce DNA were evaluated, including conjugation, chemical transformation, and electroporation. Ultimately, all successful attempts to introduce DNA into *G. xylinus* involved electroporation carried out essentially as described in Florea et. al. 2016 (19).

Transformation efficiency was never systematically determined but was estimated to range from 100 to 1000 times less efficient than similar protocols carried out with *E. coli*. This extremely poor transformation efficiency precluded the possibility of delivering ligated DNA directly to Gx. Rather, the DNA was first introduced into *E. coli*, so that the DNA could be analyzed for proper assembly, tested for gene expression when applicable, and used to generate large enough amounts of plasmid DNA to conduct transformation with Gx.

Even with the intermediate step using *E. coli*, the poor efficiency, coupled with long recovery times, lead to many failed attempts to introduce DNA. The average transformation experiment required 2 days for cell preparation, 1-3 days for recovery, 4-5 days for colony growth, and 1-3 days for colony analysis. To increase throughput and the chances of successful transformation, the standard procedure during this project was to prepare electro-competent Gx cells on three successive days and to perform 2 to 3 different transformations with each new plasmid construct.

**2.2.2.3 Knock out of key cellulose synthase gene.**

In order to better understand the mechanism by which cellulose synthesis inhibited growth, and to create one of the key mutations in the cell line to be used in later investigations, the gene (*acsA*) for the enzyme responsible for both linking glucose monomers together and secreting the cellulose fiber as it was elongated was inactivated. This was accomplished by replacing a section of the chromosomal *acsA* gene sequence by recombination with a gene encoding for ampicillin resistance (*ampR*) using PCR and standard molecular biology techniques (Figure 2.2-2).

After electroporation, cells that had taken up and incorporated the *acsA-ampR-acsA* construct were selected on media containing ampicillin. PCR with primers designed to amplify the remaining end of *acsA* and the *ampR* gene used to confirm the selected mutants carried the *acsA-ampR* construct, and to assess the orientation of *ampR*. None of the mutants that were ampicillin resistant and verified by PCR to contain the *acsA-ampR* construct made cellulose under the conditions tested. A single mutant strain, designated Gx 399A (or Gx 399A) was selected for future efforts. A similar mutation was also made using a chloramphenicol resistance gene, creating the non-cellulose producing mutant strain Gx 399C.

As a follow-on to the experiment described in 2.2.2.1 above, wildtype Gx 399 and the two cellulose negative mutants, 399A and 399C, were grown with an without cellulase to see what the effects on growth would be. Cells were grown in triplicate 125 ml flasks in 20 ml of media. The cell density was measured at set intervals after addition of cellulase to subsamples of the culture.

In presence of cellulase, a significant reduction in the growth rate of all the cultures was observed, regardless of their ability to produce cellulose (Figure 2.2-3). Further, with Gx 399C, a decrease in growth was observed after ~48 hours of incubation, which was not observed with Gx 399 or Gx 399A. It was noted that the pellicle formed by Gx 399 was not really cohesive and could be disrupted by vigorous shaking of the flask, in contrast with the previous experiment in which the pellicle could not be disrupted without the addition of cellulase. These findings highlighted how growth conditions (e.g., 50 ml conical tubes vs. 125 flask) dramatically influence cellulose production, and further indicate that cellular control of cellulose production (quantity and quality) is highly complex.

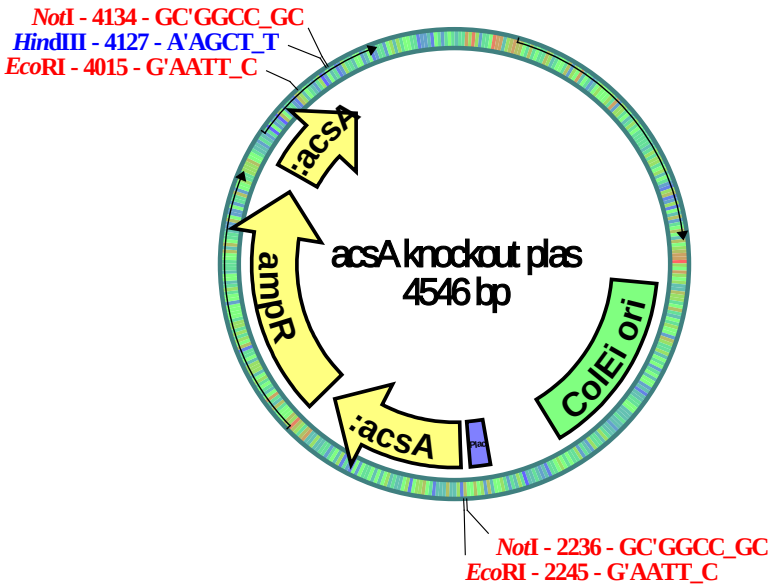



**Figure 2.2-2. Plasmid utilized to generate knockout of key cellulose synthesis gene *acsA*.**

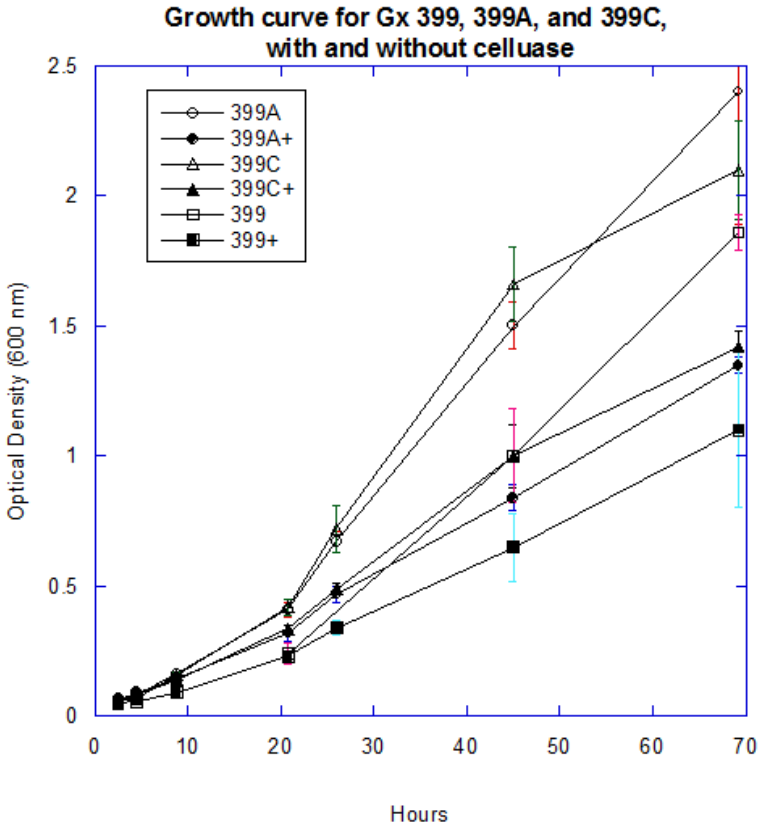



**Figure 2.2-3. Growth of wildtype Gx 399, and two cellulose *acsA* knockout mutants in the presence and absence of cellulase.**

**2.2.2.4 Design of controllable gene expression system.**
The results described above confirm that the first part of the hypothesis stating that temporarily separating cellulose production and growth will lead to an increased growth rate. The next step was to develop a controllable gene expression system (and ideally one that is tightly regulated and strongly expressed) that would allow cellulose synthesis only when the cellular biomass had been produced. The system would drive the expression of an intact *acsA* gene to compliment the deactivated copy of the gene in the chromosome of the mutants. Many similar systems are known for *E. coli*, but not a single controllable expression system had been described for *G. xylinus* at the start of this project.

Apart from an effective means to transfer DNA into the host, the desired expression system consists of a number of elements, including:

1. A selectable marker that allows for identification of transformed cells (such as an antibiotic resistance gene);
2. An origin of replication that allows the plasmid to be stably maintained and passed to daughter cells as the culture grows;
3. A DNA promoter that can be turned on and off (preferably in an incremental fashion, like a dimmer switch and with an inducer that is cheap and non-toxic), and;
4. A reporter gene that indicates whether the promoter is on or off. [The reporter gene is ultimately not part of the final production strain, but is necessary to understand how well the other parts are working during process development.]

As most of these elements are interdependent, multiple permutations are required to find an effective system. We consulted the existing published literature to select starting points for this project regarding which elements work together. However, some of these finding proved to be misleading and/or were otherwise irreproducible during this project.

**2.2.2.4.1 Determining antibiotic resistance profile of target strains.**
In order to facilitate which selectable markers could be used during this project, strains ATCC 53582, ATCC 23769 (53582 and 23769, respectively, henceforth), and Gx 399 were screened for resistance to the common laboratory antibiotics ampicillin, kanamycin, streptomycin, gentamycin, chloramphenicol, and tetracycline.

All of the cell lines tested were resistant to gentamycin. They were all susceptible to the other antibiotics in liquid culture and on solid media, but elevated concentrations (3-fold more than what is normally used in our laboratory) were required to prevent growth. The concentrations needed to reliably suppress broad based growth on agar plates were as follows:

| | |
|---|---|
| ampicillin | 300 µg/ml |
| kanamycin | 150 µg/ml |
| streptomycin | 150 µg/ml |
| chloramphenicol | 75 µg/ml |
| tetracycline | 45 µg/ml |

Even at 4-fold normal concentrations, spontaneous resistant mutants were regularly observed with kanamycin and chloramphenicol. The most reliable and effective antibiotic for these three strains was ampicillin.

Because of the strain's propensity to develop resistance to antibiotics, the validity of many of the negative results generated during this project is equivocal. Therefore, findings that are uncertain have be noted throughout the report, and are likely good places for further development.

#### 2.2.2.4.2 Identifying plasmids that replicate in chosen Gx strains.

Once the antibiotic resistance profile for these strains was determined, a literature search was conducted to identify broad-host range plasmids which carried one or more of these antibiotic resistance markers, with particular attention given to plasmids known or believed to replicate in *Gluconacetobacter* (or closely related strains).

One plasmid known to have a very broad host range and was reported to replicate in *Gluconacetobacter* (61) is pBBR1 (Figure 2.2-4) (2). The plasmid carries the kanamycin and chloramphenicol resistance markers. A derivative of pBBR1 was created using site directed mutagenesis to remove the single *Eco*RI and *Hind*III restriction sites, thus making cloning of heterologous genetic elements (e.g., *acsA* from Gx, various reporter genes) easier. The resulting plasmid was designated pBEH (Figure 2.2-5).

The first attempt to transform this plasmid into Gx strains appeared to be successful, as indicated by the growth of isolated kanamycin resistant colonies on the agar plates. This, paired with the previously reported ability of this plasmid to replicate *Gluconacetobacter*, were taken as presumptive evidence that pBEH replicated in various Gx strains.

Based on presumptive evidence indicating replication of the pBEH plasmid, work proceeded to create multiple promoter/reporter combinations to be used for testing in Gx strains. During the construction of the reporter plasmids, the incoming heterologous DNA was frequently cloned into either the reading frame of the kanamycin or chloramphenicol genes. This removed the functionality of the affected gene, leaving only one selectable marker in the plasmid, thereby increasing the chances that a spontaneous mutant would arise. Following transformation with these reporter plasmids, colonies were found growing on the plates with the appropriate antibiotics. Studies to determine if the promoter/reporter systems were functional in Gx yielded negative results. These colonies were eventually determined to be spontaneous antibiotic resistant mutants (as mentioned above) rather than actual transconjugates.

After these negative results with the pBBR-based pBEH plasmid, additional broad host range plasmids were tested for the ability to replicate in Gx. These plasmids included versions of the RSF1010 plasmid (https://www.ncbi.nlm.nih.gov/pubmed/417070) and the RK2 plasmid (https://www.ncbi.nlm.nih.gov/pubmed/7043468 ) with various antibiotic markers. A derivative of the broad-host plasmid known as RK2 (mini-RK2, Figure 2.2-6) was successfully transferred into and replicated in Gx 399, as supported by the presence of colonies on ampicillin containing plates. PCR amplification of the *ampR* gene, and isolation and transfer of mini-RK2 back into *E. coli*, confirmed that mini-RK2 was a viable plasmid vector for controllable biocellulose

synthesis. This was further corroborated with recent literature reports that the RK2 plasmid can be stably maintained in Gx strains (43).

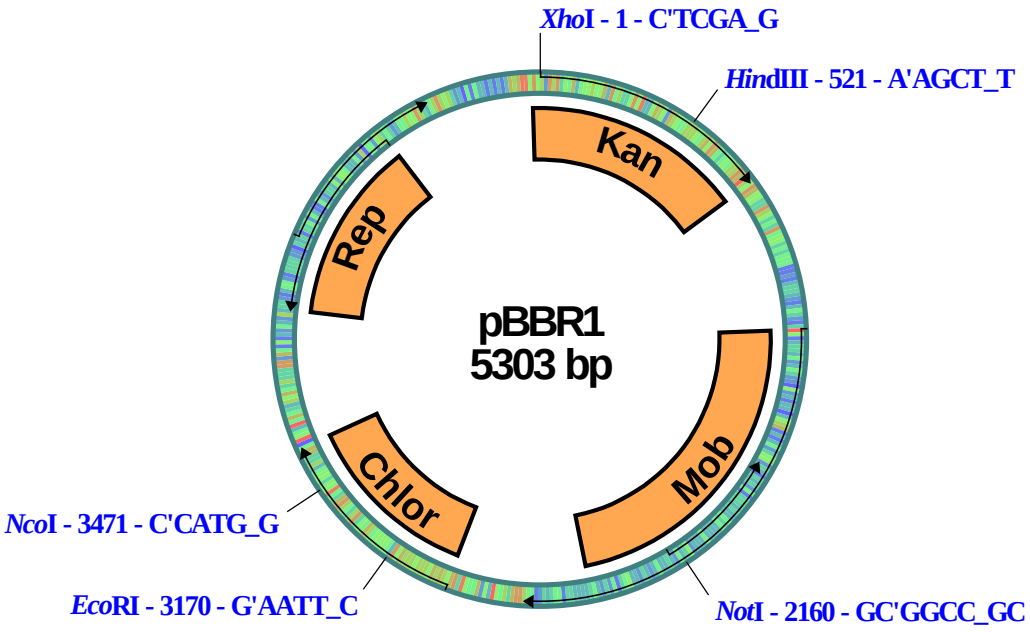



**Figure 2.2-4. Map of plasmid pBBR1.**

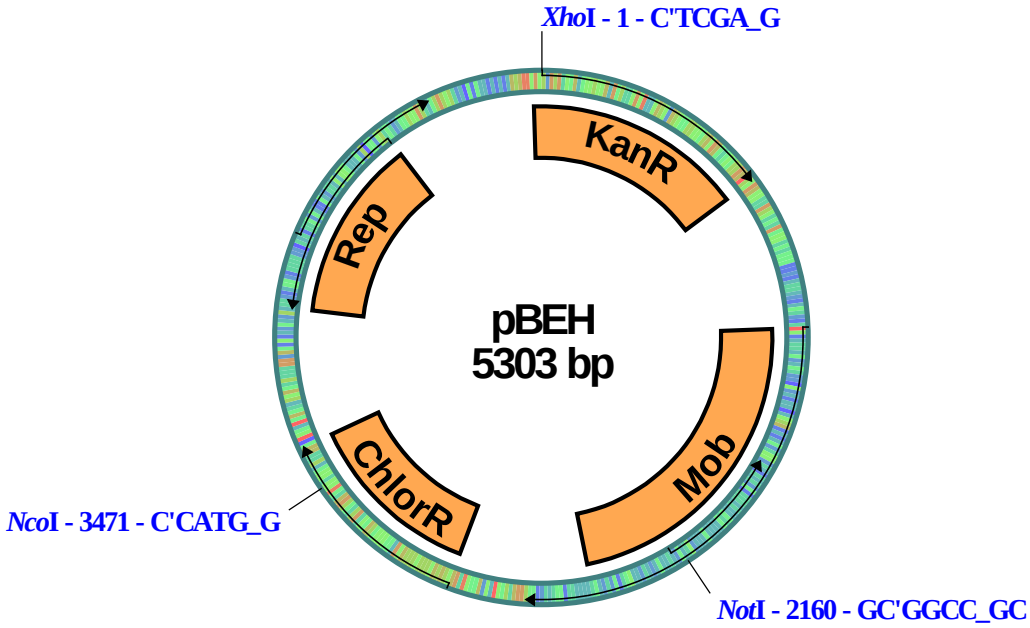


**Figure 2.2-5. Map of plasmid pBEH, a derivative of pBBR1 with the recognition site for the endorestriction nucleases *Eco*RI and *Hind*III removed via PCR.**

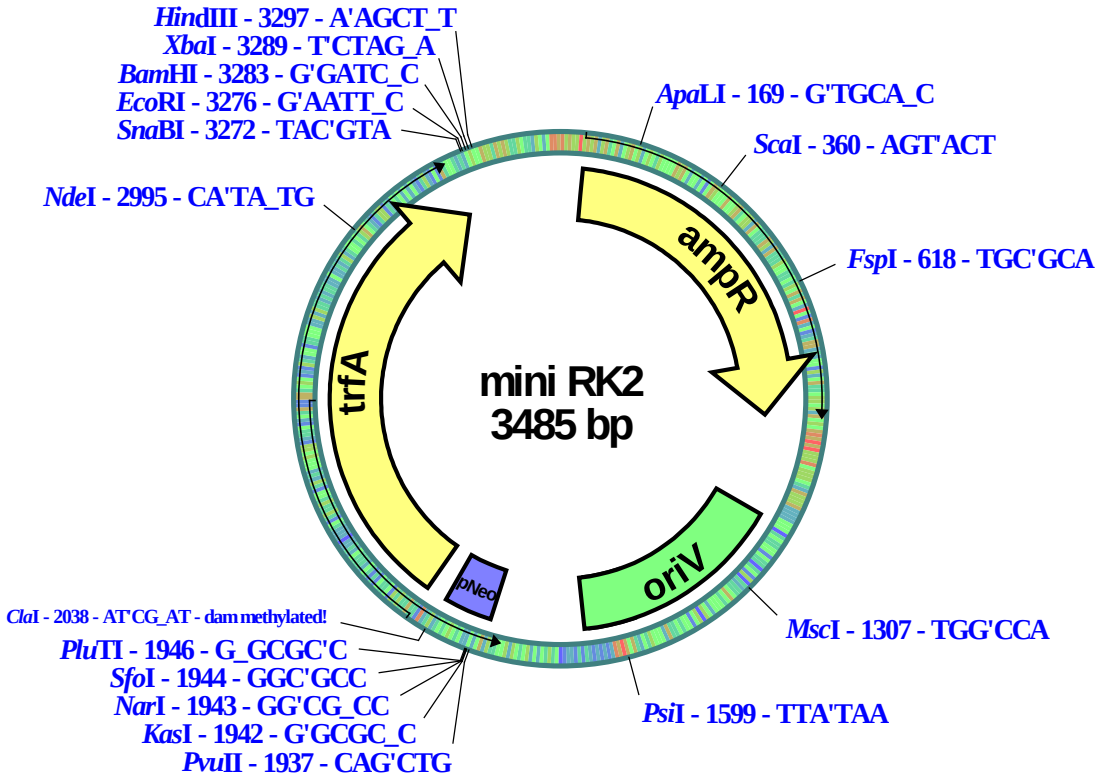



**Figure 2.2-6. Map of the mini-RK2 plasmid.**

Further experiments were performed to replace the *ampR* gene with the tetracycline resistance gene (*tetR*) into mini-RK2. This was done to create a plasmid that was compatible with biocellulose negative mutants Gx 399A and Gx 399C. Efforts to introduce the new plasmid into these mutants were not successful. When a mini-RK2 construct which included both *ampR* and *tetR* was used, not only could tetracycline resistance never be established in Gx 399, but ampicillin resistant colonies were never isolated. This seemed to indicate that the *tetR* gene product itself was toxic to Gx 399, thus preventing the otherwise stable mini-RK2 from establishing itself. This conclusion was supported by observations that the tetR derivative of mini-RK2 decreased the growth rate of the *E. coli* host, and seemed to cause the cells to be more prone to lysis. The RK2 origin of replication paired with the *ampR* was deemed the most reliable construct for genetic engineering.

These results, put into the context of the antibiotic resistance profile of the strains that were being studied, lead to the following conclusions:

- The tetracycline resistance gene proved to be problematic and would not be recommended for use in Gx expression vectors.

- Kanamycin and chloramphenicol resistant markers were deemed useful, provided that routine monitoring of plasmid retention was performed. For example, the chloramphenicol resistance gene seems to be stable when inserted into the chromosome (as evidenced by strain Gx 399C mentioned above).
- When used as a plasmid selectable marker, PCR confirmation of plasmid retention would need to be done frequently to make sure spontaneous resistance did not arise and result in plasmid loss.

#### 2.2.2.4.3 Identification of reporter genes that function in Gx 399

While the promoter and reporter in genetic expression system are intimately tied together, the results of reporter identification and screening will be discussed here, and the results of the promoter identification and screening will be discussed in the following section.

The main results of the reporter screening are summarized in Table 2.2-1.

**Table 2.2-1. Summary of promoter and reporter screening results.**

| Reporter | Promoter | Functional in *E.coli*? | Functional in Gx? |
|---|---|---|---|
| Luciferase | pM | Y | N* |
| Esterase | pM | Y | N* |
| | pDctA | ND | N* |
| Lipase | pPOB | Y | N* |
| T4MO | pM | Y | N* |
| | pDctA | ND | N* |
| | pTac | Y | N* |
| GFP | pM | Y | N* |
| | pAra (pBad) | Y | N* |
| | pAlk | Y | N* |
| | pVan | ND | N* |
| Beta-galactosidase | pVan | Y[a] | Y[a] |
| | pAlk | Y | TBD[c] |
| | pNeo | Y[b] | Y[b] |
| | pL | TBD[c] | TBD[c] |
| | pTbu | TBD[c] | TBD[c] |
| | | | |

*Indicates results were equivocal due to possible artifacts arising from spontaneous antibiotic resistance in Gx strains.
ND, not determined.
[a] Expressed, but with inducer interference (see below).
[b] Constituitive promoter.
[c] To be determined; experiments in progress at time of reporting.

*2.2.2.4.3a Luciferase Reporter*

The luciferase operon from *Vibrio fischeri* (3), which has been employed as a reporter gene to create many biosensors and reporters, is ~9 kb long and is comprised of the five genes designated *luxCDABE*. The essential components are *luxA* and *luxB*, which code for the two proteins which simultaneously convert fatty acid aldehydes to fatty acids and consume $O_2$ to convert flavin mononucleotide ($FMNH_2$) to FMN, producing light in the process. The *luxCDE* genes encode for proteins which convert the long chain fatty acids back into fatty acid aldehydes and are not necessary if exogenous fatty acid aldehydes (typically nonyl aldehyde) are added to the assay. The other reactants (FMNH2 and $O_2$) are taken from the host and medium.

During this project, the *luxAB* genes were amplified via PCR and placed under the control of the pM promoter from the broad-host plasmid pNM185 (42) and placed into *E. coli*. Upon addition of benzoic acid (the inducer of the pM promoter), light was produced and the promoter-reporter was deemed to be fully functional. However, several attempts to transform Gx 399 with this plasmid were made, but no evidence of luciferase activity was ever detected.

*2.2.2.4.3b Esterase Reporters*

Several esterases, with varying substrate specificities, were screened as potential reporters. Esterases produce a detectable signal by several mechanisms, including 1) cleaving long or short acyl groups from lipids (such as tributyrin) to produce a clear zone from an emulsified substrate contained in agar plates; 2) causing a pH change by the release of fatty acids, or; 3) liberating a chromogenic or fluorogenic compound from acylated precursors.

A two stage search protocol for putative esterases was used, starting with the complete genome sequence of Gx 23769. Five putative esterases were identified in the chromosome of Gx 23769 and were then subsequently found in the newly sequenced genome of Gx 399. The esterases were amplified via PCR and cloned in various plasmids for testing in *E. coli.*

Esterase activity from esterase #2 and esterase #5 was verified in *E. coli* using tributyrin as a substrate. Esterase #2 was cloned behind the Gx specific *dctA* promoter in the pBEH plasmid while esterase #5 was cloned behind the pM promoter in the same genetic vehicle. However, activity from these added esterases in Gx 399 was not definitively detected due to the high basal level of Gx 399's native esterase activity. Removing all of Gx 399's native esterases was beyond the scope of this project, so esterase reporters were removed from consideration.

*2.2.2.4.3c Lipase Reporter*

The lipase (a specific type of esterase) from *Bacillus subtilis* cleaves ester bonds (40) was tested as a reporter. Similar to esterases, activity of lipases can be detected using a number of substrates and formats including the cleavage of the long chain triacyl glycerol triolein (clearing zone from emulsified substrates in agar plates) or cleavage of long chain acyl groups linked to fluorescent or chromogenic substrates (p-nitrophenol or methylumbelliferone). This was deemed a good prospect, given that the native esterases from Gx 399 would not likely react with these substrates.

The lipase gene was cloned behind the pPOB promoter taken from *Acinetobacter baylyi* ADP in a derivative of the pBEH plasmid. The reporter plasmid was tested in *E. coli* and found to be

fully functional. Attempts to transform Gx 399 with this plasmid were made, but no evidence of lipase activity was detected.

*2.2.2.4.3d Toluene-4-monooxygenase (T4MO) Reporter*

Another reporter evaluated was the holozyme of toluene-4-monooxygenase, which is made up of five separate peptides that catalyze the first step in converting toluene into TCA cycle intermediates (68). The enzyme also converts the colorless precursor indole into the intensely pigmented textile dye indigo.

The reporter gene was cloned under the control of the pM promoter in plasmid pNM185 and a pBEH derivative as well as the *dctA* promoter in a pBEH derivative. The dctA promoter was not expected to be functional given its reliance of host specific protein for regulation, so it was not tested in *E. coli*, but the T4MO reporter was fully functional under the pM promoter in *E. coli*. However, potential Gx 399 transformants incubated in the presence of indole yielded no indication of indigo formation under either promoter.

*2.2.2.4.3e Green Fluorescent Protein Reporter*

The green fluorescent protein (GFP) and its derivatives have become one of the most widely used reporter genes and were thus screened during this project. The wildtype GFP (wtGFP) is naturally produced by the jellyfish *Aequorea victoria i*s a relative small gene of less than 800 bp that produces a protein that emits green light (509 nm) when illuminated with light of a shorter wavelength (~395 nm) (51). Once folded and activated, GFP does not require any co-factors (unlike luciferase) to produce a signal. Oxygen is required to activate the fluorophore of the protein by promoting the cyclization of serine 65, tyrosine 66 and glycine 67, but is not needed thereafter. Many variants have been discovered or created that activate faster, emit more light upon excitation, or possess different excitation and emission wavelengths.

For this project, the wtGFP was placed behind the pM promoter, the arabinose promoter, alk promoter, and the vanillate promoter in various combinations with the pBEH origin and the RK2 origin. GFP was active in E. coli under the pM, pAlk, pAra, and pVan promoter(s). However, none of these plasmid based constructs resulted in detectable expression of the GFP gene in Gx 399.

Most notable among these nonfunctional constructs was the plasmid comprised of the *ampR* resistance marker, the RK2 origin of replication, that arabinose promoter, and GFP (mini-RAG). The RK2 origin and *ampR* gene both function in Gx 399,But for some reason, the mini-RAG could not be transformed into this strain, yet other mini-RK2 plasmids without GFP (used as positive controls) could be transformed. Interestingly, during our work on this, two articles were published by the same group presenting conflicting findings as they pertain to the expression of GFP in Gx strains.

In the first paper, the group attempted to transform *G. hansenii* ATCC 53582 (formerly *G. xylinus*) with a plasmid backbone (no reporter gene) designated pSEVA331-Bb, which proved difficult, but was ultimately successful (19). However, all their attempts transform 53582 with derivatives of pSEVA331-Bb that carried fluorescent proteins failed. In an attempt to explain this phenomena, they examined the genome of 53582 for the presence of DNA restriction

enzymes and found two (*Mrr* and *Pst* I). Because of the substrate specificity of *Mrr*, it was deemed unlikely that this enzyme was causing the problem. However, a *Pst I* site was present in the plasmids carrying the fluorescent protein. Site directed mutagenesis was used to remove the site from the plasmids, which did not fix the problem. Further, only one plasmid carrying a gene for a fluorescent protein (pBAVIK) was able to replicate in 53582, but no GFP expression was observed. Sequencing of the plasmid revealed that there was a deletion in the promoter that drives the expression of GFP, thus there was no expression of GFP. This implies that fluorescent proteins are themselves the problem for 53582.

In a second paper by the same group, it was noted that *K. rhaeticus* (another member of cellulose producing Aceotobacter group to which *Gluconacetobacter* belongs) can be transformed with plasmids carrying non-standard fluorescent protein including the red fluorescent protein variant mRFP1(functional as a monomer as opposed to the wildtype tetrameric form), and two variants of GFP that have been mutated to have enhanced brightness and undergo a more rapid folding/activation process (GFPmut3 and superfolder GFP), while four other fluorescent protein were not expressed (18).

The reason the GFP was problematic during this project with Gx 399, and for the closely related 53582, is not completely understood. However, it is of interest to note that only enhanced versions of fluorescent proteins worked in *K. rhaeticus*, which implies that some form of proteopathy related to misfolding or slow folding RFP and GFP might be at play. The use of these alternative fluorescent protein genes as reporters was not able to be investigated during this project, but warrants further consideration.

*2.2.2.4.3f LacZ Reporter*

The *E. coli* enzyme beta-galactosidase (coded by the *lacZ* gene) is also a very commonly used reporter gene. Chromogenic (X-gal, ONPG, etc.) and fluorescent (Lumigal-530) substrates have been developed so that LacZ enzymatic activity can be measured in numerous ways. It has been adapted to function well in a broad range of bacteria, plants, viruses, mammals, and as a histochemical stain.

A genome search revealed there were no obvious homologs of *lacZ* in the genome of Gx 399 or 23769, and neither of these cell lines converted the X-gal substrate into the indigoid chromogen on agar plates. This indicated that LacZ would make a useful reporter if it was functional. The *lacZ* gene was cloned behind a constitutive promoter (pNeo) and later the pVan promoter from 399 in the mini-RK2 plasmid and transformed into Gx 399. Possible transformants were spread on agar plates with ampicillin and X-gal. The presence of blue colonies were observed, indicating a functioning reporter gene was identified for Gx 399. Pairing this reporter with a functional inducible promoter would be expected to allow study of the induction dynamics of the promoter, but was not completed during this project due to time limitations.

**2.2.2.4.4 Identifying promoters that functional and regulated in Gx 399.**

The promoters selected and screened for use in strain Gx 399 are detailed below. A range of promoters previously described to function in a broad range of bacterial strains were considered. In addition, a large amount of effort was directed towards identifying and screening promoters that were functional and inducible in our Gx strains.

*2.2.2.4.4a Antibiotic Resistance Promoters (pAmp and pChlor)*

The constitutive promoters that drive the expression of the antibiotic resistance genes *ampR* and *chlorR* both function in Gx 399 as evidenced by the fact that when theses gene were inserted into the chromosome, the cells were resistant to the antibiotics ampicillin and chloramphenicol. This is not a particularly surprising finding, as these genes are frequently found of transmissible genetic elements that are often passed from one species to another and thus, have evolved to work well in many genetic backgrounds. These promoters are of limited use in this project as they are in an always "on" and do not represent a useful mean of control protein production. However, they could be of use in driving the constituitive expression of CDG-insensitive isoforms of the *acsA* gene as described in the following section of this report, if needed.

*2.2.2.4.4b pM Promoter*

The pM promoter, originally identified in the TOL plasmid of *Pseudomonas putida* pWW0, is positively regulated by XylS protein when the inducer, benzoic acid is present (69). This promoter was selected because it has been used successfully to regulate the expression of various genes in multiple species (7, 42).

The pM promoter successfully drove expression of luciferase, T4MO, GFP, and esterase #5 using the RSF1010 based plasmid pNM185 (42) in *E. coli*, However, all these constructs failed in Gx 399 in the RSF1010 based plasmid, as well the pBEH plasmid.

*2.2.2.4.4c pAlk Promoter*

The pAlk promoter is activated by the AlkS protein in the presence of octane. In its native host, *P. oleovarans*, this system regulates the expression of the genes needed to grow on octane (8). Like the pM promoter, pAlk has been deployed in broad host plasmids to drive the expression of various genes in multiple Gram-negative species.

It was hypothesized that the hydrophobic octane inducer of pAlk would have less difficulty in passing through cell membranes than more polar and/or charged inducers (e.g., sugars and organic acids) that Gx 399 may not effectively transport into their cells. All of the constructs created and tested using pAlk where linked to reporters GFP and LacZ. While these constructs were very well regulated in *E. coli*, none of them worked in Gx 399 (*lacZ* test in progress at time of this report). The cause of this failure was not further investigated.

*2.2.2.4.4d pPOB Promoter*

The pPOB promoter functions as the driver for the p-hydroxybenzoate hydroxylase (PobA) in *Acinetobacter calcoaceticus* ADP1 (15). The regulator of this operon, PobR, is a mixed function genetic regulator, negatively regulating its own expression, while serving as a positive regulator of *pobA*. In the absence of PobR, the promoter that drives expression of *pobR* gene is in the "on" position. When PobR is produced, it binds to the promoter region, shutting down PobR production. Once PobR is present and bound to the promoter/operator region of the operon, the binding of benzoic acid causes a conformational rearrangement of PobR, turning on the expression of the PobA.

In previous work, we found that a pPOB-*lacZ* regulated by *pobR* had essentially zero background expression and was highly inducible in *E. coli* (>400 fold), making it an interesting promoter candidate. Further it has been reported that in strain ADP1, there was a transporter for p-hydroxybenzoate that had a high degree of substrate overlap with the vanillate transporter present in Gx 399. The promoter was paired with the lipase reporter in the pBEH plasmid. This construct functioned well in *E. coli*, but no lipase activity was ever seen in Gx 399 due to unknown reasons.

*2.2.2.4.4e pAra Promoter*

The native host of the arabinose responsive promoter, pAra (also known as pBad), is *E. coli*. It has a scalable induction range of 1200 fold (21). Broad-host genetic elements have been developed that function in *E. coli*, *Pseudomonas aeruginosa*, *Xanthomonas campestris* (62), and *Agrobacterium tumefacians* (47). It should be noted that in some cases the arabinose specific transporter, AraE, must be present for the system to function (71).

During this project, the pAra promoter was cloned along with the wtGFP reporter into the mini-RK2 plasmid, with and without *araE*. The system worked in *E. coli*, but did not lead to detectable expression of GFP in Gx 399, again for unknown reasons.

*2.2.2.4.4f pTac Promoter*

The pTac promoter (a hybrid of the *trc* promoter and *lac* promoter (39)) has been deployed in RSF1010 type broad-host plasmids and in transposons in different bacterial species (14). It can be induced with either lactose (which requires transport) or with IPTG, which freely penetrates many cell types.

This promoter was tested in the broad-host plasmid pVLT31de Lorenzo, 1993 #278} or pBEH. The only reporter tested in these constructs was T4MO, which worked in *E. coli*, but failed to produce a detectable signal in Gx 399.

*2.2.2.4.4g Endogenous Promoters*

The genomes of 23769 and Gx 399 were manually examined for the presence of endogenous inducible operons that met the following criteria:

i) promoters must not be associated with genes that are essential to the survival of the cells;
ii) the promoters must function under the anticipated conditions in which the BC will be produced;
iii) the expected promoter induction mechanism must be simple and safe to implement (e.g., the inducer is a single, non-toxic compound), and;
iv) the induction of the promoter would be expected to result in minimal metabolic burden and no undesirable shift in the organism's metabolism (e.g., assure that the given inducer, if it is a utilizable carbon source, does not adversely impact growth or cellulose synthesis).

The first endogenous promoter, pDctA, was identified as a region assumed to be turned "on" in the presence of four carbon dicarboxylic acids in the external environment (or "C4", e.g.,

succinate and fumarate), and would probably be inactive in the presence glucose (or glycerol) that would be fed to the Gx strains during BC production, was selected as the initial promoter candidate. Two pBEH based constructs were assembled using T4MO and esterase #2, and one genomic insert using the kanamycin resistance gene selectable marker were made. All three showed some basal (e.g., non-inducible) level of activity in *E. coli*. However, no activity was ever observed in Gx 399.

A second inducible operon in both 23769 and GX 399 that met the criteria was also identified. This operon consisted of a promoter, regulator, and the genes which catalyze the demethylation of vanillate. The general regulatory scheme for this operon is the same as the one described for pPOB above but is native to Gx 399.

The inducibility of the putative pVan promoter was tested by measuring the expression of *vanA* via reverse transcriptase-qPCR (RT-qPCR) before and after exposure to vanillate. Results of this test showed that the operon had some degree of basal activity (which is common), but the expression could be increased by at least 100 fold in the presence of vanillate (Table 2.2-2). While this level of analysis was not applied to every promoter tested, this was the first inducible promoter that was proven to function in Gx 399, and provided a required element needed to make a reporter plasmid that functioned in Gx 399.

| Vanillate concentration (µM) | Crossing threshold |
|---|---|
| 0 | 20.5 |
| 60 | 13.6 |
| 120 | 13.1 |
| 240 | 13.7 |
| 480 | 12.8 |
| 960 | 13.4 |

**Table 2.2-2. RT-qPCR analysis of chromosomal based production of vanillate demethylase mRNA as a function of vanillate concentration.**

Following confirmation of pVan inducibility, a suite of mini-RK2 plasmids were constructed with various iterations of the vanillate promoter and the *lacZ* reporter. The first was mini-RK2VZ. When *E. coli* cells carrying the plasmid were grown on agar plates containing X-gal (chromogenic substrate for LacZ) the colonies turned blue with no obvious difference caused by the presence or absence of vanillate in the vapor phase (very low concentration), or when incorporated into the growth media (higher concentration). However, when this construct was inserted into Gx 399, the cells turned blue faster when vanillate was present in the vapor phase than when it was absent, indicative of induction and in line with what was expected based on the RT-qPCR results. However, adding vanillate to the agar did not did not cause early onset color development for unknown reasons.

More quantitative induction studies were conducted, in which the GX 399 cells containing the pVan-LacZ construct were grown in liquid culture with varying concentrations of vanillate, followed by a measurement of LacZ activity using a luminometer to detect the conversion of the Lumigal-530 substrate (Figure 2.2-7). Unexpectedly, increasing concentrations of vanillate caused a decrease in LacZ activity. A similar effect was seen with the first metabolite of the vanillate degradation pathway, protocatachuate (Figure 2.2-8).

The apparent high level of constitutive activity of the synthetic pVan-*lacZ* operon, paired with anomalous reduction in activity caused by the inducer of the operon, instigated a reexamination of the parts used to create the synthetic operon. As shown in Figure 2.2-9, the first attempt to build an synthetic operon based pVan, its cognate regulatory protein vanR, and lacZ contained the -35 and -10 sequences (as identified by Neural Network Promoter Prediction) required for gene transcription. However, a careful manual examination of the DNA sequence between the intended promoter and the start codon for LacZ revealed that the cloning procedures might have inadvertently generated a second promoter region by placing two partial TATAAT (-10) sequences downstream from a naturally occurring partial -35 sequence.

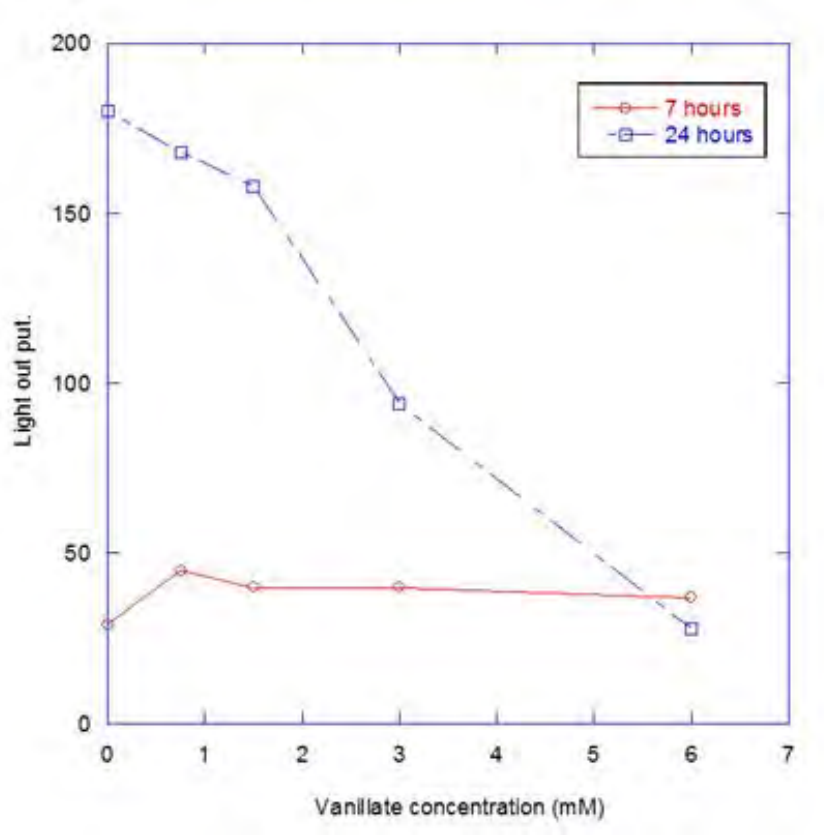



**Figure 2.2-7. Inhibition of LacZ activity in Gx 399 (mini-RK2VZ) as a function of vanillate concentration and duration of incubation period.**

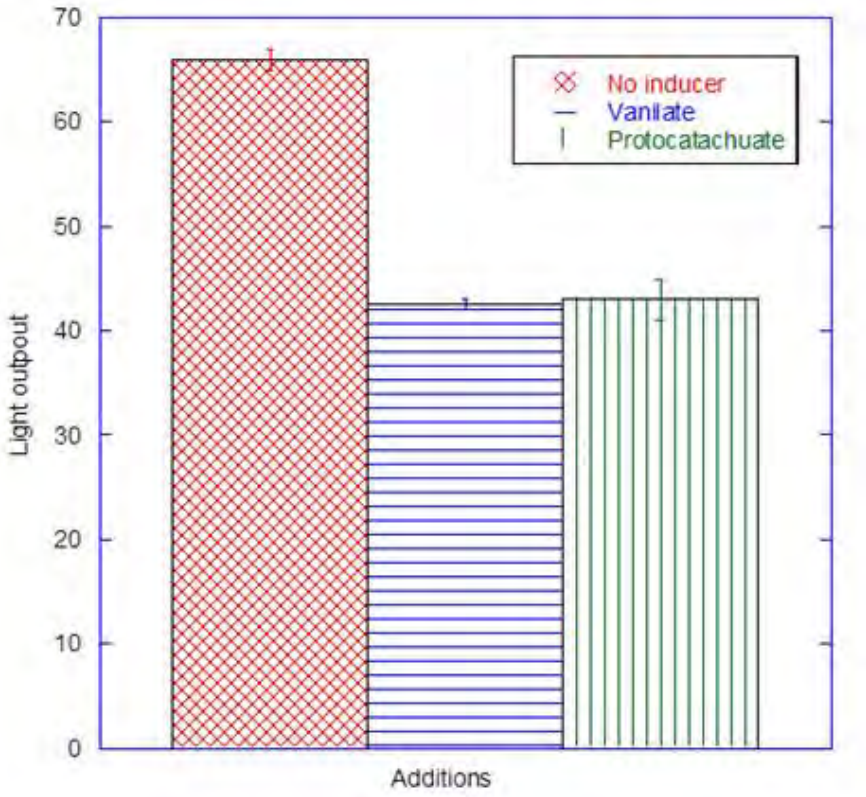



**Figure 2.2-8. Inhibition of LacZ activity in Gx 399(mini-RK2VZ) by vanillate and the first downstream metabolite, protocatachuate.**

To preclude the possibility that a second synthetic operon was created during cloning, the leader sequence preceding the lacZ ribosomal binding site was shortened by 16 bp. Furthermore, the pVan/lacZ fusion site was moved downstream, relative to the native vanillate operon, such that DNA coding for the first 23 or 30 amino acids of the VanA peptide precedes the ribosomal binding site for lacZ. The uncertainty regarding the number of amino acids of VanA being produced is related to the fact that in the GenBank annotation database, there is a disagreement about whether the true start codon of *vanA* is the less frequently used GTG start codon, which is 21 base pairs (equivalent to 7 amino acids in the peptide) upstream from the canonical ATG. Based on the location of the ribosomal binding site, it appears that GTG is the start codon. The resulting synthetic operon was left with only one obvious promoter, the native pVan (Figure 2.2-9). The partial -35 sequence was still present but the -10 sequences were no longer included.


**Figure 2.2-9. Sequences of variants of the *pVan-lacZ* fusion.**
DNA sequence derived from the *E. coli lacZ* gene has a gray highlight. Confirmed start codons are in red text with directional arrows underneath. Possible start codons are in green text with directional arrows underneath. Promoter and potential promoter sequences are highlighted in yellow. Putative mRNA transcription start sites are marked with a blue C.

Unfortunately, the pVan-lacZ version 2 still had significant background expression. Three other plasmids were created which utilized other putative elements of vanillate operon. The piece of DNA that makes up promoter region of mini-dual1Z2 and mini-dual2Z2 contained two different promoters and regulators. The promoter regions being at opposite ends of the engineered DNA both point outwards from their respective ends. One of the promoters was again pVan, with essentially the same make up as it has in pVan-lacZ version 2. The other was the promoter/regulator for a putative alcohol/aldehyde reductase.

Given the proximity of the two promoters in the genome of Gx 399, and the fact that a number of alcohols and aldehydes appear in the vanillate degradation pathway, it was hypothesized that the

second promoter might also respond to vanillate and/or its metabolites. This dual promoter piece was constructed so that it could be cloned into the mini-RK2 plasmid that already carried *lacZ* in either direction. The final plasmid created in this series thus far was a derivative of mini-RK2V2Z2, but with the *SnaB*I regulatory piece removed. The resulting plasmid had the potential to be used as both a constitutive *lacZ* expression platform, as well as a promoter-probe vector wherein putative promoters could be cloned in front of *lacZ* using any blunt enzyme paired with *Xba* I or *Nhe* I.

To test the variants' functionality, all of the plasmids were transferred to Gx 399 and grown in the presence and absence of vanillate. In each case, the presence of the vanillate decreased the level of LacZ activity measured at all but the lowest concentrations. A logical genetic explanation for these five plasmids to all have promoter systems down-regulated by the inducer of the operon could not be envisioned, especially in light of the up-regulation of the promoter observed in the presence of vanillate based on RT-qPCR results cited above.

One possible explanation that was explored further was that vanillate (or its metabolites) were directly inhibitory to LacZ enzyme activity. This idea had been considered earlier, but dismissed because there were literature reports describing the use of vanillate responsive promoters to drive expression of a *lacZ* reporter in other bacterial species, including *Sphingomonas*, *Myxococcus*, and *Caulobacter* (26, 29, 65). These previous reports describe "essentially no expression in the absence of an inducer" (26), and they all were reported to exhibit inducer "dose-response" characteristics (29). None of these studies reported inhibition of LacZ by vanillate at doses up to 800 µM vanillate, in direct contrast to what was seen here with whole cells.

This explanation was tested using *E. coli* and Gx 399 cells carrying the plasmid mini-RK2Z2(*SnaBI*) which constitutively produces LacZ (to eliminate any sort of genetic/induction effects). Cells were grown in 70 ml of LB medium or ATCC Medium 1765. The cells were concentrated to 3 ml in 10 mM tris:1 mM EDTA buffer pH 8.0, and lysed via bead beating, followed by centrifugation at 17,000 *x g* for 10 minutes to pellet cellular debris. The extracts were then incubated for 1hour in the presence of 3 mM vanillate (equivalent 130 µM if added to the original 70 ml culture volume) and the LacZ activity was measured using the Lumigal-530 assay. The presence of vanillate caused a ~90% decrease in LacZ activity after 1 hour with Gx 399, whereas the LacZ from *E. coli* experienced only a 10% decrease. The difference in the magnitude of LacZ inhibition in this experiment lead to the hypothesis that Gx 399 converts vanillate to an inhibitory compound that is not produced by *E. coli* (which is not known to metabolize vanillate). However, it was also noted that the *E. coli* extract had a much higher starting LacZ activity.

Follow-on experiments were designed to determine if the apparent resistance of LacZ from *E. coli* to vanillate inhibition was merely a function the amount of LacZ present at the start of the assay. This was done by diluting the *E. coli* extract with buffer and conducting dose/response experiments, measuring LacZ activity as a function of the concentration of vanillate added (Figure 2.2-10). A separate control experiment demonstrated that DMSO, which was used as a carrier solvent for the vanillate, had a minimal negative impact an LacZ (Figure 2.2-11). These dose/response studies showed that LacZ was susceptible to inhibition by vanillate regardless of the source of the LacZ (e.g., *E.coli* or Gx 399 transformant). Further, given that LacZ extracted

from *E. coli* (which is not known to metabolize vanillate) was also inhibited by vanillate, the hypothesis that Gx 399 converted vanillate to a more inhibitory compounds was also likely false. Taken together, the results suggested that vanillate itself was inhibitory to LacZ activity; side experiments also indicated that protocatechuate, the first metabolite of vanillate degradation, appears to be equally inhibitory. These results showed that pVan activity could not be measured using a *lacZ* reporter using standard LacZ assay conditions due to the previously undocumented inhibition of LacZ by vanillate.

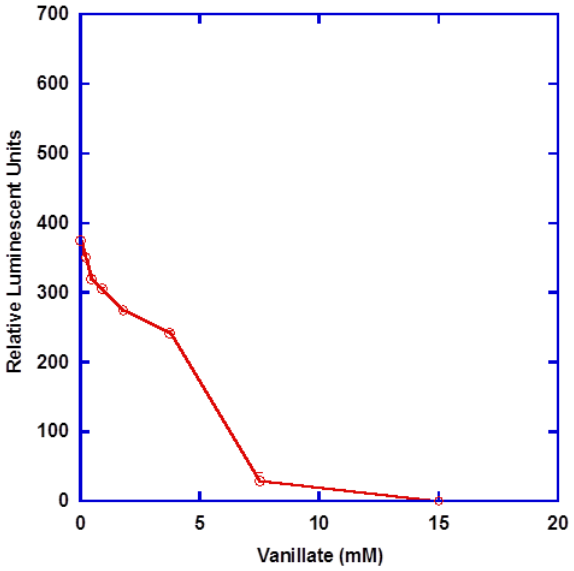



**Figure 2.2-10. Residual LacZ activity in Gx 399 extract as a function of vanillate concentration.**

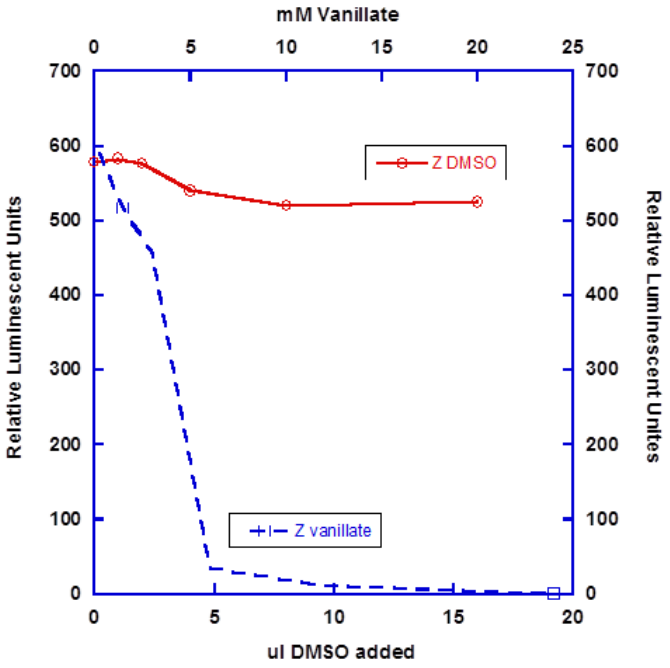



**Figure 2.2-11. Residual LacZ activity in *E. coli* extract as a function of increasing DMSO and DMSO+vanillate concentration.**

### *2.2.3 Conclusions*

The duration of this project precluded full development of a controllable gene expression system for BC production in *G. xylinus*. However, much progress was made towards that goal, with all four elements of a possible system being identified.

Research in this area will continue to build on the successful identification of the four genetic elements needed to build a robust and controllable expression system in Gx. Some of this work will involve the reexamination of negative results which might have been caused by the spontaneous antibiotic resistance. Additional work may also be performed based on the observation that GFP variants with enhanced folding characteristics seem to function in at least some Gx strains.

Three question categories will be explored:

1. Can pBBR1 (plasmid pBEH) be stably inherited in Gx 399 if the ampicillin resistance marker is incorporated?
2. Can the pPob, pAlk, etc., be used to drive scalable expression of lacZ in the mini-RK2 plasmid?
3. Is the folding enhanced variant of GFP (GFPmut3) functional in Gx 399?

The answers to these questions will shape the direction of future efforts to create the controllable expression system in Gx 399. For example, if ampicillin resistance makes the pBEH plasmids

stable, it will allow for the rapid re-examination of many promoter/reporter constructs by inserting the *ampR* gene into available restriction sites. A positive outcome from any of the three questions should allow for rapid progress in this area.

An unrelated question that is worthy of further examination involves the low transformation efficiencies in Gx strains, and the requirement for extended recovery times to isolate transformants. This situation raises the question of whether the plasmid DNA must undergo some sort of genetic or epigenetic change for the plasmid to replicate and be stably maintained. A series of tests to see if plasmids isolated from Gx 399 (as opposed to *E. coli*) are easier to introduce back into Gx 399 is planned during follow-on research. If genetic or epigenetic changes appear to be involved, studying the genetics and epigenetics (methylation patterns) of the plasmid would be a logical next step.

Work done during this project and in other unaffiliated laboratories has highlighted that strain genetic and phenotypic variability in the broader Acetobacter group means previously published findings might be noninformative at best, and misleading at worst, if different species or even different strains are used for similar work. Given these findings, it seems that a concerted effort to develop a broader suite of genetic tools applicable to a chosen biocellulose producing strain must be undertaken before the greater challenges of improving cellulose production can realistically be pursued.

## 2.3 Task 3 – Remove CDG control of biocellulose synthesis

### *2.3.1 Goal and Introduction*

In addition to controlling fermentation conditions, redirecting carbon flux pathways (metabolic engineering), and timing of when the cellulose synthase (CS) enzyme is expressed, it was postulated that modulating the cellulase synthases need for an allosteric activator could improve cellulose production.

The wildtype cellulose synthase will not produce cellulose in the absence of a small molecule (di-cyclic-GMP, CDG). While the exact mechanism by which CDG activates CS was not known when the project started, a number of different paths to address this challenge were conceived. The first was to create an isoform of CS that no longer required CDG to be active. If this were accomplished, the biggest factor controlling cellulose synthesis would be the amount of cellulose synthase produced.

If the host for CS production were as genetically amendable as *E. coli*, any number of techniques could be used to create random mutations in the CS gene, followed by screening for improved cellulose production. To help insure that the improved activity was associated with decreased dependence on CDG levels, enzymes that break down CDG could be over expressed in the host (57). However, Gx CS is minimally functional, if at all, in heterologous hosts like *E. coli*, and the selected Gx strains that produce cellulose naturally are genetically intransigent due to poor transformation efficiency.

A second way to remove dependence of CS for an allosteric activator would be to alter the intracellular concentration of CDG. This could be accomplished by knocking out CDG metabolizing enzymes (e.g., phosphodiesterases (9)), which would increase the levels of CDG. It was anticipated that this approach was worthy of investigation, but there were concerns that altering CDG concentrations might have unintended consequences, as CDG levels are known to regulate many different cellular processes (56).

A third approach would be to increase the CDG binding affinity of CS, thereby increasing the apparent CDG concentration, without impacting other cell processes. There was support for this approach in the literature that indicated that modification of a portion of the CS enzyme responsible for the binding of CDG, called the PilZ domain, was both needed and possible to affect this change (34, 54). The envisioned mutations would be created around the highly conserved five amino acid sequence (RXXXD and/or RXXXR[where R=arginine, D=aspartic acid, and X=any amino acid]) within the PilZ domain, using PCR primers to create a limited number of mutations which targeted the slightly longer 9 amino acid string QKRNSHRIP within the CS enzyme. While this approach shares some similarity to creating random mutations througout the entire CS gene (still requires transformation of mutant plasmids and screening), the number of mutations that would need to be screened would be greatly reduced.

### *2.3.2 Methods, Results, and Discussion*

Before this task started during this project, a separate research group discovered the mechanism by which CDG imparts it allosteric effect. In the absence of CDG, a “gating loop”, which is part

of the CS peptide chain, blocks the entrance to the active site of the CS (45), thereby preventing the glucose monomers from entering the active site and being added to the cellulose chain. The closed position of the gating loop is stabilized by an interaction with the first arginine in the PilZ domain (R574 of the CS) which forms an ionic bond with a glutamic acid (D) at position 369. Binding of CDG breaks the R-D salt bridge, allowing the gating loop to move out of the way as the enzyme adopts an active conformation. It was further found that replacing R574 with the smaller, neutral, alanine prevents the salt bridge from forming in the first place and results in a constitutively active cellulose synthase. The resulting synthase is ~40% as active as the fully activated wildtype enzyme, regardless of the concentration of CDG.

We incorporated these new findings into our research plan, and created two different version of *acsA* which mimic the R574 mutation referenced above: *acsA* R580A and *acsA* R580G (R574 and R580 are equivalent amino acids in the two cellulose synthases). These CS isoforms would serve as the basis for further investigation to determine if their controlled overexpression would increase the cellulose production characteristics of Gx 399. They would also serve as the basis for further investigation to see if the random mutations in and around the PilZ domain, which R580 is part of, would result in greater that 40% activity of the fully activated CS.

A partial exploration of the functionality of these *acsA* isoforms was be carried out by placing the R580A and G isoforms under the control of one of the constitutive promoters in the mini-RK2 plasmids and introducing them into the cellulose deficient strain Gx 399C. While this construct did not fulfill all of the requirements of the original project plan, it did provide proof-of-concept for modifications to AscA that make CS not dependent on CDG. The isoforms were never fully tested because the inducible expression system needed to fully test them was not completed, as detailed in section 2.2 above.

### *2.3.3 Conclusions*

As mentioned above, the scientific community has only a limited understanding of the genetics of the broader Acetobacter group, and useful genetic tools that can be applied to this group of bacteria are lacking.

An ambitious and comprehensive program for improving cellulose production was put forth for this project, which could not be fully realized because of these limitations. Through this work it has become clear that a suite of basic genetic tools must be available for the specific strain being studied, prior to attempting to apply the concepts of synthetic biology for the purpose of improving production of bacterial cellulose.

## 2.4 Task 4 – Metabolic modeling for optimization of biocellulose synthesis

The results summarized in this section here will be published in the following scientific articles:

Kumar P., McClay K., Fuller M.E., and Reed, J.L. Complete genome sequence of biocellulose producing bacteria *Gluconacetobacter xylinus* ATCC 53582 and 399 (in preparation for submission to Genome Announcements).

Kumar P., Pan S., McClay K., Zhang X., Fuller M.E., and Reed, J.L. Strategies to enhance the yield of biocellulose in *Gluconacetobacter xylinus* ATCC 53582 and 399 (in preparation for submission to Biotechnology & Bioengineering or BMC Systems Biology).

### *2.4.1 Goals and Introduction*

The importance of biocellulose is clear from its numerous applications. Hence, it is important to increase the yield of biocellulose from *Gluconacetobacter.* The goal of this task was to study the metabolic network of the two strains of *Gluconacetobacter xylinus,* ATCC 53582 and Gx 399 that were being used during the other tasks, and to identify gene deletion strategies to enhance the yield of biocellulose.

### *2.4.2 Methods, Results, and Discussion*

#### 2.4.2.1 Genome sequencing of *Gluconacetobacter xylinus* ATCC 53582 and Gx 399

Sequencing of *G. xylinus* ATCC 53582 was performed at the Biotechnology Center at the University of Wisconsin-Madison. The SPAdes assembler (4) was used to assemble the sequences into contigs (a set of overlapping DNA segments that together represent a consensus region of DNA) using the U.S. Department of Energy Systems Biology Knowledgebase (Kbase accessed at kbase.us). The genome annotation of *G. xylinus* ATCC 53582 and *G. xylinus* 399 was performed on the KBase server (28, 60).

*G. xylinus* ATCC 53582 contained 267 contigs with a maximum length of 598,932 bp and minimum length of 128 bp, and 20 contigs having more than 1000 bp. The G+C content of *G. xylinus* ATCC 53582 chromosome is 59.42% with a total of 3,506,630 bp. CheckM (52) with Acetobacter marker genes to analyze the completeness of the *G. xylinus* ATCC 53582 genome and the result shows that the completeness of the genome was 93.02%. The entire genome of *G. xylinus* ATCC 53582 had an estimated 4,051 genes. Out of these 4,051 genes, 53 were RNA genes (194 are repeats, 2 crispr_array, 37 crispr_repeat, and 35 crispr_spacer) and 3,730 were protein coding genes, including six cellulose synthase genes (CDS.5823, CDS.5824, CDS.5825, CDS.6350, CDS.4576 and CDS.4577) and an endogluconase gene (CDS.5821). The chromosome of *G. xylinus* ATCC 53582 was 226,228 bp shorter than *Gluconacetobacter hansenii* ATCC 23769 (GenBank: CM000920.1).

Genome sequencing of Gx 399 showed that the strain has a total of 4,121 genes. The sequence had 682 contigs with a maximum length of 570,077 bp and minimum length of 128 bp, with 24 contigs having more than 1000 bp. The G+C content of this strain was 59.54% with a total of 3,378,262 bp. There were 3,836 protein coding genes and 159 RNA genes. The genome

sequences of *G. xylinus* ATCC 53582 and Gx 399 were submitted to GenBank.

The quality of the sequencing, assembly, and annotation was evaluated using the recently published CheckM algorithm (4). The CheckM algorithm looks for the abundance of marker genes in an organism given its position in the phylogenetic tree. The CheckM results for the assembled *G. xylinus* ATCC 53582 and Gx 399 genomes and an already published genome for *G. hansenii* ATCC 23769 (for comparison) are shown in **Table 2.4.1**. Based on these results, the assembled genome was of similar quality to previously published *Gluconacetobacter* genomes.

Bidirectional proteome comparison was also performed for *G. xylinus* ATCC 53582 with five other *Gluconacetobacter* strains, namely, *G. hansenii* ATCC 23769, *Gluconacetobacter diazotrophicus* PAI 5 (6), *G. xylinus* NBRC 3288 (49), *G. xylinus* 399 and *G. xylinus* E25 (36). Of sequenced strains, *G. xylinus* ATCC 53582 was closest to the Gx 399 (isolated from Kombucha tea and sequenced in this project) with 2,758 common genes (**Figure 2.4-1**).

| | *G. hansenii ATCC 23769* | *G. xylinus ATCC 53582* | *G. xylinus 399* |
|---|---|---|---|
| Marker Lineage | Acetobacter | Acetobacter | Acetobacter |
| No. of Genomes Examined | 11 | 11 | 11 |
| No. of Markers | 1037 | 1037 | 1037 |
| Completeness | 92.34 | 93.02 | 93.07 |
| Contamination | 5.70 | 7.09 | 6.43 |
| Strain Heterogeneity | 0.00 | 3.95 | 1.43 |

**Table 2.4-1. Summary of CheckM results for two *Gluconacetobacter* strains.**

### 2.4.2.2 Minimal media and Biolog growth experiments

*Gluconacetobacter* strains like ATCC 53582 and Gx 399 have traditionally been cultivated in Hestrin-Schramm (HS) media (58). However, the high concentrations of yeast extract and peptone in HS media lead to chances of less pure biocellulose and do not allow an accurate assessment of the native metabolism of the strains. Therefore, *G. xylinus* ATCC 53582 and Gx 399 were grown in a basal salts minimal media (BSM) (23) containing (g/L): $K_2HPO_4.3H_2O$, $NaH_2PO_4.H_2O$, $NH_4Cl$, $MgSO_4.7H_2O$, $FeSO_4.7H_2O$, $MnSO_4.7H_2O$, $ZnSO_4.7H_2O$, $CoCl_2.6H_2O$, nitriloacetic acid, and a carbon source.

Only 26 carbon sources tested supported growth (summarized in **Table 2.4-2**), indicating that *G. xylinus* ATCC 53582 had a modest substrate utilization range. *G. xylinus* ATCC 53582 could use a variety of sugars and amino acids, but seemed to lack the ability to use organic acids (e.g., acetate, pyruvate, lactate, succinate). This is somewhat consistent with the literature information for this class of bacteria, which indicates as buildup of acidic products in the medium during

growth (37, 72), and indicates a loss of substrate carbon that is not going into biocellulose. A possible strategy would be to add organic acid utilization pathways into these organisms to increase biocellulose yield and reduce the negative effects of acid production on growth.

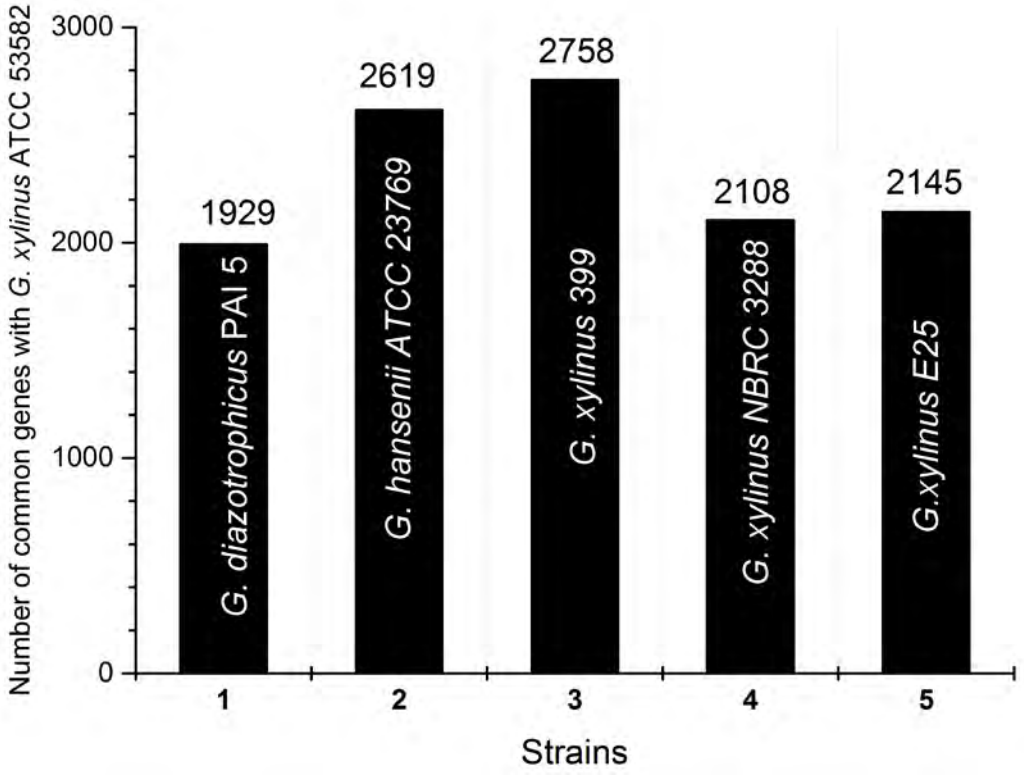



**Figure 2.4-1. Proteome comparison of five G*luconacetobacter* strains with *Gluconacetobacter xylinus* ATCC 53582.**

**Biolog experiments**

Biolog PM1 plates were used to screen the carbon sources that can be used by *G. xylinus* ATCC 53582 for growth, which gave insight into what genes and metabolic capabilities the strain possessed. The experiments were performed in media with different carbon sources supplemented with 5 g/L yeast extract. As a result, some growth was also observed in the negative control (where no additional carbon source was present). Carbon source utilization was determined by monitoring the optical density at 600 nm ($OD_{600}$). Any substrate resulting in a change in $OD_{600}$ that was 20% greater than that observed in the no carbon control was considered positive for supporting growth of *G. xylinus*. The ratio of growth on the substrate compared to the negative control was calculated as in Equation 1:

$$\text{Ratio} = \frac{[\text{Max OD}_{600} - \text{Min OD}_{600}]_{\text{carbon source}}}{[\text{Max OD}_{600} - \text{Min OD}_{600}]_{\text{negative control}}} \quad (1)$$

### 2.4.2.3 Reconstruction of the genome scale metabolic network of ATCC 53582

In order to identify strategies to couple biocellulose production with bacterial growth and enhance the yield of biocellulose, it was important to reconstruct the metabolic network of the two strains. The process for reconstructing a genome-scale metabolic network is shown in the flowchart below (**Figure 2.4-2**).

| Carbon source (Growth) | Average ratio of the replicates | Carbon source (No Growth) | Average ratio of the replicates |
|---|---|---|---|
| N-Acetyl-D-Glucosamine | 1.42 | D-Serine | 0.85 |
| Glycerol | 1.71 | L-Asparagine | 1.01 |
| m-Inositol | 1.72 | Glycyl-L-AsparticAcid | 0.95 |
| L-Fucose | 1.66 | Glycyl-LGlutamicAcid | 0.99 |
| D-Fructose-6-Phosphate | 1.23 | Glycyl-L-Proline | 1.09 |
| Tween40 | 1.37 | D-GalactonicAcid-g-Lactone | 0.47 |
| Tween80 | 1.74 | D-AsparticAcid | 0.92 |
| FumaricAcid | 1.87 | m-TartaricAcid | 0.95 |
| L-Alanine | 1.64 | CitricAcid | 0.84 |
| D-Galactose | 1.74 | p-HydroxyPhenylAceticAcid | 1.06 |
| a-Hydroxy Glutaric Acid-g-Lactone | 1.51 | D-Glucose-1-Phosphate | 1.01 |
| L-Alanyl-Glycine | 1.46 | Tyramine | 0.94 |
| L-AsparticAcid | 1.35 | SuccinicAcid | 0.90 |
| D-Fructose | 1.83 | D-Psicose | 0.72 |
| a-HydroxyButyricAcid | 1.44 | PropionicAcid | 0.50 |
| L-Proline | 1.60 | AcetoaceticAcid | 0.81 |
| D-Xylose | 1.74 | Glucuronamide | 0.48 |
| a-D-Glucose | 1.65 | GlycolicAcid | 1.01 |
| MonoMethylSuccinate | 1.22 | FormicAcid | 0.98 |
| L-GalactonicAcid-g-Lactone | 1.36 | GlyoxylicAcid | 0.74 |
| Lactulose | 1.96 | 2-DeoxyAdenosine | 0.72 |
| MethylPyruvate | 1.34 | D-MalicAcid | 0.99 |
| D-Mannose | 1.52 | Phenylethylamine | 0.50 |
| D-Melibiose | 1.86 | Dulcitol | 1.01 |
| Sucrose | 1.75 | Adenosine | 1.02 |
| Uridine | 1.26 | L-MalicAcid | 1.05 |
| | | 2-Aminoethanol | 0.84 |

**Table 2.4-2. Biolog growth experiments to screen carbon sources for growth.**

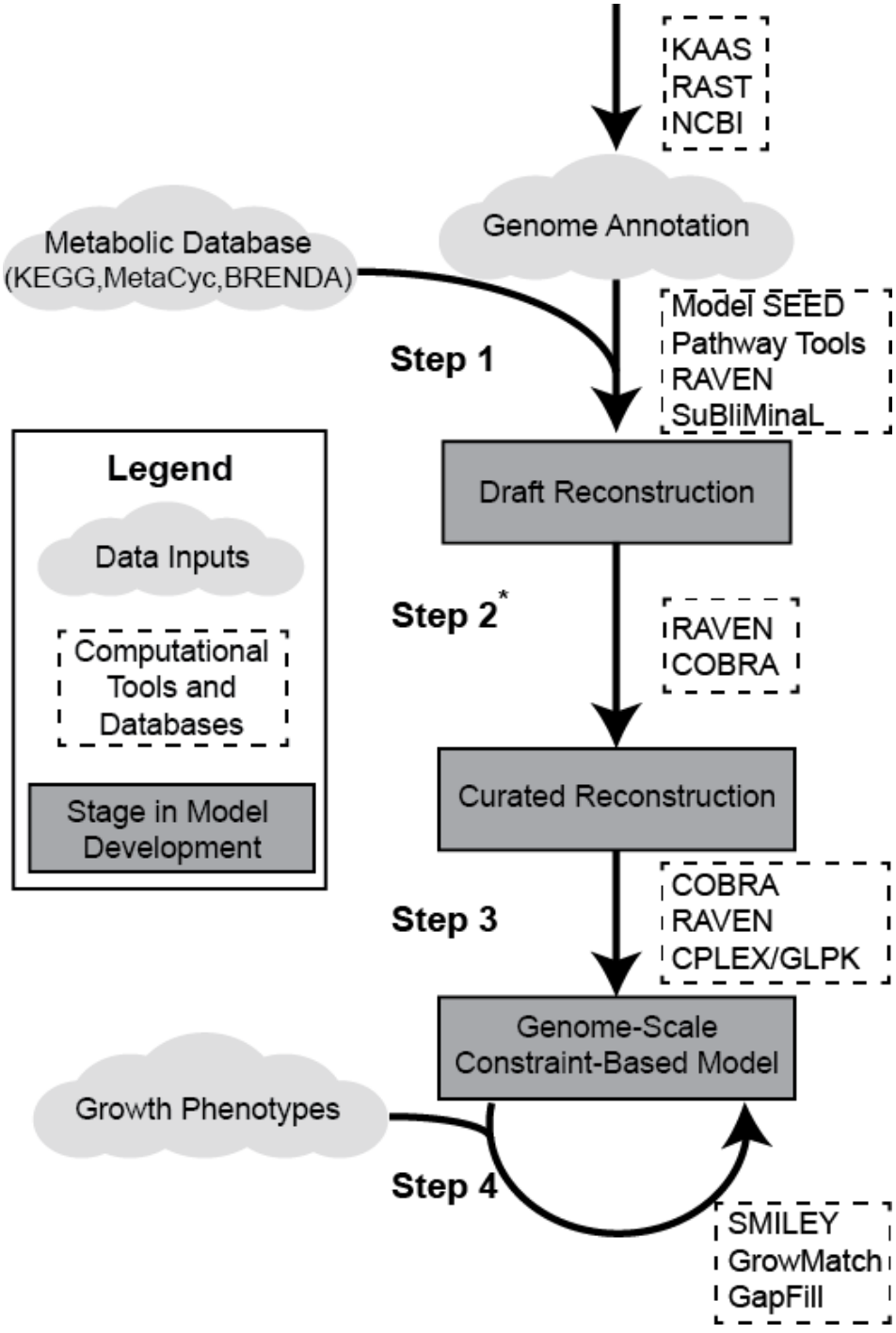



**Figure 2.4-2. Figure 2.4-2 Process of genome scale metabolic model reconstruction.** First step is to build a draft network and then gap-fill it based on an objective to reach to the final genome-scale metabolic network

The first step after genome annotation was the reconstruction of a draft metabolic network for *G. xylinus* ATCC 53582 which could be improved later to build the genome-scale metabolic network. The *Escherichia coli* metabolic model (iJO1366) was used as a template to build a draft metabolic network for *G. xylinus* ATCC 53582. Proteome comparison showed that *G. hansenii* ATCC 23769 was the closest previously sequenced strain to our *G. xylinus* ATCC 53582 (**Figure 2.4-1**). However, neither genome was available in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (30) sequence similarity database (SSDB).

Therefore, starting with the highly curated *E. coli* metabolic model (iJO1366), a metabolic model for strain *G. xylinus* E25 (which is present in KEGG SSDB) was generated, which was then used as a template to build a draft model for *G. xylinus* ATCC 53582. Briefly, the KEGG SSDB database was used to map *E. coli* metabolic genes to their orthologs in *G. xylinus* E25. Then the KBase proteome comparison was used to map these metabolic *G. xylinus* E25 genes to genes in *G. xylinus* ATCC 53582. This resulted in a draft metabolic model for *G. xylinus* ATCC 53582 with 742 reactions. This draft metabolic model had a large number of missing and isolated reactions as it was a subset of the *E. coli*'s metabolic model. Hence, in the next step, missing reactions were identified and added (a process called gap-filling), facilitated using the Biolog experiment data.

**Gap-filling the draft metabolic model of *G. xylinus* ATCC 53582.**

Using the Biolog experiments, reactions that were missing from the draft model - those needed to synthesize all biomass components (e.g., lipids, RNA, DNA, cofactors, and protein) from media components - were identified. A gap-filling algorithm was used to fill any missing reactions in the draft metabolic network of *G. xylinus* ATCC 53582 needed to support growth with glucose in BSM medium. The algorithm identified 314 reactions from the *G. xylinus* E25 draft metabolic model and 133 extra reactions from the *E. coli* iJO1366 model. To find genes that could catalyze these gap-filled reactions, NCBI-Blast+ was used to blast the genes of *G. xylinus* ATCC 53582 against *E. coli* K-12 MG1655 and *G. hansenii* ATCC 23769 genomes. *G. hansenii* ATCC 23769 was chosen for performing the blast process because it possessed the greatest overlap with *G. xylinus* ATCC 53582 based on the proteome comparison. Bit-scores, percent identities and e-values from the blast results, along with MetaCyc searches, were used to identify the responsible metabolic genes in *G. xylinus* ATCC 53582.

Out of the 314 reactions that were gap-filled using E25 draft model reactions, 298 were enzymatic reactions. Out of these 298 reactions, 273 were predicted to occur in *G. xylinus* ATCC 53582 based on proteome comparisons between *G. xylinus* E25 and *G. xylinus* ATCC 53582 using an e-value cut-off of 1e-30. Genes were also be found for 23 of the remaining 25 reactions by searching through the *G. xylinus* ATCC 53582 genome annotations. The two remaining reactions were predicted to be 'essential' for cell growth on glucose in BSM. The still unidentified genes were responsible for reactions, which were related to synthesizing molybdenum and bis-molybdopterin guanine dinucleotide, both of which are basic biomass components.

In addition to using reactions that occur in E25 model, the gap-filling algorithm also predicted that an additional 133 reactions from the *E. coli* model were needed to enable growth of *G. xylinus* ATCC 53582 in BSM glucose medium. About 25% of the reactions (35/133) were

responsible for transport of compounds between the extracellular and periplasmic space; 34 of were associated with one or more outer membrane porins and 1 was a symporter. The genome annotation for *G. xylinus* ATCC 53582 indicated there were 14 potential porins in *G. xylinus* ATCC 53582. There were 7 spontaneous reactions which did not need any gene to catalyze them. Genes could be found for 74 of the 91 remaining gap-filled *E. coli* reactions using NCBI-Blast+ with *E. coli and G. hansenii* ATCC 23769. Three genes of the remaining 18 gap-filled reactions genes were identified using the gene complementation experiments with *E. coli* (listed in **Table 2.4-3**); however genes for 15 reactions remained identified.

As mentioned above, the gap-filling algorithm was run with the objective of identifying the minimum number of reactions that could enable cell growth. Even after completing the gap-filling process, there a significant number of important metabolic genes which might not have been associated to any reaction because they were not required for cell growth per se. Adding these additional reactions might better explain more capabilities of the organism, and may also be important because the network after gap-filling only had 494 metabolic genes out of a total of 3,731 protein coding genes indicated by the sequencing data.

To identify the reactions catalyzed by additional genes, the gene annotation and EC numbers were cross-referenced with the BiGG database (http://bigg.ucsd.edu/) and KEGG's reaction database. After these additions, there were 1077 reactions in the metabolic model of *G. xylinus* ATCC53582.

| Reaction | | Gene | *E. coli* gene |
|---|---|---|---|
| asp-L[c] + h[c] -> ala-B[c] + co2[c] | Aspartate 1-decarboxylase | CDS.6673 | PanD |
| 4per[c] + nad[c] <=> h[c] + nadh[c] + ohpb[c] | Erythronate 4-phosphate (4per) dehydrogenase | CDS.6145 | PdxB |
| atp[c] + skm[c] -> adp[c] + h[c] + skm5p[c] | Shikimate kinase | CDS.5873 | AroK |

**Table 2.4-3. Gene identification using complementation in *E. coli.***

### 2.4.2.4 Reconstruction of the genome scale metabolic network of *G. xylinus* 399

The genome-scale metabolic network of *G. xylinus* 399 was reconstructed from the metabolic network of *G. xylinus* ATCC 53582. Proteome comparison in **Figure 2.4-1** shows that there were many orthologs between *G. xylinus* ATCC 53582 and Gx 399. The *G. xylinus* ATCC 53582 genes were mapped to the *G. xylinus* 399 genes using the proteome comparison feature in KBase. Genes were found for all but 13 the reactions in *G. xylinus* ATCC 53582, which was reduced to just 5 unidentified reactions by manual genomic analysis. **Table 2.4-4** lists the numbers of reactions in different metabolic subsystems for the *G. xylinus* ATCC 53582 and Gx 399 models. The genome-scale metabolic model of *G. xylinus* 399 had five fewer reactions that

the *G. xylinus* ATCC 53582 model (listed in **Table 2.4-5**).

| Subsystem | # of reactions | Subsystem | # of reactions |
|---|---|---|---|
| Exchange | 110/110 | Glycolysis/Gluconeogenesis | 15/15 |
| Cofactor and Prosthetic Group Biosynthesis | 161/160 | Membrane Lipid Metabolism | 9/9 |
| Purine and Pyrimidine Biosynthesis | 23/23 | Lipopolysaccharide Biosynthesis / Recycling | 18/18 |
| Folate Metabolism | 10/10 | Glycine and Serine Metabolism | 7/6 |
| tRNA Charging | 18/18 | Cysteine Metabolism | 11/11 |
| Transport, Inner Membrane | 78/78 | Methionine Metabolism | 13/13 |
| Arginine and Proline Metabolism | 24/24 | Methylglyoxal Metabolism | 2/2 |
| Glycerophospholipid Metabolism | 58/58 | Pyruvate Metabolism | 6/6 |
| Cell Envelope Biosynthesis | 85/85 | Transport, Outer Membrane Porin | 47/47 |
| Inorganic Ion Transport and Metabolism | 33/32 | Anaplerotic Reactions | 6/6 |
| Unassigned | 8/8 | Citric Acid Cycle | 9/9 |
| Murein Biosynthesis | 10/10 | Nucleotide Salvage Pathway | 57/57 |
| Alternate Carbon Metabolism | 50/49 | Alanine and Aspartate Metabolism | 8/8 |
| Threonine and Lysine Metabolism | 15/15 | Glutamate Metabolism | 2/1 |
| Valine, Leucine, and Isoleucine Metabolism | 15/15 | Murein Recycling | 11/11 |
| Tyrosine, Tryptophan, and Phenylalanine Metabolism | 19/19 | Transport, Outer Membrane | 10/10 |
| Pentose Phosphate Pathway | 11/11 | Histidine Metabolism | 10/10 |
| Oxidative Phosphorylation | 13/13 | Glutamate Metabolism | 3/3 |
| Nitrogen Metabolism | 2/2 | Glyoxylate Metabolism | 1/1 |
| Additional Unassigned | 66/66 | | |

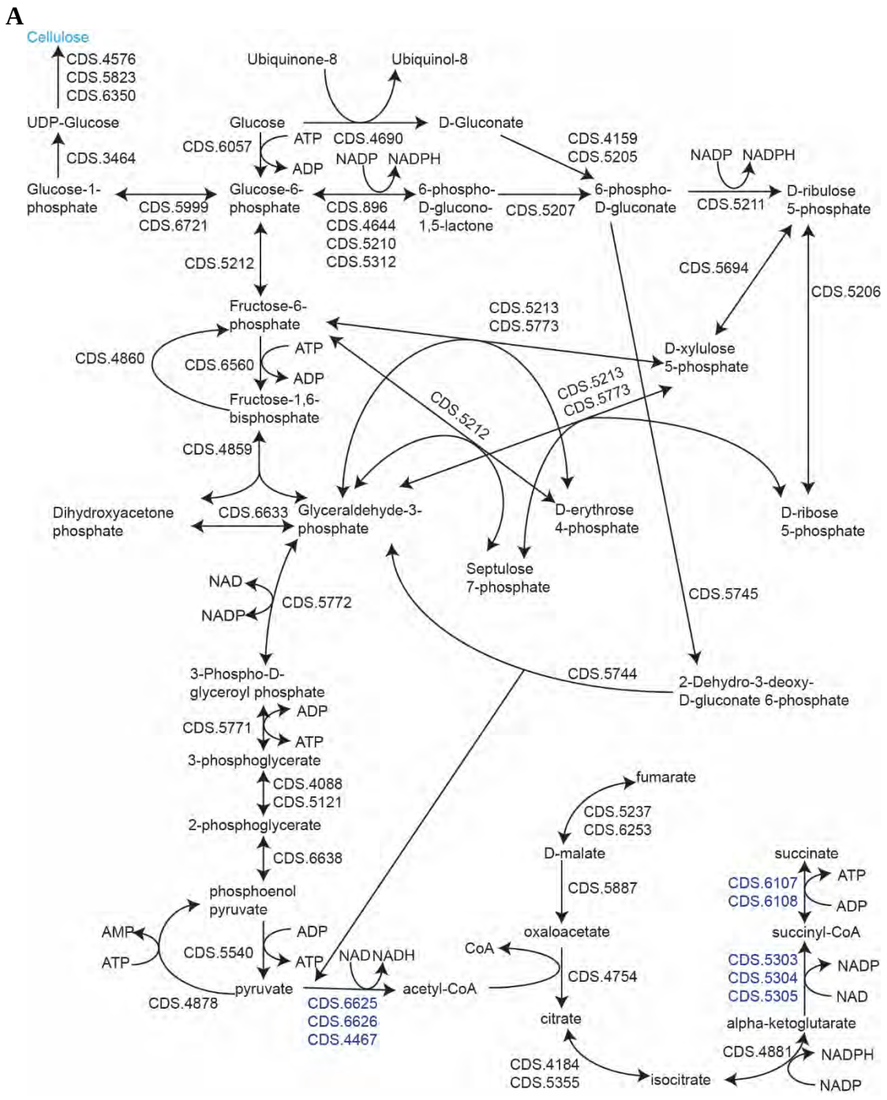
**Table 2.4-4. Breakdown of reactions by subsystems for *G. xylinus* ATCC53582/Gx 399.**

| Reaction Name | Metabolic Reaction |
|---|---|
| Arsenate reductase | Arsenate+ 2 Reduced Glutathione -> Arsenite + Oxidized Glutathione + $H_2O$ |
| Glutamate dehydrogenase | L-glutamate + $H_2O$ + NADP <=> 2-oxoglutarate + $H^+$ + NADPH + $NH_4^+$ |
| Sarcosine oxidase | $H_2O$ + $O_2$ + Sarcosine -> Formaldehyde + Glycine + $H_2O_2$ |
| Flavin reductase (NAD) | $H^+$ + NADH + Riboflavin -> NAD + Reduced Riboflavin |
| Phenylacetate-CoA ligase | ATP + coa + Phenylacetic acid -> AMP + Phenylacetyl-CoA + PPi |

**Table 2.4-5. Missing reactions in *G. xylinus* 399 which are present in *G. xylinus* ATCC 53582.**

### 2.4.2.5 Metabolic pathway and genetic differences between *G. xylinus* ATCC 53582 and Gx 399

From the **Figure 2.4-3**, it can be seen that both the strains had an incomplete TCA cycle because of the absence of fumarate reductase and succinate dehydrogenase. Even though both strains also had a complete glycolytic pathway, reports suggest that *Gluconacetobacter* strains have a weak phosphofructokinase activity and instead use the Entner-Doudoroff pathway (72, 73).



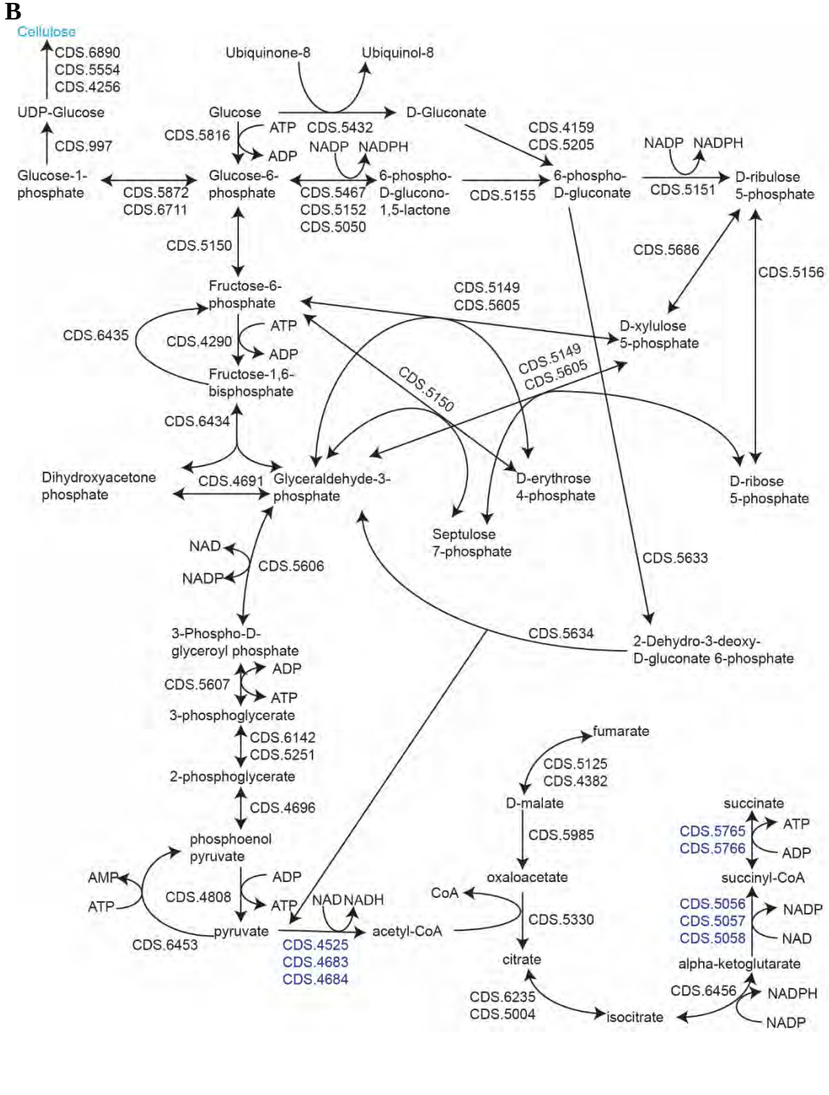




**Figure 2.4-3. Core metabolism of the strains A) *G. xylinus* ATCC 53582 and B) Gx 399.**

### 2.4.2.6 Strain design to enhance biocellulose production from *G. xylinus* ATCC 53582

SimOptStrain (33) was used to identify gene deletions and reaction addition strategies to enhance the production of biocellulose from *G. xylinus* ATCC 53582 and Gx 399 when the corresponding strains are growing at their maximal growth rate. All simulations were done assuming growth in BSM with glucose only (no added YE). SimOptStrain only found solutions when a minimum of four or five gene deletions were made. The resulting strain designs corresponding to 4 and 5 deletions are summarized in **Tables 2.4-6** and **2.4-8**, respectively.

It is important to note that the metabolic model does not incorporate regulatory interactions which direct substrate utilization to either biomass production or biocellulose synthesis. Therefore, the models cannot at this point predict biocellulose production in wildtype for comparison to the various designs. Therefore, comparisons were made to empirical data from BC production experiments.

**Strain design solutions with five gene deletions**

The deletions are mainly focused on reactions involved in the isomerization of glucose-6-phosphate into glucose-1-phosphate, and preventing the resulting metabolites consumption by the non-biocellulose arm of the pentose phosphate pathway (which does not lead to biocellulose precursors). Flux into the pentose phosphate pathway is redirected through D-gluconate made from D-glucose (by glucose dehydrogenase, which uses ubiquinone-8 as an electron acceptor). As shown in **Table 2.4-6,** the model predicted growth rates decrease from 0.95 $hr^{-1}$ in the wildtype strain to 0.11 $hr^{-1}$ in the proposed mutant as a result of increased flux through the biocellulose synthesis reaction. The biocellulose yield for the wildtype was 0.46 ± 0.24 g BC / g glucose consumed (n = 16), compared to 0.70 for the mutant. While these estimated increases in yield are encouraging, experiments directly comparing the mutant and the wildtype would be needed to validate them.

To make sure that the cell used this glucose dehydrogenase reaction, SimOptStrain suggested deletions of other reactions which can reduce ubiquinone. Ubiquinol is required for the reduction of $O_2$ to $H_2O$ in the cytochrome oxidase reaction (an essential reaction) shown below.

> Cytochrome oxidase bo3 (ubiquinol-8: 4 protons) (periplasm):
> $4\ H^+ + 0.5\ O_2$ + ubiquinol-8 -> $H_2O$ + ubiquinone-8 + $4\ H^+[p]$

The genes listed were associated with more than five reactions because some of the enzymes catalyze multiple reactions (associated reactions are listed in **Table 2.4-7**). The reactions which were required for coupling biocellulose production with growth are: (i) NADH dehydrogenase (ubiquinone-8 & 3 protons) (periplasm) (NADH16pp); (ii) D-lactate dehydrogenase; (iii) 6-phosphogluconolactonase; (iv) D-fructose-6-phosphate D-erythrose-4-phosphate-lyase, and; (v) Glycolate dehydrogenase (NAD).

The genes listed were associated with more than five reactions because some of the enzymes catalyze multiple reactions (associated reactions are listed in **Table 2.4-7**). The reactions which are required for coupling biocellulose production with growth are: (i) NADH dehydrogenase (ubiquinone-8 & 3 protons) (periplasm) (NADH16pp); (ii) D-lactate dehydrogenase; (iii) 6-

phosphogluconolactonase; (iv) D-fructose-6-phosphate D-erythrose-4-phosphate-lyase, and; (v) Glycolate dehydrogenase (NAD).

| | | | | | Growth Rate ($hr^{-1}$) | Biocellulose Yield (g BC / g glucose consumed) |
|---|---|---|---|---|---|---|
| WT | | | | | 0.95 | 0.46 ± 0.24 (n = 16) |
| ΔGene1 | ΔGene2 | ΔGene3 | ΔGene4 | ΔGene5 | | |
| CDS.4235 | CDS.5783 | CDS.4853 | CDS.5552 | CDS.5207 | 0.11 | 0.70 |
| CDS.4236 | CDS.5783 | CDS.4853 | CDS.5552 | CDS.5207 | | |
| CDS.4237 | CDS.5783 | CDS.4853 | CDS.5552 | CDS.5207 | | |
| CDS.4238 | CDS.5783 | CDS.4853 | CDS.5552 | CDS.5207 | | |
| CDS.4240 | CDS.5783 | CDS.4853 | CDS.5552 | CDS.5207 | | |

**Table 2.4-6. Strain design solutions with five gene deletions in *G. xylinus* ATCC 53582.** Shown are the SimOptStrain suggested gene deletions along with the resulting cellular growth rate and the flux through the cellulose synthesis reaction. Any of the five genes in the first column can be deleted, as they are all subunits of NADH dehydrogenase. Each row is a different solution, and all solutions have the same predicted growth and biocellulose production rate.

| Gene | Reaction | Formula |
|---|---|---|
| CDS.4238<br>CDS.4237<br>CDS.4236<br>CDS.4235<br>CDS.4240 | NADH dehydrogenase (demethylmenaquinone-8 & 3 protons) (periplasm) (NADH18pp) | 2-Demethylmenaquinone-8 + 4 $H^+$ + NADH -> 2-Demethylmenaquinol-8 + NAD + 3 $H^+$ [p] |
| CDS.4238<br>CDS.4237<br>CDS.4236<br>CDS.4235<br>CDS.4240 | NADH dehydrogenase (menaquinone-8 & 3 protons) (periplasm) (NADH17pp) | 4 $H^+$ + Menaquinone-8 + NADH -> Menaquinol-8 + NAD + 3 $H^+$ [p] |
| CDS.4238<br>CDS.4237<br>CDS.4236<br>CDS.4235<br>CDS.4240 | NADH dehydrogenase (ubiquinone-8 & 3 protons) (periplasm) (NADH16pp) | 4 $H^+$ + NADH + Ubiquinone-8 -> NAD + Ubiquinol-8 + 3 $H^+$ [p] |
| CDS.4853 | D-lactate dehydrogenase (LDH_D2) | D-lactate + Ubiquinone-8 -> Pyruvate + Ubiquinol-8 |
| CDS.5207 | 6-phosphogluconolactonase (PGL) | 6-Phospho-D-glucono-1,5-lactone+ $H_2O$ -> 6-Phospho-D-gluconate + $H^+$ |
| CDS.5552 | D-xylulose 5-phosphate D-glyceraldehyde-3-phosphate-lyase (R01621) | D-Xylulose 5-phosphate + Orthophosphate <=> Acetyl phosphate + D-Glyceraldehyde 3-phosphate + $H_2O$ |
| CDS.5552 | D-fructose-6-phosphate D-erythrose-4-phosphate-lyase (R00761) | D-Fructose 6-phosphate + Orthophosphate <=> Acetyl phosphate + D-Erythrose 4-phosphate + $H_2O$ |
| CDS.5783 | Glycolate dehydrogenase (NAD) (GLYCLTDx) | Glyoxylate + $H^+$ + NADH -> Glycolate + NAD |

**Table 2.4-7. Corresponding reactions of the suggested five gene knockouts in *G. xylinus* ATCC 53582.**

Shown are the reactions which are predicted to result in no flux as a result of the five gene deletions.

**Figure 2.4-4** shows the flux maps for the wildtype *G. xylinus* ATCC 53582 and the proposed mutant with five gene deletions. In the wildtype, glucose was predicted to be metabolized by the glycolysis pathway, with no flux through the phosphoglucomutase reaction (glucose-6-phosphate to glucose-1-phosphate), which results in no flux through the biocellulose synthesis reaction. This means that the model predicts no coupling between biocellulose synthesis and growth. However, a link between growth and biocellulose synthesis was demonstrated empirically during Task 2.2, indicating that the wildtype model is still not a completely accurate representation of the wildtype strain.

In contrast, in the five deletion mutant, glucose was converted to gluconate (also reducing ubiquinone), which entered the pentose phosphate cycle and produced fructose-6-phosphate and glyceraldehyde-3-phosphate (through transketolase, transaldolase and 2-dehydro-3-deoxy-phosphogluconate aldolase reactions). Glyceraldehyde-3-phosphate was consumed eventually through the remaining reactions of glycolysis to make pyruvate. Fructose-6-phosphate was converted into glucose-6-phosphate which goes into the biocellulose synthesis pathway to make biocellulose.

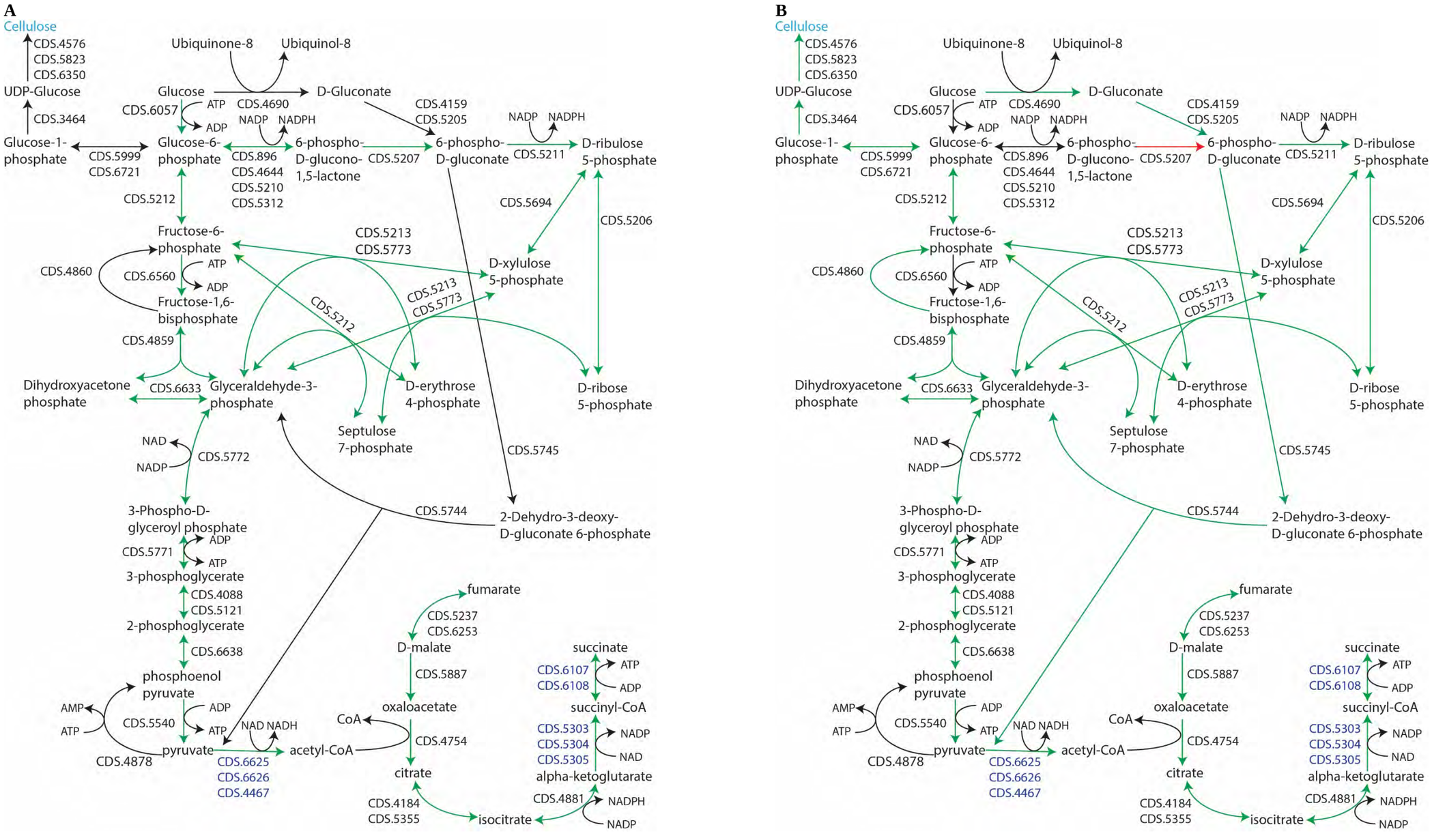



**Figure 2.4-4. Flux map of A) *G. xylinus* ATCC 53582 and B) the mutant strain with five gene deletions.**
Reactions with green arrows have flux, black arrows have no flux and red arrows are deleted due to the deleted genes. The enzymes marked with blue text are subunits.

**Strain design solutions with four gene deletions**

SimOptStrain suggested solutions involving four gene deletions which are listed in **Table 2.4-8**. The reaction deletions required for coupling biocellulose production with growth are (i) phosphoglycerate kinase, (ii) D-Glucose: D-fructose oxidoreductase, (iii) 6-phosphogluconolactonase, and (iv) D-fructose-6-phosphate D-erythrose-4-phosphate-lyase (**Table 2.4-9**). In this case, the growth rate decreased from 0.95 $hr^{-1}$ in the wildtype to 0.41 $hr^{-1}$ in the mutant, while the yield of biocellulose in the wildtype was 0.46 ± 0.24 (n = 16) compared to an estimated 0.41 g BC / g glucose consumed in the mutant. A final assessment of the effectiveness of these modifications with respect to increasing BC yields would require actual experiments with the wildtype and mutant, due to the large uncertainty in the measured wildtype yield.

| | | | | Growth Rate ($hr^{-1}$) | Biocellulose Yield (g BC / g glucose consumed) |
|---|---|---|---|---|---|
| WT | | | | 0.95 | 0.46 ± 0.24 (n = 16) |
| ΔGene1 | ΔGene2 | ΔGene3 | ΔGene4 | | |
| CDS.6862 | CDS.5552 | CDS.5771 | CDS.5207 | 0.41 | 0.41 |
| CDS.6254 | CDS.5552 | CDS.5771 | CDS.5207 | | |
| CDS.3717 | CDS.5552 | CDS.5771 | CDS.5207 | | |
| CDS.5370 | CDS.5552 | CDS.5771 | CDS.5207 | | |
| CDS.5148 | CDS.5552 | CDS.5771 | CDS.5207 | | |

**Table 2.4-8. Strain design solutions with four gene deletions in *G. xylinus* ATCC 53582.**

| Gene | Reaction | Formula |
|---|---|---|
| CDS.5771 | phosphoglycerate kinase (PGK) | 3-Phospho-D-glycerate+ ATP <=> 3-Phospho-D-glyceroyl phosphate + ADP |
| CDS.5207 | 6-phosphogluconolactonase (PGL) | 6-phospho-D-glucono-1,5-lactone + $H_2O$ → 6-Phospho-D-gluconate + $H^+$ |
| CDS.5552 | D-xylulose 5-phosphate D-glyceraldehyde-3-phosphate-lyase (R01621) | D-Xylulose 5-phosphate + Orthophosphate <=> Acetyl phosphate + D-Glyceraldehyde 3-phosphate + $H_2O$ |
| CDS.5552 | D-fructose-6-phosphate D-erythrose-4-phosphate-lyase (R00761) | D-Fructose 6-phosphate + Orthophosphate <=> Acetyl phosphate + D-Erythrose 4-phosphate + $H_2O$ |
| CDS.6862 | D-Glucose:D-fructose oxidoreductase (R00874) | D-Fructose + D-Glucose → Glucono-1,5-lactone+ D-Sorbitol |
| CDS.6254 | D-Glucono-1,5-lactone lactonohydrolase | D-Glucono-1,5-lactone + $H_2O$ <=> D-Gluconate + $H^+$ |
| CDS.3717 | | L-Iditol + NAD+ <=> L-Sorbose + NADH + $H^+$ |
| CDS.3717 | D-Glucitol:NAD+ 2-oxidoreductase | D-Fructose+ NADH+ $H^+$ <=> D-Sorbitol+NAD |
| CDS.3717 | | Xylitol+NAD <=> D-Xylulose + NADH |
| CDS.5370 | Alpha, alpha-trehalase | $H_2O$ + Trehalose → 2 D-Glucose |
| CDS.5370 | Alpha, alpha-trehalase (periplasm) | $H_2O$[p] + Trehalose[p] → 2 D-Glucose[p] |
| CDS.5148 | Trehalose-phosphatase | $H_2O$ + Trehalose-6-phosphate → pi + Trehalose |

**Table 2.4-9. Corresponding reactions of the suggested four gene knockouts in *G. xylinus* ATCC 53582.**
Shown are the reactions which will have no flux as a result of the five gene deletions.

There were five different solutions and three gene knock outs (phosphoglycerate kinase (CDS.5771), D-fructose-6-phosphate D-erythrose-4-phosphate-lyase (CDS.5552) and 6-phosphogluconolactonase (CDS.5207)) that were common across the five solutions. These three common gene knock outs suggest again directing glucose-6-phosphate for the synthesis of biocellulose and not for other cellular processes via the pentose phosphate pathway.

Similar to the five deletion solution, glucose was converted into gluconate which enters the pentose phosphate pathway and into the Entner-Doudoroff pathway to make glyceraldehyde-3-phosphate and pyruvate. Because phosphoglycerate kinase is knocked out, glyceraldehyde-3-phosphate followed the gluconeogenesis pathway to make glucose-6-phosphate which could only be utilized for biocellulose synthesis pathway due to the removal of 6-phosphogluconolactonase (**Figure 2.4-5**).

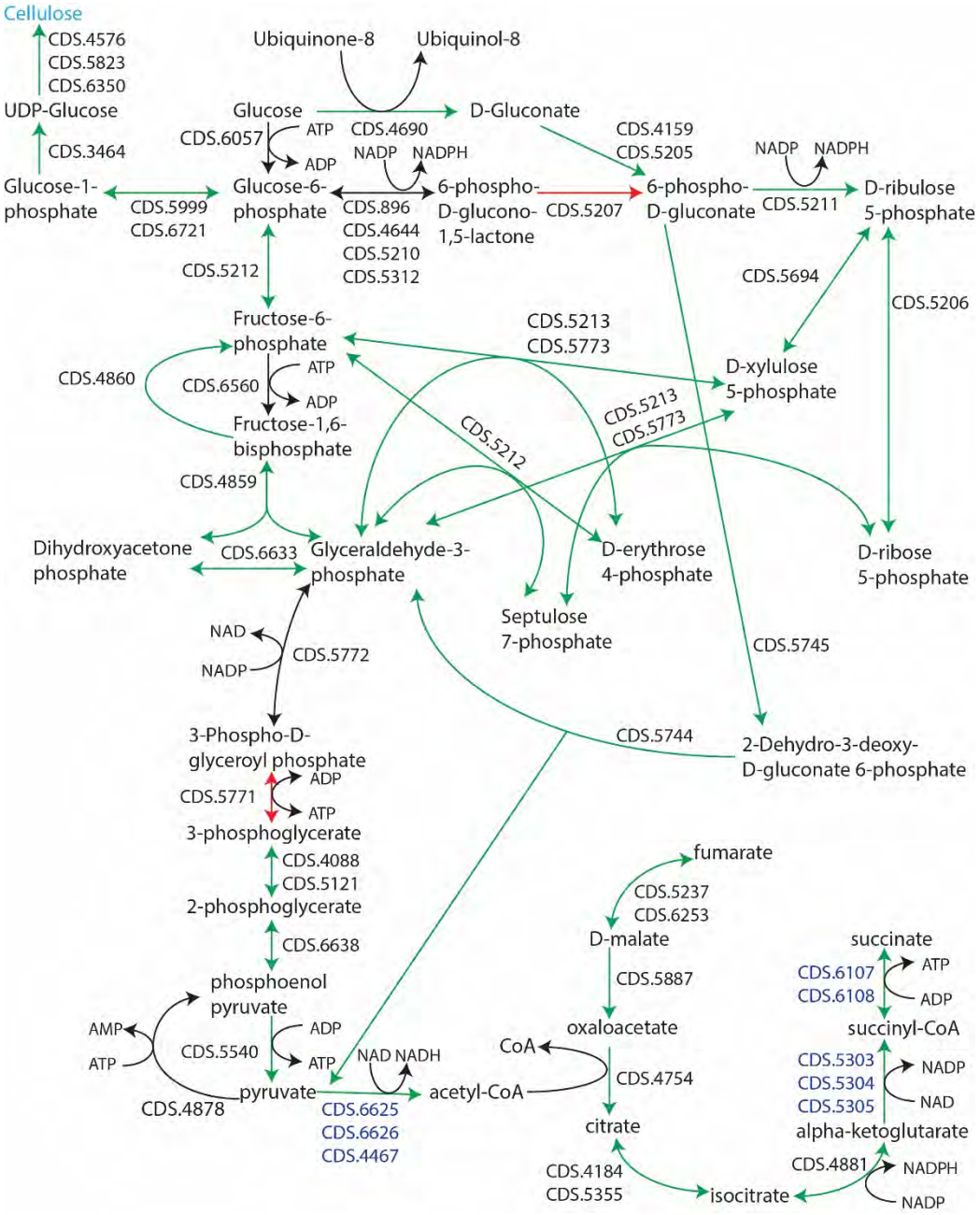



**Figure 2.4-5. Flux map of the mutant strain after the suggest four deletions in *G. xylinus* ATCC 53582.**

### 2.4.2.7 Strain design to enhance biocellulose production from *G. xylinus* 399

As with *G. xylinus* ATCC 53582, SimOptStrain found solutions for stain improvement of Gx 399 only with four or more gene deletions.

**Strain design solutions with five gene deletions**

The SimOptStrain strain designs with five gene deletions are listed in **Table 2.4-10** and the reactions associated with these genes are listed in **Table 2.4-11**. The predicted growth rate in the mutants decreases from 0.92 $hr^{-1}$ in the wildtype to 0.27 $hr^{-1}$ in the mutant, and the biocellulose yield goes up from 0.32 ± 0.16 (n = 16) to 0.58 (g biocellulose)/(g glucose). Actual experiments comparing the mutant and the wildtype would be needed to assess the effectiveness of these modifications due to the large uncertainty in the measured wildtype yield. The predicted flux maps for both *G. xylinus* 399 and the proposed five deletion mutants are shown in **Figure 2.4-6**.

Like the *G. xylinus* ATCC 53582 mutants, deletion of CDS.5155 (6-phosphogluconolactonase) meant that glucose-6-phosphate could only be utilized to produce biocellulose. Deletion of CDS.4796 (D-fructose-6-phosphate D-erythrose-4-phosphate-lyase) prevented the conversion of fructose-6-phosphate to erythronate-4-phosphate, so that the fructose-6-phosphate could be only used to produce glucose-6-phosphate and eventually biocellulose.

Deletion of CDS.5634 (2-dehydro-3-deoxy-phosphogluconate aldolase) or CDS.5633 (6-phosphogluconate dehydratase) ensured that there was production of both glyceraldehyde-3-phosphate and fructose-6-phosphate. Deleting CDS.4290 (phosphofructokinase) or CDS.6434 (fructose-bisphosphate aldolase) ensured that fructose-6-phosphate was not utilized by glycolysis and could only be used to produce biocellulose.

| | | | | | Growth Rate ($hr^{-1}$) | Biocellulose Yield (g BC / g glucose consumed) |
|---|---|---|---|---|---|---|
| WT | | | | | 0.92 | 0.32 ± 0.16 (n = 16) |
| ΔGene1 | ΔGene2 | ΔGene3 | ΔGene4 | ΔGene5 | | |
| CDS.5155 | CDS.4796 | CDS.5634 | CDS.6434 | CDS.6851 | 0.27 | 0.58 |
| CDS.5155 | CDS.4796 | CDS.5633 | CDS.4290 | CDS.6851 | | |
| CDS.5155 | CDS.4796 | CDS.5634 | CDS.6434 | CDS.4381 | | |
| CDS.5155 | CDS.4796 | CDS.5634 | CDS.4290 | CDS.4381 | | |
| CDS.5155 | CDS.4796 | CDS.5633 | CDS.4290 | CDS.4381 | | |

**Table 2.4-10. Strain design solutions with five gene deletions in Gx 399.**
Shown are the solutions with five gene deletions and the corresponding predictions for cell growth rate and flux through the biocellulose synthesis reaction. Each row is a different solution and all solutions had the same predicted growth and biocellulose production rate.

| Gene | Reaction | Formula |
|---|---|---|
| CDS.5155 | 6-phosphogluconolactonase (PGL) | 6-phospho-D-glucono-1,5-lactone + $H_2O$ → 6-Phospho-D-gluconate + $H^+$ |
| CDS.4796 | D-xylulose 5-phosphate D-glyceraldehyde-3-phosphate-lyase (R01621) | D-Xylulose 5-phosphate + Orthophosphate <=> Acetyl phosphate + D-Glyceraldehyde 3-phosphate + $H_2O$ |
| CDS.4796 | D-fructose-6-phosphate D-erythrose-4-phosphate-lyase (R00761) | D-Fructose 6-phosphate + Orthophosphate <=> Acetyl phosphate + D-Erythrose 4-phosphate + $H_2O$ |
| CDS.5634 | 2-dehydro-3-deoxy-phosphogluconate aldolase (EDA) | 2-Dehydro-3-deoxy-D-gluconate 6-phosphate → Glyceraldehyde 3-phosphate + Pyruvate |
| CDS.5634 | Oxaloacetate decarboxylase (OAADC) | $H^+$ + Oxaloacetate → $CO_2$ + Pyruvate |
| CDS.5633 | 6-phosphogluconate dehydratase (EDD) | 6-Phospho-D-gluconate → 2-Dehydro-3-deoxy-D-gluconate 6-phosphate + $H_2O$ |
| CDS.6434 | Fructose-bisphosphate aldolase (FBA) | D-Fructose 1,6-bisphosphate <=> Dihydroxyacetone phosphate + Glyceraldehyde 3-phosphate |
| CDS.6434 | Sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase (FBA3) | Sedoheptulose 1,7-bisphosphate <=> Dihydroxyacetone phosphate+ D-Erythrose 4-phosphate |
| CDS.4290 | Phosphofructokinase | ATP + D-Fructose 6-phosphate → ADP + D-Fructose 1,6-bisphosphate + $H^+$ |
| CDS.6851 | D-Glucose:D-fructose oxidoreductase (R00874) | D-Fructose+D-glucose → glucono-1,5-lactone + D-Sorbitol |
| CDS.4381 | D-Glucono-1,5-lactone lactonohydrolase (R01519) | D-Glucono-1,5-lactone+ $H_2O$ → D-Gluconate + $H^+$ |

**Table 2.4-11. Corresponding reactions of the suggested five gene knockouts in Gx 399.** Shown are the reactions which will have no flux as a result of the five gene deletions.

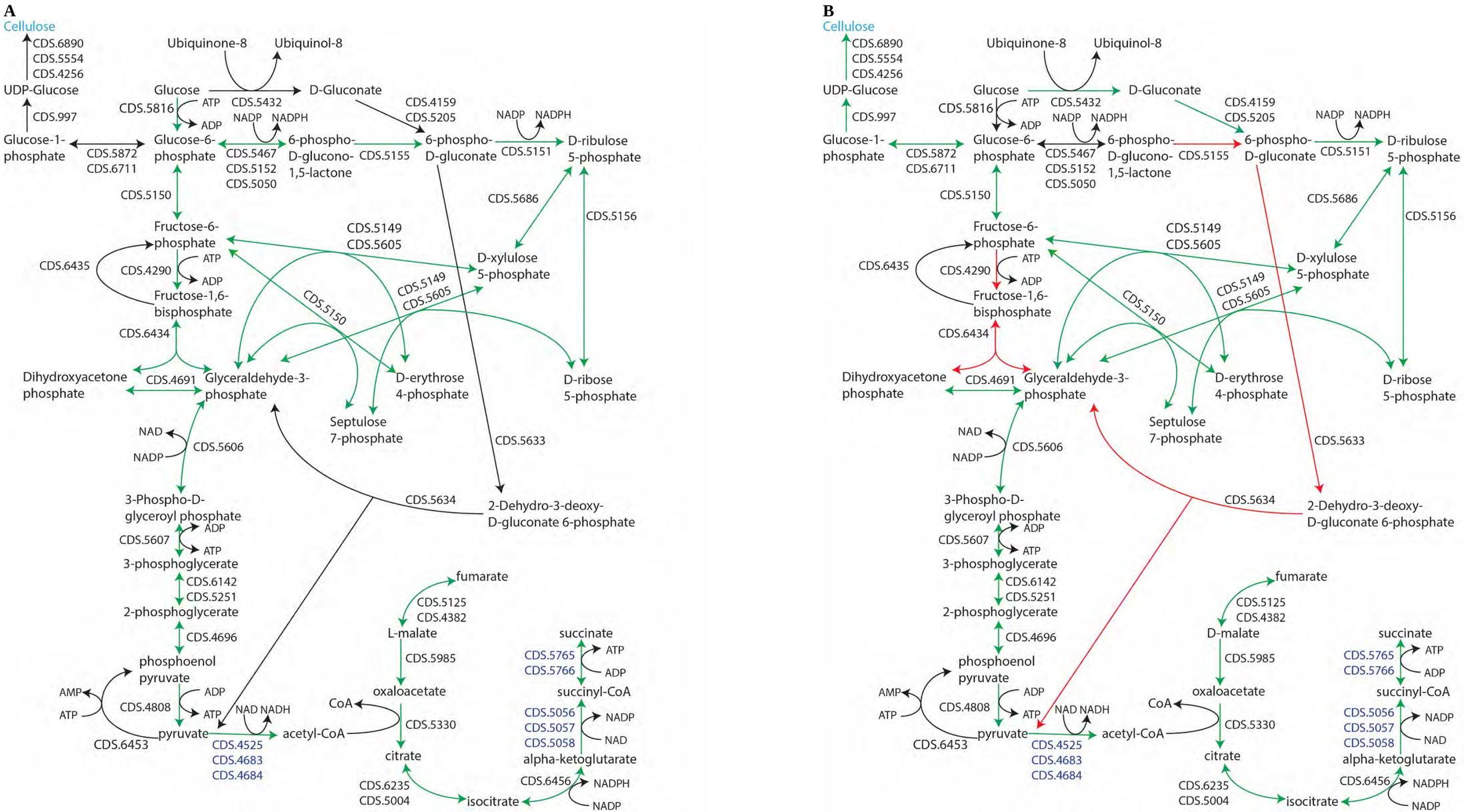



**Figure 2.4-6. Flux map for A) Gx 399 and B) the mutant with five gene deletions.**
Reactions with green arrows have flux, black arrows have no flux and red arrows should be deleted based on the deleted genes.

**Strain design solutions with four gene deletions**

SimOptStrain found solutions with four gene deletions, but the predicted biocellulose yields were lower than in mutants with five gene deletions (**Table 2.4-12**). SimOptStrain suggests deleting CDS.5155 (6-phosphogluconolactonase), CDS.4796 (D-fructose-6-phosphate D-erythrose-4-phosphate-lyase), CDS.4381 (D-Glucono-1,5-lactone lactonohydrolase), and CDS.6851 (D-Glucose: D-fructose oxidoreductase) like in the five deletions solutions. However, these four deletion mutants use the ED pathway because of the deletion of CDS.5607 (phosphoglycerate kinase) since glyceraldehyde-3-phosphate cannot be consumed to produce pyruvate.

The predicted growth rate in the four deletion mutants is 0.40 $hr^{-1}$ as compared to the 0.92 $hr^{-1}$ in wildtype, and the yield of biocellulose production in the wildtype was 0.32 ± 0.16 (n = 16) consumed, compared to 0.41 g BC / g glucose predicted for the mutant. As stated in previous sections, side-by-side comparisons of the wildtype and the mutant would be needed to validate these estimates, especially given the large variability in the available wildtype data. **Figure 2.4-7** shows the predicted flux map for the four gene deletion mutants.

| | | | | Growth Rate ($hr^{-1}$) | Biocellulose Yield (g BC / g glucose consumed) |
|---|---|---|---|---|---|
| WT | | | | 0.92 | 0.32 ± 0.16 (n = 16) |
| ΔGene1 | ΔGene2 | ΔGene3 | ΔGene4 | | |
| CDS.5155 | CDS.4796 | CDS.5607 | CDS.4381 | 0.40 | 0.41 |
| CDS.5155 | CDS.4796 | CDS.5607 | CDS.6851 | | |
| CDS.5155 | CDS.4796 | CDS.5607 | CDS.4121 | | |
| CDS.5155 | CDS.4796 | CDS.5607 | CDS.4989 | | |
| CDS.5155 | CDS.4796 | CDS.5607 | CDS.5224 | | |

**Table 2.4-12. Strain design solutions with four gene deletions in Gx 399.**

Shown are the solutions with four gene deletions and the corresponding predictions for cell growth rate and flux through the biocellulose synthesis reaction. Each row is a different solution and all solutions had the same predicted growth and biocellulose production rate.

| Gene | Reaction | Formula |
|---|---|---|
| CDS.5155 | 6-phosphogluconolactonase (PGL) | 6-phospho-D-glucono-1,5-lactone + $H_2O$ → 6-phospho-D-glucono-1,5-lactone+ $H^+$ |
| CDS.4796 | D-xylulose 5-phosphate D-glyceraldehyde-3-phosphate-lyase (R01621) | D-Xylulose 5-phosphate + Orthophosphate <=> Acetyl phosphate + D-Glyceraldehyde 3-phosphate + $H_2O$ |
| CDS.4796 | D-fructose-6-phosphate D-erythrose-4-phosphate-lyase (R00761) | D-Fructose 6-phosphate + Orthophosphate <=> Acetyl phosphate + D-Erythrose 4-phosphate + $H_2O$ |
| CDS.5607 | Phosphoglycerate kinase (PGK) | 3-Phospho-D-glycerate + ATP <=> 3-Phospho-D-glyceroyl phosphate + ADP |
| CDS.4381 | D-Glucono-1,5-lactone lactonohydrolase (R01519) | D-Glucono-1,5-lactone+$H_2O$ → D-Gluconate + $H^+$ |
| CDS.6851 | D-Glucose: D-fructose oxidoreductase (R00874) | D-Fructose + D-Glucose → Glucono-1,5-lactone + D-Sorbitol |
| CDS.4121 | L-iditol: NAD+ 2-oxidoreductase (R07145) | L-Sorbose + NADH + $H^+$ <=> L-Iditol + NAD |
| CDS.4121 | D-Glucitol: NAD+ 2-oxidoreductase (R00875) | D-Fructose + NADH + $H^+$ <=> D-Sorbitol + NAD |
| CDS.4121 | Xylitol: NAD+ 2-oxidoreductase (R01896) | Xylitol + NAD <=> xylu-D+ NADH |
| CDS.4989 | Alpha, alpha-trehalase (TREH) | $H_2O$ + Trehalose → 2 D-Glucose |
| CDS.4989 | Alpha, alpha-trehalase (periplasm) (TREHpp) | $H_2O$[p] + Trehalose[p] → 2 D-Glucose[p] |
| CDS.5224 | Trehalose-phosphatase (TRE6PP) | $H_2O$ + tre6p → pi + Trehalose |

**Table 2.4-13. Corresponding reactions of the suggested four gene knockouts in Gx 399.** Shown are the reactions which will have no flux as a result of the four gene deletions.

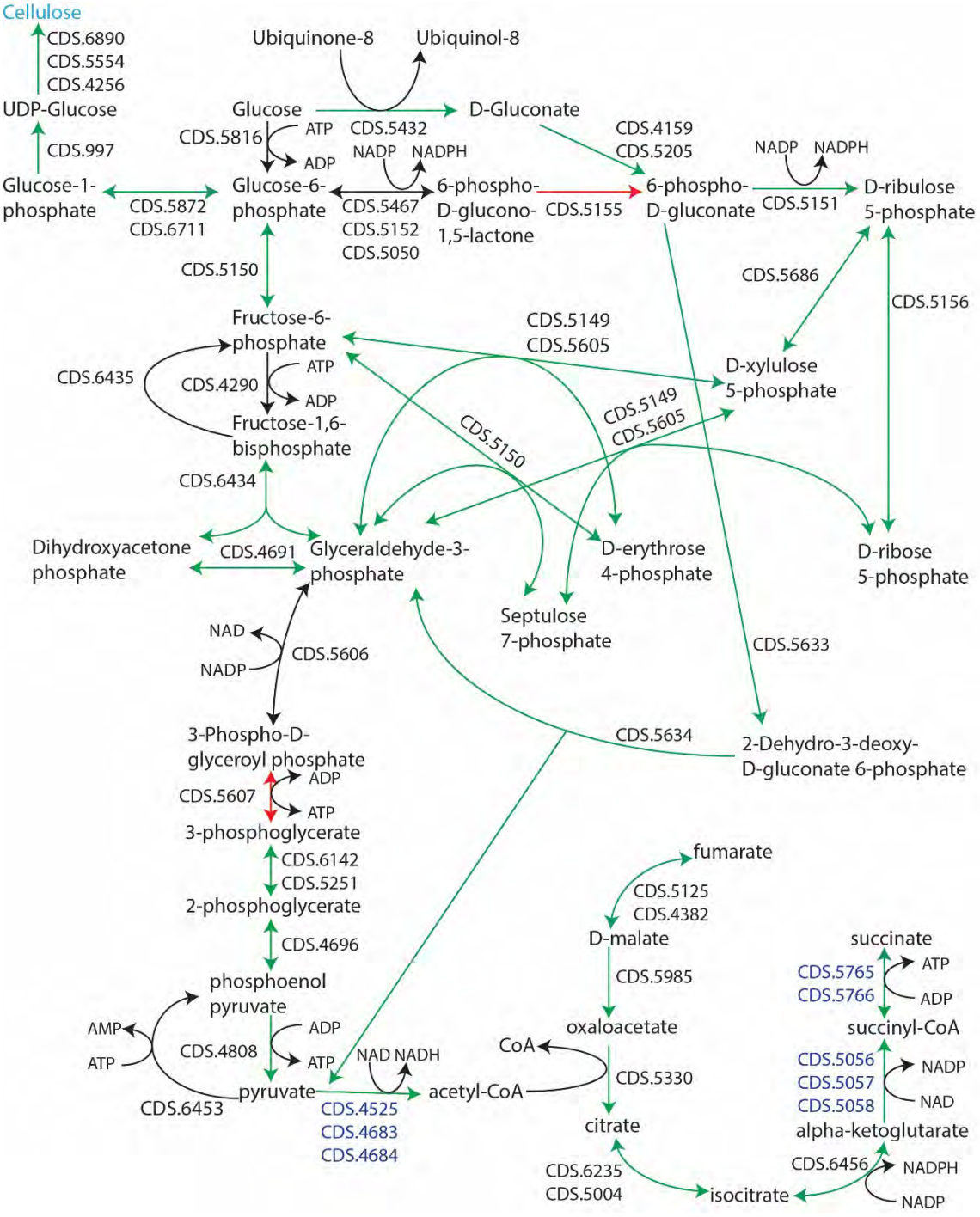



**Figure 2.4-7. Flux map for the Gx 399 mutant with four gene knock outs.**
Reactions with green arrows have flux, black arrows have no flux and red arrows should be deleted.

### *2.4.3 Conclusions*

In this study, the reconstructed metabolic networks of *G. xylinus* ATCC 53582 and Gx 399 were used to efficiently identify gene deletion strategies were which were predicted to enhance the yield of biocellulose. For *G. xylinus* ATCC 53582, a biocellulose yield of 0.70 g biocellulose/g glucose was predicted with five gene deletions, while and 0.41 g biocellulose/g glucose was predicted with four gene deletions. In Gx 399, a yield of 0.58 g biocellulose/g glucose when five deletions, and 0.41 g biocellulose/g glucose with four deletions.

In all instances, the predicted growth rate of the deletion mutants was reduced (by 50 to 90%), and five deletion solutions resulted in a greater reduction in the predicted growth rates than four deletion solutions.

While difficulties in genetic manipulation of the *Gluconacetobacter* strains impeded progress on this project, the metabolic modeling work points to alternative ways that biocellulose production can be increased, and warrants further study.

## 2.5 Task 5 – Fermentation optimization for biocellulose production

The results summarized in this section here will be published in the following scientific articles:

> Fuller M.E., Andaya, C, and McClay, K. Evaluation of ATR-FTIR for analysis of biocellulose impurities (in preparation for submission to Journal of Microbiological Methods).

### *2.5.1 Goal and Introduction*

The goal of this task was to develop optimized media and scalable fermentation procedures for producing biocellulose from the engineered bacterial strains. Additionally, post-fermentation processing methods and standardized analyses for assessing biocellulose purity were developed and evaluated.

### *2.5.2 Methods, Results, and Discussion*

#### 2.5.2.1 Fermentation Optimization

While multiple experiments were planned for this subtask, the difficulties encountered during the other tasks precluded this work on optimization to be performed.

#### 2.5.2.2 Biocellulose Processing

**Methods - Biocellulose Processing**

Based on the existing literature, a basic biocellulose processing protocol was developed as follows:

| | | | | |
|---|---|---|---|---|
| A | B | C | 1) | Biocellulose pellicles were removed from the incubation vessel (flask, ziplock bag, etc.) and squeezed by hand to wring out excess medium |
| | | | 2) | The pellicle was rinsed with distilled water |
| | | | 3) | The pellicle was placed in 1 M NaOH overnight at room temperature |
| | | | 4) | The pellicle was rinsed with distilled water until the pH of the rinsate was approximately 8. |
| | | | 5) | The pellicle was placed in 96% ethanol for 2 hours |
| | | | 6) | The pellicle was rinsed thoroughly with distilled water |

Pellicle were dried after processing (room temp or 35°C), and stored in sealed ziplock bags until analysis or further use.

Strains ATCC 53582 and Gx 399 were grown in standard minimal medium with glucose in 2 L ziplock bags. Triplicate ziplock bags of each strain were grown. Pellicles were harvested from each ziplock and rinsed with distilled water. Each pellicle was then split into three roughly equal portions. One portion underwent no further processing and represented minimal processing (A). The other two portions of each pellicle were treated with NaOH overnight, then rinsed with distilled water until the pH was neutral. These were designated as partial processing (B). The final portion of each pellicle was then soaked in ethanol overnight, then rinsed with distilled water. These were designated as complete processing (C). All the biocellulose was then dried and stored in sealed ziplocks until analysis.

An approximately 1" x 1.5" coupon was cut from each dried pellicle, labeled, placed in a desiccator, and allowed to equilibrate for at least 24 hr outside the desiccator. The coupons were analyzed (10 scans per coupon) using an Agilent Cary 660 Fourier-Transform infrared (FTIR) spectrophotometer equipped with a Pike MIRacle attenuated total reflection (ATR) assembly with a ZnSe crystal and a liquid nitrogen cooled linearized mercury cadmium telluride (MCT) detector (Agilent Technologies, Wilmington, DE, USA). Data was collected using Agilent Resolutions software (v. 5.2 and 5.3). Parameters of the standard analysis method are shown in Figure 2.5-1. The ATR clamp and platform were cleaned with a cotton swab dampened with isopropyl alcohol between analysis of each coupon

**Results and Discussion - Biocellulose Processing**

The FTIR spectra for the differentially processed biocellulose from strain ATCC 53582 and Gx 399 are shown in Figure 2.5-2 to Figure 2.5-3. For reference, spectra from experimentally contaminated biocellulose (as described in section 2.5.4 below) are also shown. A qualitative assessment indicated that even minimal processing (rinsing the pellicle with just distilled water) removed impurities to below 0.1% (w:w).

When the data were analyzed using the quantitative impurity models (as described in section 2.5.4 below), the results also indicated that minimal processing resulted in the removal of almost all impurities (Table 2.5-1).

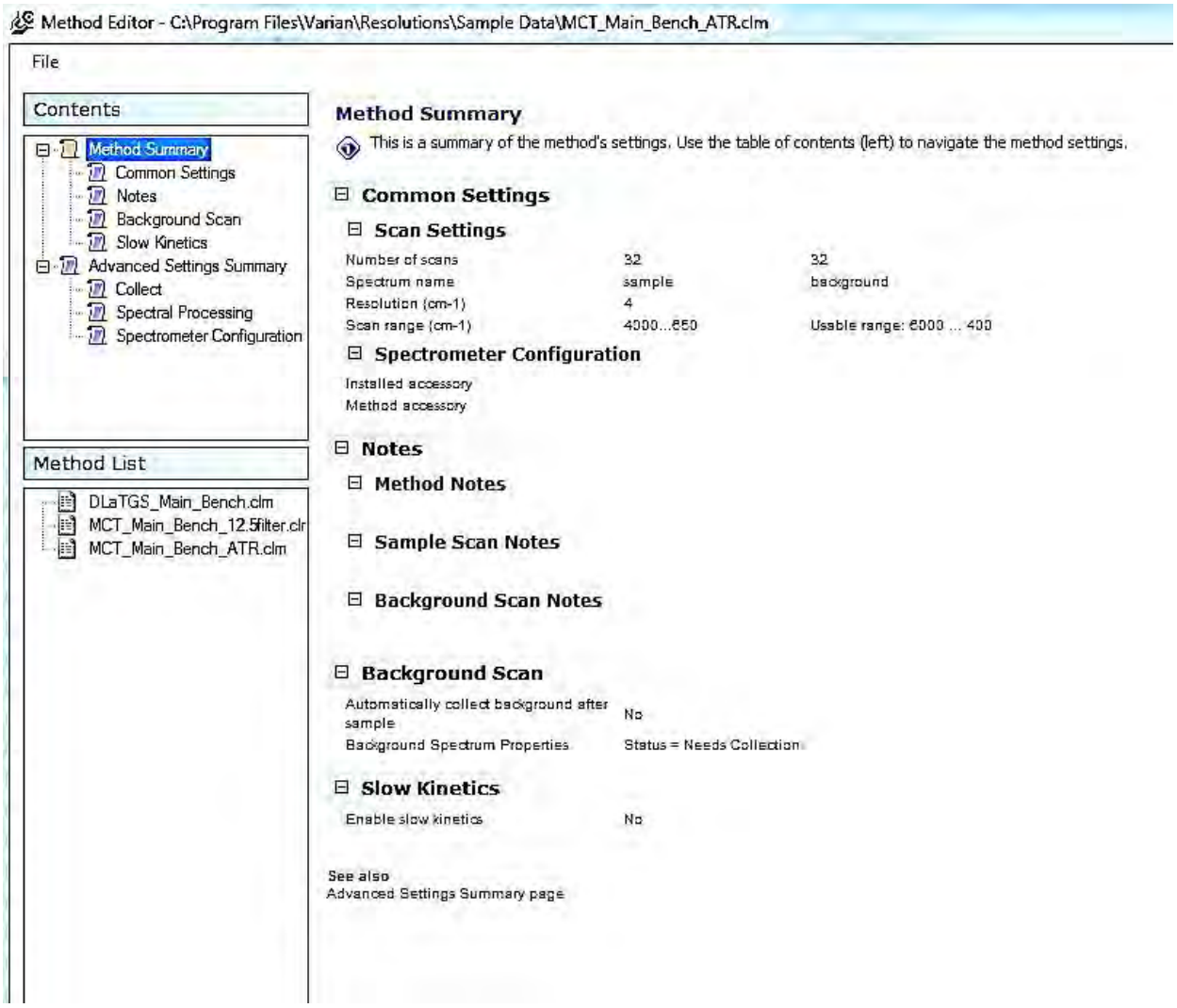



**Figure 2.5-1. Method parameters for FTIR analysis of biocellulose samples.**

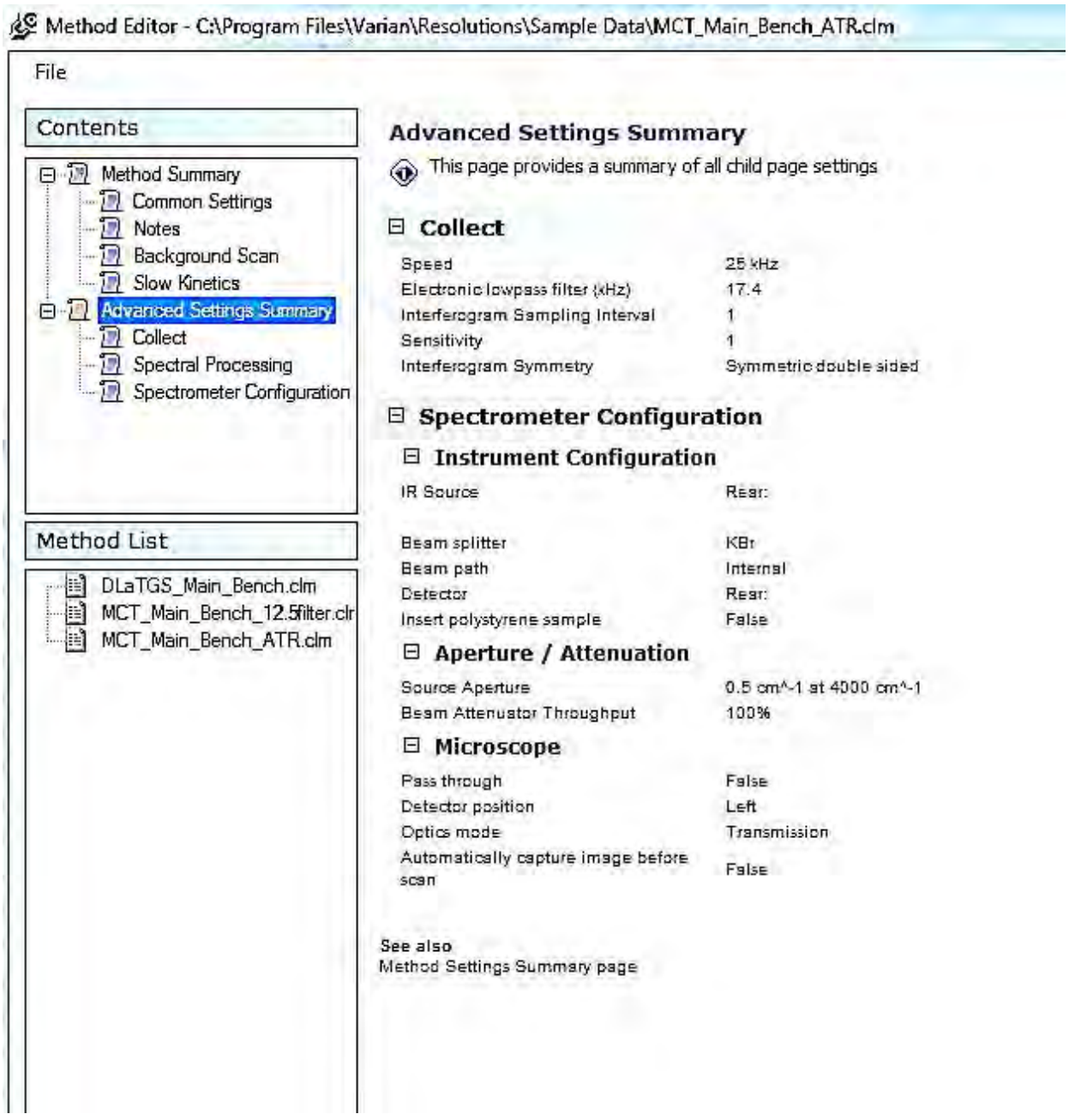



**Figure 2.5-1 (cont.)**

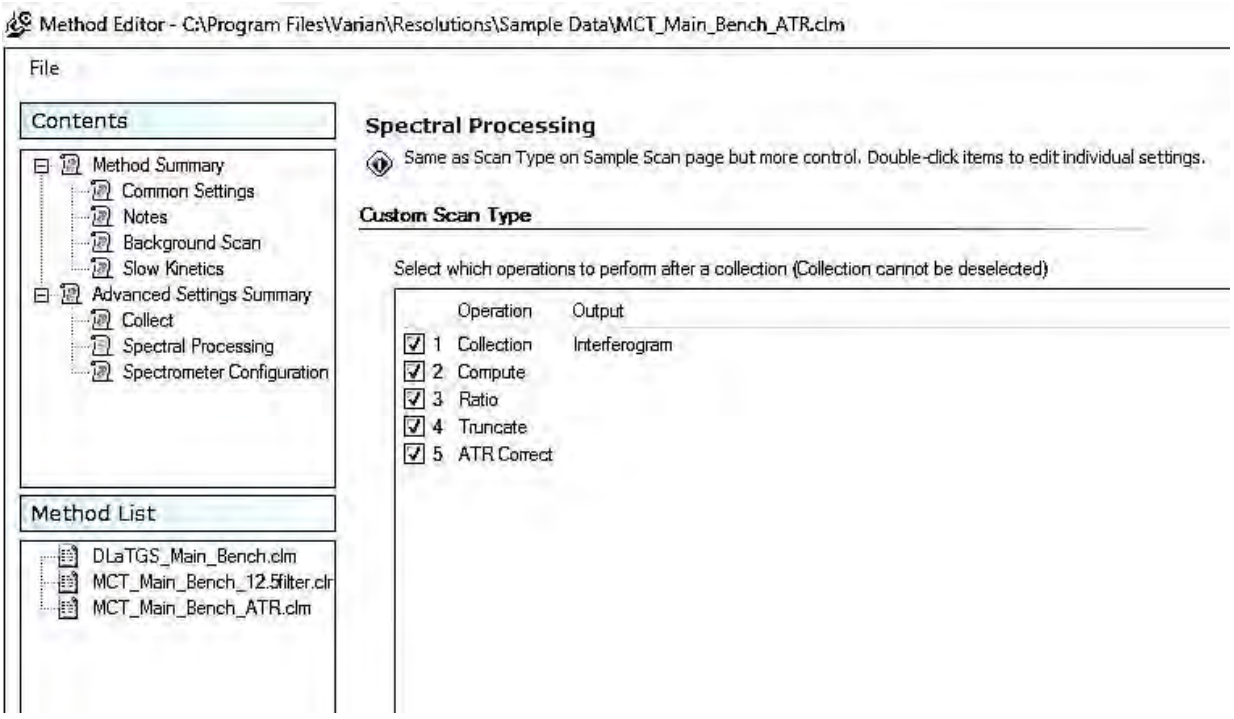



**Figure 2.5-1 (cont.)**

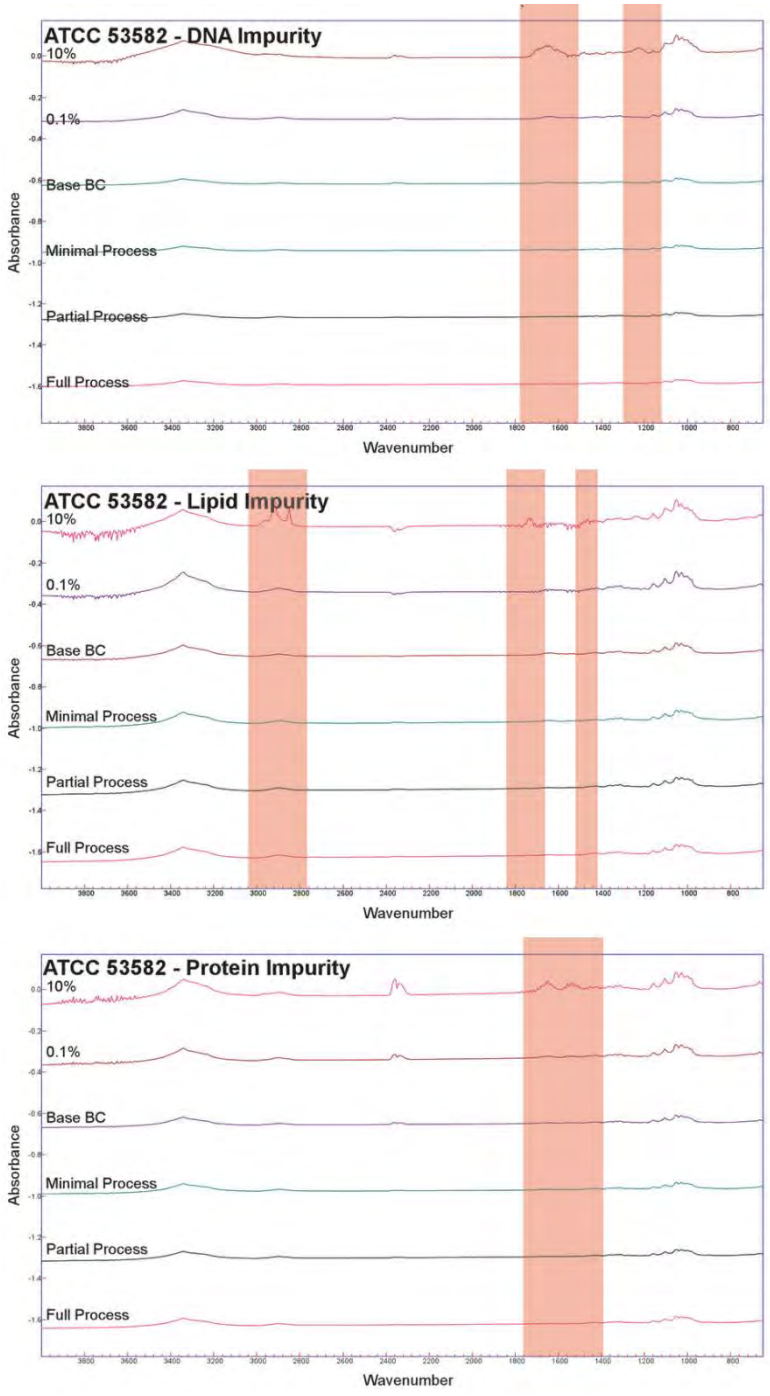



**Figure 2.5-2. FTIR spectra of biocellulose samples produced by strain ATCC 53582 and subjected to three levels of processing.**
Spectra for biocellulose experimentally contaminated with 0.1% (w:w) and 10% (w:w) DNA, lipid, and protein are shown for comparison

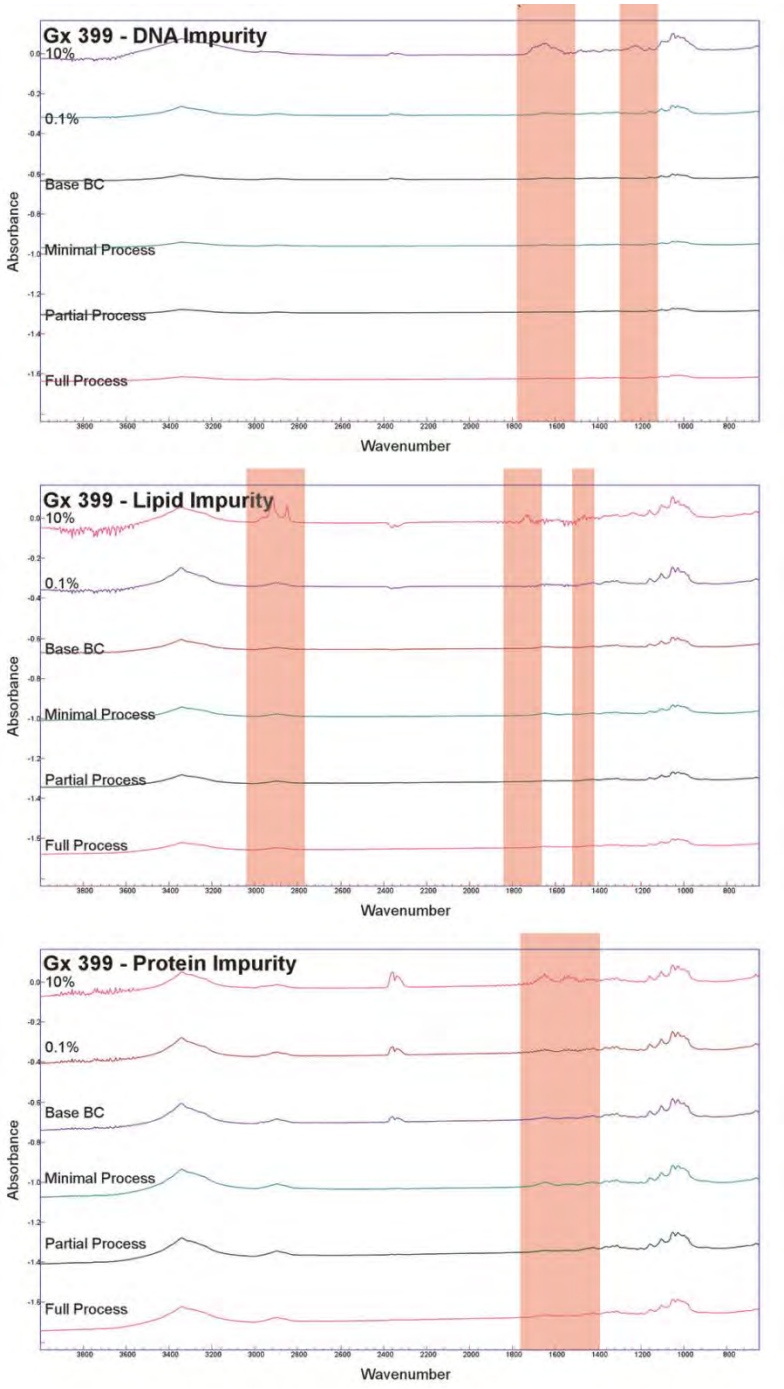



**Figure 2.5-3. FTIR spectra of biocellulose samples produced by strain Gx 399 and subjected to three levels of processing.**
Spectra for biocellulose experimentally contaminated with 0.1% (w:w) and 10% (w:w) DNA, lipid, and protein are shown for comparison

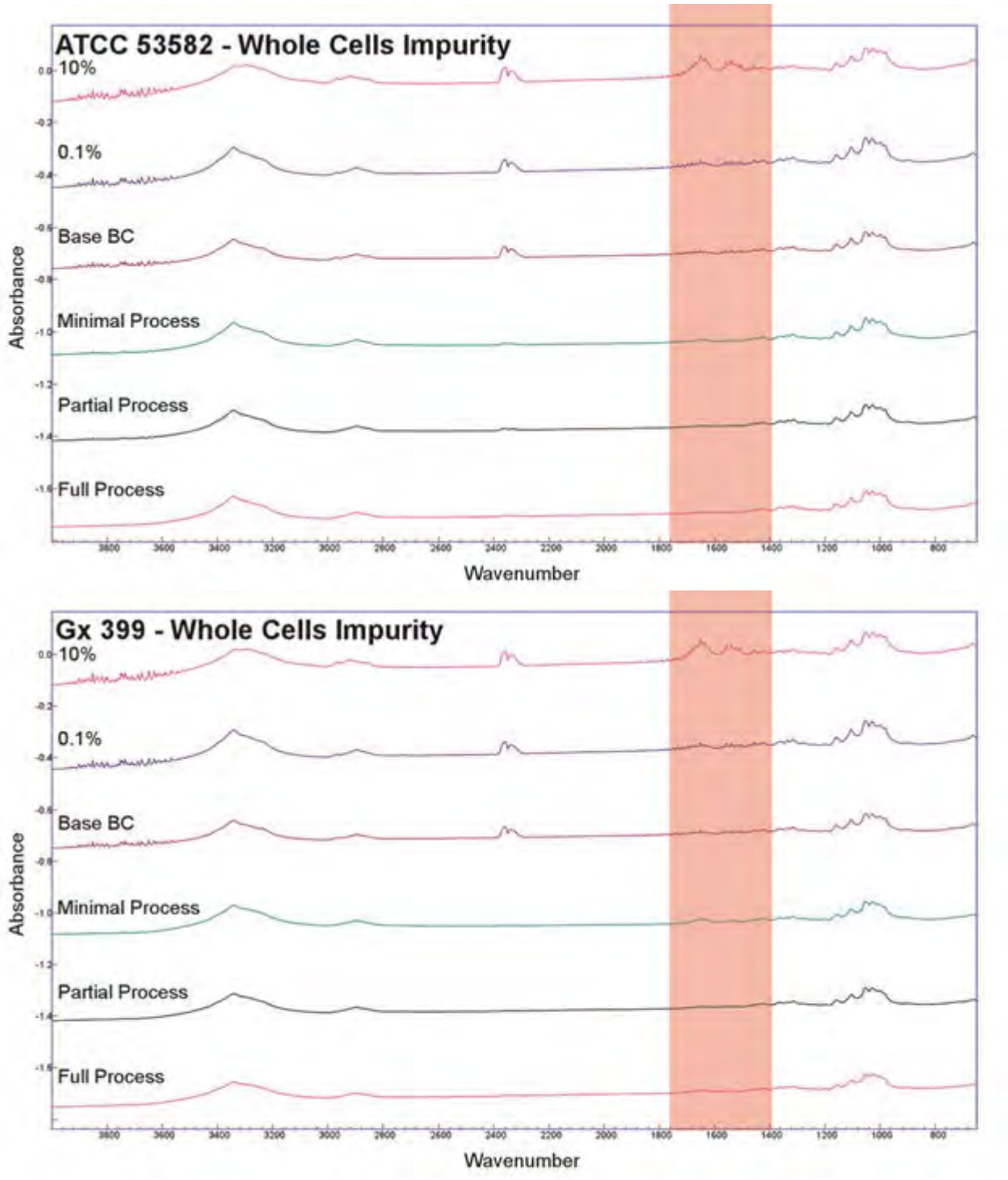



**Figure 2.5-4. FTIR spectra of biocellulose samples produced by strain ATCC 53582 and Gx 399 and subjected to three levels of processing.**
Spectra for biocellulose experimentally contaminated with 0.1% (w:w) and 10% (w:w) whole cells are shown for comparison

| Strain | Processing | Replicate | % (w:w)* DNA | Lipid | Protein |
|---|---|---|---|---|---|
| ATCC 53582 | Minimal | 1 | <0.1 | <0.1 | <0.1 |
| | | 2 | <0.1 | <0.1 | <0.1 |
| | | 3 | <0.1 | <0.1 | <0.1 |
| | Partial | 1 | <0.1 | <0.1 | <0.1 |
| | | 2 | <0.1 | <0.1 | <0.1 |
| | | 3 | <0.1 | <0.1 | <0.1 |
| | Full | 1 | <0.1 | 0.1 | <0.1 |
| | | 2 | 0.03 | <0.1 | <0.1 |
| | | 3 | <0.1 | <0.1 | <0.1 |
| Gx 399 | Minimal | 1 | <0.1 | <0.1 | <0.1 |
| | | 2 | <0.1 | <0.1 | <0.1 |
| | | 3 | <0.1 | <0.1 | <0.1 |
| | Partial | 1 | <0.1 | <0.1 | <0.1 |
| | | 2 | <0.1 | <0.1 | <0.1 |
| | | 3 | <0.1 | <0.1 | <0.1 |
| | Full | 1 | <0.1 | <0.1 | <0.1 |
| | | 2 | <0.1 | <0.1 | <0.1 |
| | | 3 | <0.1 | <0.1 | <0.1 |

*Negative impurity estimates were interpreted as less than the lowest impurity standard, <0.1%

**Table 2.5-1. Impurity analysis of biocellulose from two strains processed to different levels.**

### 2.5.2.3 Biocellulose Purity Assessment

**Methods - Biocellulose Purity Assessment**

Base biocellulose was produced by strain ATCC 53582 under static incubation in 2 L ziplock bags. The complete biocellulose washing procedure (steps 1 to 6) cited in section 2.5.2.2 above was used, followed by several rinses in chloroform to assure maximum purity.

The sheets of base biocellulose were cut into coupons (approximately 3 cm x 3 cm). Solutions of potential impurities that may be encountered (e.g., nucleic acid (DNA), lipid, protein, alginate (an extracellular polymer) were applied to the coupons to achieve a range of nominal impurity levels (0.1% to 10%, w:w). Coupons with no added impurity were included as the control (e.g., 0% impurity) treatment.

The impurities used are listed below.

| | |
|---|---|
| Lipid | 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (Sigma-Aldrich, P4329), in chloroform |
| DNA | salmon testes DNA (Sigma-Aldrich, D7656), applied as an aqueous solution |
| Protein | bovine serum albumin (Sigma-Aldrich, A0281), applied as an aqueous solution |
| Alginate | alginic acid, sodium salt (Sigma-Aldrich, A0682), applied as an aqueous solution |

The coupons were thoroughly dried at room temperature, then stored in a desiccator until FTIR analysis.

Coupons were also contaminated with whole bacterial cells as the complex impurity most likely to be expected in the biocellulose. Strain Gx 399A (biocellulose negative mutant) cells were grown on glucose, washed twice in sterile purified water, and then diluted to an optical density at 600 nm of 5. Assuming that a bacterial cell weighs $1 \times 10^{-9}$ mg (dry weight), then the OD 5 cell solution contained 2.5 mg/ml of bacterial dry mass. The cell solution was applied to the coupon, allowed to dry, then the coupon was place in a desiccator as with the other coupons.

Each coupon was subjected to twenty scans at different locations on coupon to capture spatial variability of the applied impurity distribution. The ATR clamp and platform were cleaned with a cotton swab dampened with isopropyl alcohol between analysis of each coupon. The data was transferred into Panorama Pro for multiple regression analysis, building a model that could then be used to determine the levels of the different impurities in a given BC sample.

**Results and Discussion - Biocellulose Purity Assessment**

The FTIR spectra were examined visually to evaluate the ability of the FTIR to detect the various impurity types, A qualitative assessment of purity was made by averaging replicate scans for each impurity at each concentration. Spectra for the lipid impurity, showing increasing peak size in key regions of the spectra with increasing percentage of lipid impurity, are shown in Figure 2.5-5. The spectra for the other impurity types also evidenced a good correlation between impurity concentration and peak height, except the changes were in different key regions of the spectra (data not shown). The spectra obtained from whole cells also showed some changes in specific regions of the spectra indicative of protein functional groups as the concentration increased (Figure 2.5-6). Comparison of spectra obtained from biocellulose contaminated with 10% (w:w) of the different impurities is presented in Figure 2.5-7.

The Partial Least Squares 2 calibration model in Panorama Pro was used as it allowed multiple regression of several of the impurities at once. Using data from selected regions of the FTIR spectra from a total of 400 scans (e.g., 20 scans per impurity concentration, 5 concentrations, 4 impurities), the PLS2 model appeared to yield good results for alginate, DNA, and lipid (Figure 2.5-8 to Figure 2.5-11), with regression coefficients (actual vs. predicted concentration) of greater than 0.8. This initial PLS2 calibration model did not work well for the protein (regression coefficient of less than 0.6). However, use of only the protein dataset with another model, PLS1, yielded much better results (Figure 2-4), including a regression coefficient of 0.9 for actual vs. predicted concentrations.

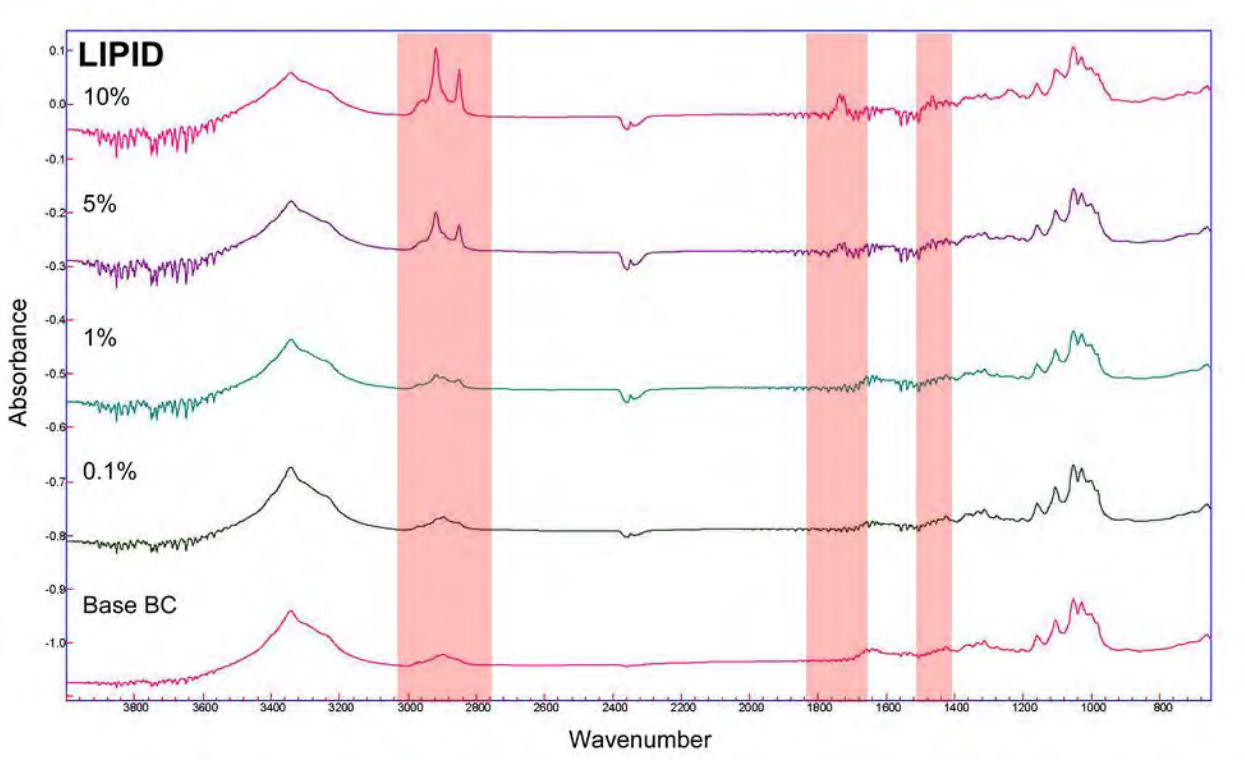



**Figure 2.5-5. FTIR spectra of biocellulose samples contaminated with 0% to 10% lipid (w:w).**
Each spectra represents the average of 20 scans. Shaded regions indicated where the FTIR signals of lipid functional groups are expected to be observed.

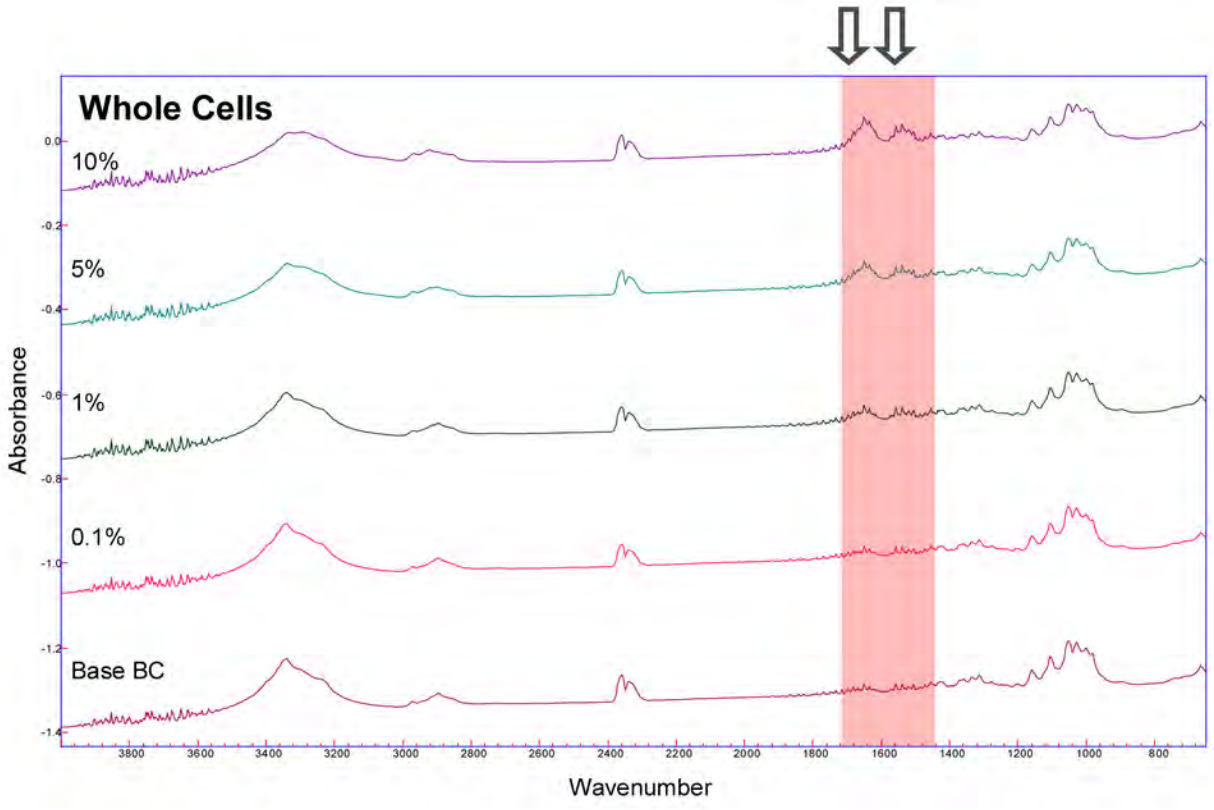



**Figure 2.5-6. FTIR spectra of biocellulose samples contaminated with 0% to 10% whole bacteria cells (w:w).**
Each spectra represents the average of 20 scans. Shaded regions indicated where differences in the FTIR spectra varied with cell concentration.

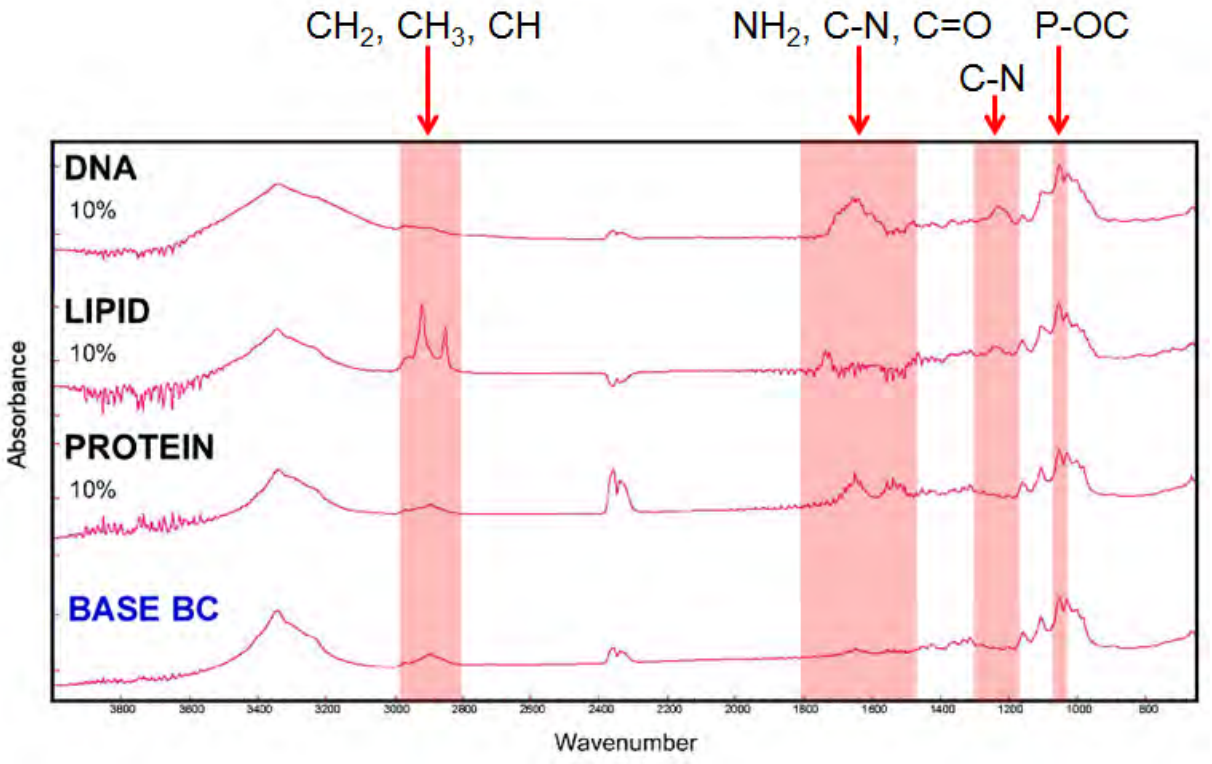



**Figure 2.5-7. FTIR spectra of biocellulose samples contaminated at 10% (w:w) with different impurity types.**
Each spectra represents the average of 20 scans. Shaded regions indicate where the FTIR signals of specific functional groups are expected to be detected.

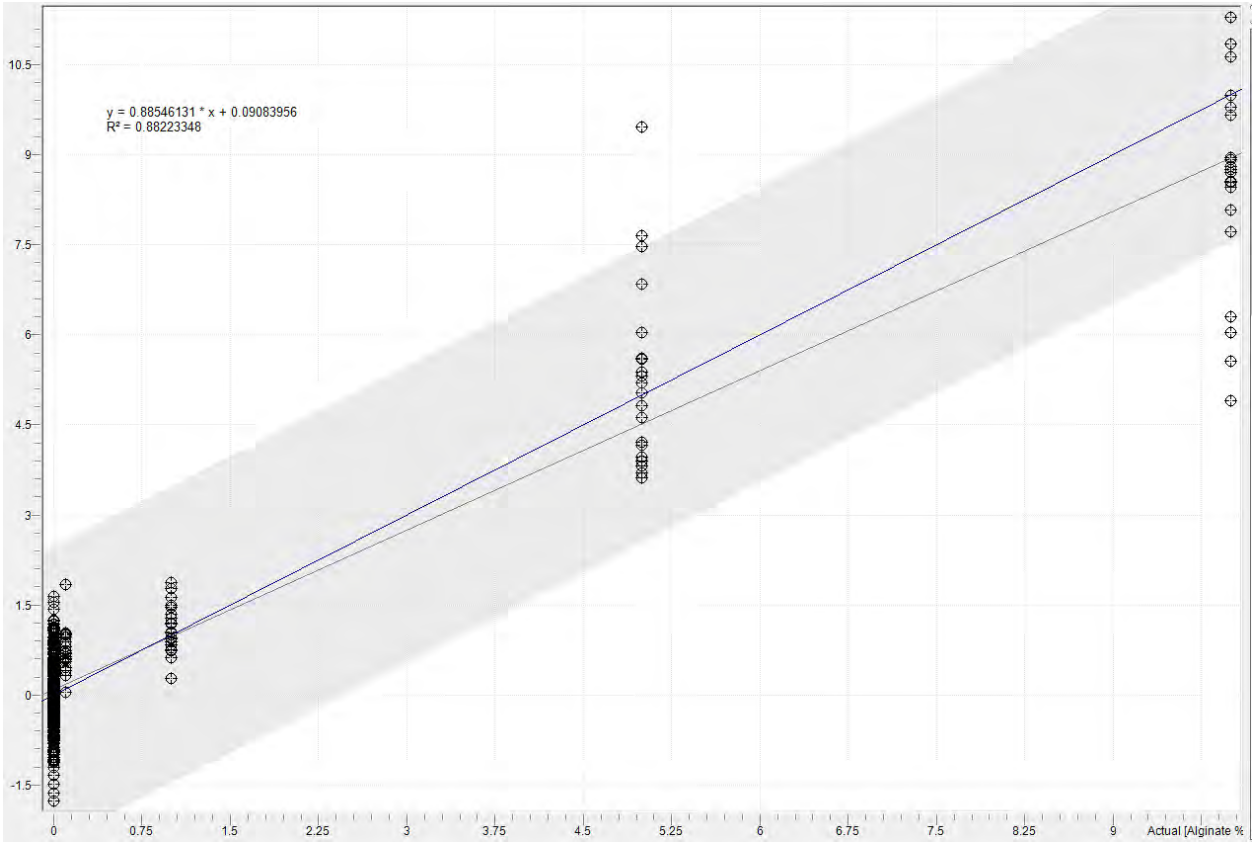



**Figure 2.5-8. PLS2 calibration model of actual vs. predicted values for alginate impurity.**

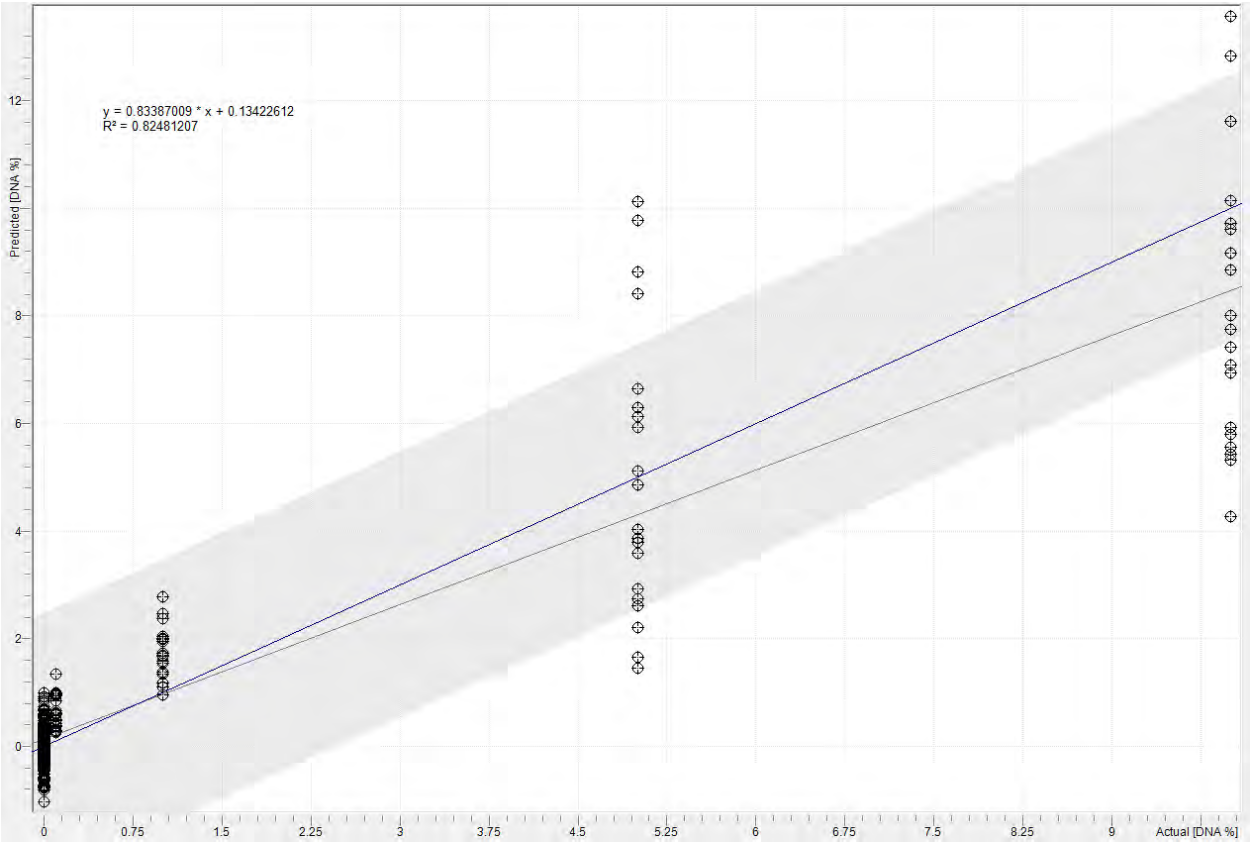



**Figure 2.5-9. PLS2 calibration model of actual vs. predicted values for DNA impurity.**

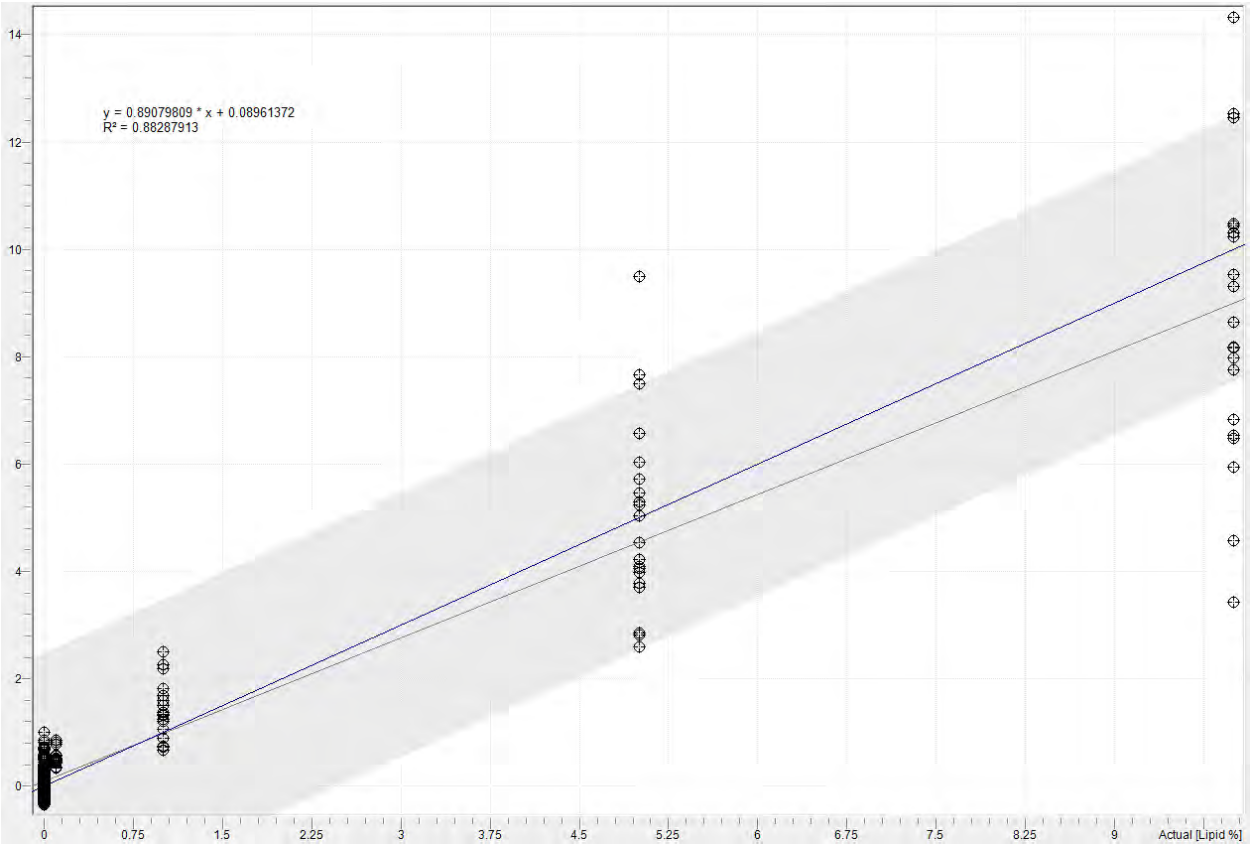



**Figure 2.5-10. PLS2 calibration model of actual vs. predicted values for lipid impurity.**

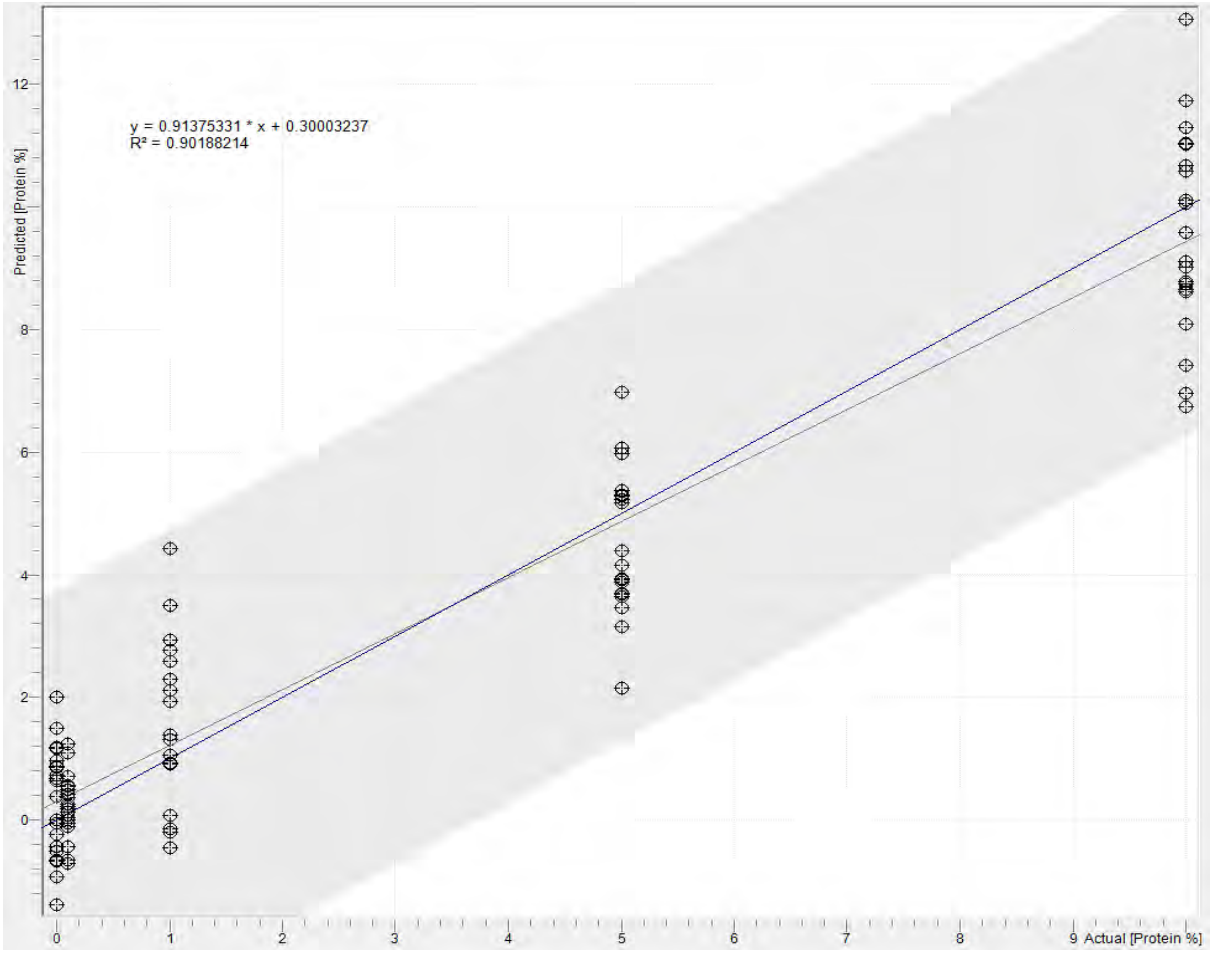



**Figure 2.5-11. PLS1 calibration model of actual vs. predicted values for protein impurity.**

### *2.5.3 Conclusions*

The results presented here indicated that minimal processing of the BC resulted in the removal of almost all impurities. While it could not be fully evaluated during this project, these data indicate that even minimally processed BC should be a suitable substrate for production of NC with very few impurities, which would be expected to extend NC shelf life and performance.

Further, the ATR-FTIR method developed during this project was shown to be effective for detecting the type and relative abundance of bacterial-derived impurities in biocellulose. This extends the use of ATR-FTIR analysis beyond simply measuring generic BC purity as reported in previously published literature. This method should be applicable to analysis of BC purity in the broader biocellulose research community to assess the effects of fermentation conditions and BC processing procedures.

## 2.6 Task 6 – Production of nitrocellulose from biocellulose

### *2.6.1 Goal and Introduction*

The goal of this task was to develop a biocellulose processing and nitration protocol to produce a usable NC powder for energetic formulations if a lacquer is not used. To achieve this goal, several approaches were attempted.

### *2.6.2 Methods, Results, and Discussion*

#### 2.6.2.1 Initial processing and standard nitration procedures

The initial biocellulose sheets produced were very tough, low density material that was difficult to process into a power with the current facilities. Initial attempts at grinding the material were not successful. These included:

a) Using a mortar and pestle with and without solvents and freezing in liquid nitrogen.

b) Using a dry mill, which produced a rather coarse and non-uniform material (Figure 2.6-1), which was used in the first nitration reaction described below. The material was lightweight enough to be entrained in the air flow generated by the high speed blades and the chopping efficiency was very low as the material mostly followed the blades. A great deal of static charge was also generated and the material was difficult to control as it readily attached itself to surfaces.

c) Using a SS laboratory blender with a liquid medium. Although a water mixture provided satisfactory chopping action simply collecting the BC on a filter and drying re-produced a sheet like form. Using an aprotic liquid such as chloroform in the blender produced material that was only slightly more uniform than dry chopping (Figure 2.6-1).

Initial attempts to nitrate the chopped biocellulose were performed based on the optimized methods presented in Sun et al. (2010) (63), which showed a degree of substitution (DS) of 2.85 (where 3.0 is fully nitrated NC), using a reaction temperature of 30°C, esterification coefficient of 58 (weight ratio of acid mixture to cellulose), water content of 8%, ratio of sulfuric acid to nitric acid of 3:1, and a reaction time of 30 minutes.

In a 100 ml round bottom flask fitted with a magnetic stir bar, a 26.1 g nitrating solution was prepared using 5.89 g 100% nitric acid, 18.26 g sulfuric acid and 1.94 g water. The flask was immersed in a 30°C water bath and the solution was stirred for 30 min. A 0.45 g sample of the dry chopped BC (dried for 24 h at 100°C) was added and stirring continued to wet the solid. The solid appeared to swell in the nitrating mixture, which prevented the liquid fraction flowing uniformly in the mix (Figure 2.6-2). Manual rotation of the flask was used to attempt to provide uniform mixing during the 30 min reaction period.

After 30 minutes the flask contents were poured over 100 g of crushed ice, and the flask was rinsed with distilled water and decanted over the ice. After stirring the quenched reaction for an additional 30 min, the nitrated BC product was collected on a fritted glass disk. The white solid was transferred to a 200 ml beaker with 100 ml distilled water and the mixture brought to a boil

for 5 min with stirring. The solid was collected by suction filtration and washed with water and this boiling/washing repeated two more times to remove any remaining acid. The particle size and shape of the white solid product appeared very similar to the starting material (Figure 2.6-2).

The solid bionitrocellulose (BNC) was stored wet for further analysis. A proton NMR spectrum of the BNC in $D_6$-DMSO (Figure 2.6-3) is consistent with a DS higher than 2.3 but less than 3. For comparison, a spectra for cellulose nitrate from Wu et al. (1980) (70) is also shown. The FTIR spectra of two batches of the starting BC are shown in Figure 2.6-4.

Based on feedback from China Lake, some additional post-harvest work was performed to generate a BC feedstock that would be more amenable to nitration. The first approach was to use an analytical mill to manually cut up the sheets of BC. This resulted in ribbon-like strips of material and was not deemed successful.

The second approach was to re-wet the BC and form wadded up balls, dry the wads, and grind them in the analytical mill after cooling with liquid nitrogen. This produced smaller BC particles (Figure 2.6-5), which could be further reduced to a powder-like material with prolonged grinding. In order to facilitate making BC wads that were amenable to grinding in the mill, the use of a Waring blender was examined to perform an initial maceration of wet BC sheets. This approach worked very well, and the macerated BC could easily be formed in wads <5 cm across prior to being dried and ground. This material was sent to China Lake for evaluation.

**Figure 2.6-1. BC chopped using a dry method (left) and using biocellulose suspended in chloroform.**

**Figure 2.6-2. Appearance of the nitration mixture (left) and wet BNC product (right).**

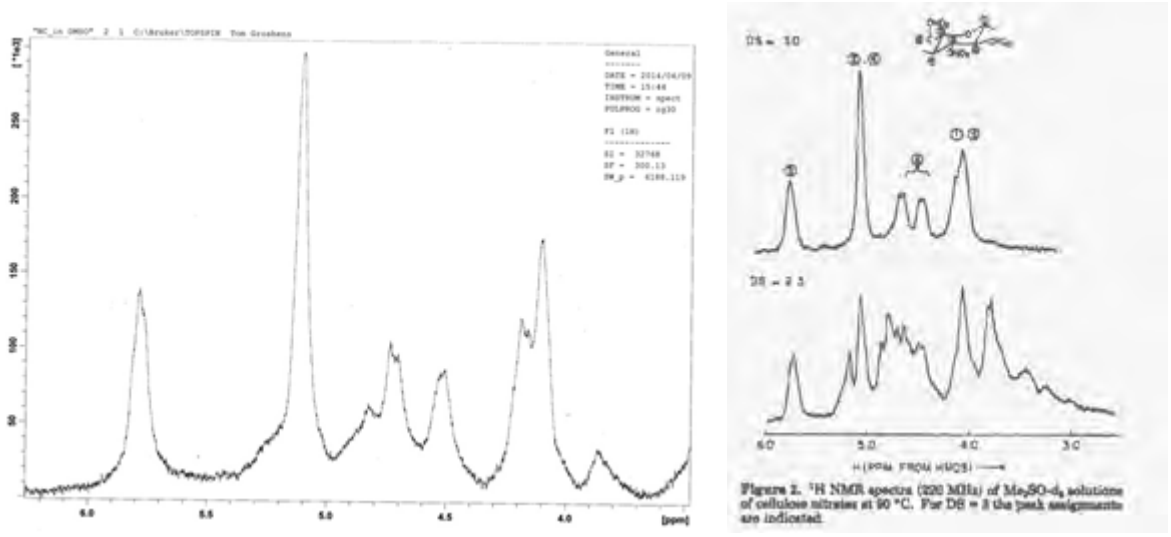


**Figure 2.6-3. Proton NMR spectra of BNC product in DMSO (left) and the cellulose nitrate of Wu et al (1980) (right).**

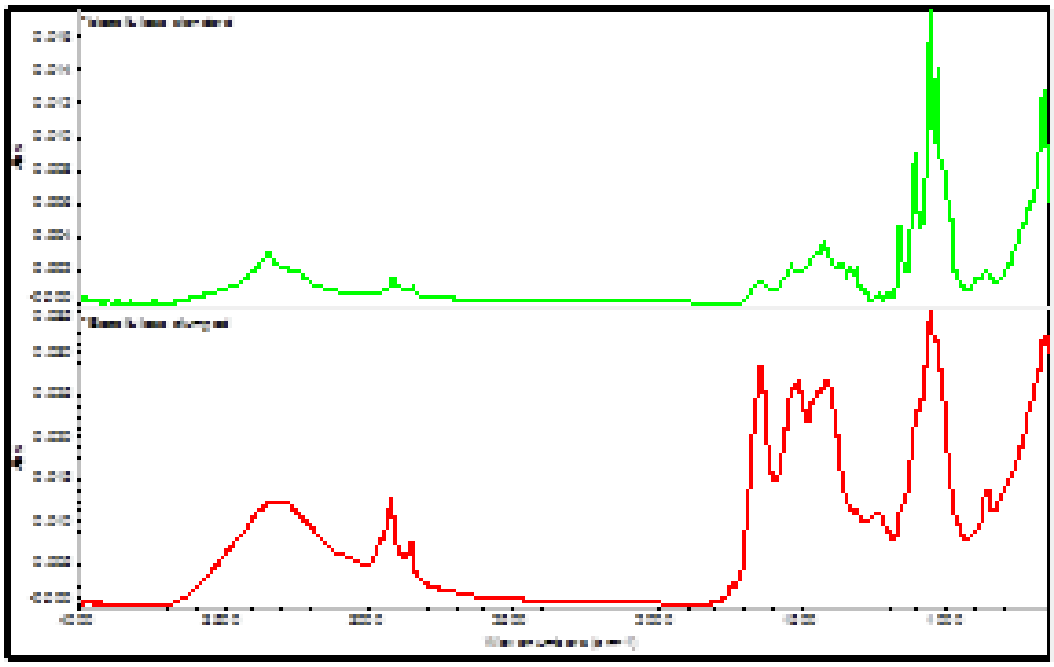


**Figure 2.6-4. FTIR of BC starting materials Jan2014-1 (top) and Jan2014-2 (bottom)**

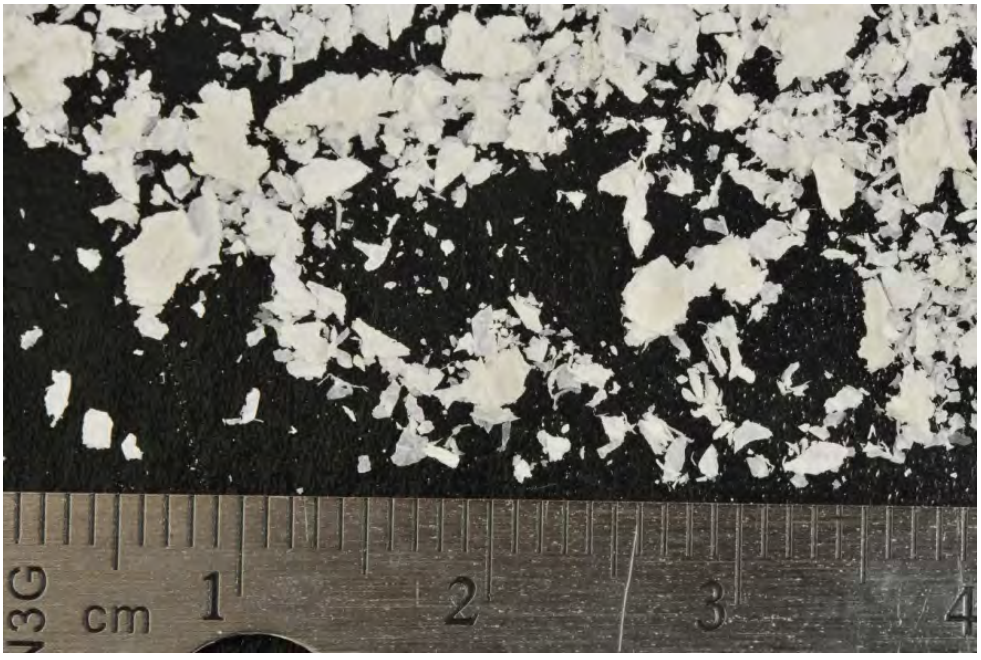



**Figure 2.6-5. Appearance of BC after extended grinding in an analytical mill with liquid nitrogen.**

### 2.6.2.2 Evaluation of nitrogen content in BNC

No nitrometer was available to determine the nitrogen levels in the BNC products by the preferred standard method and, shipping the anticipated large number of energetic samples out for elemental analysis was prohibitively expensive. Therefore, another method to provide satisfactory quantitative results on the degree of substitution (DS) in the BNC products was evaluated.

Several literature sources (13, 20, 48) detailed an FTIR method that provides nitrogen analysis results that were consistent with standard methods. Dr. Lee Cambria, Infrared Spectroscopist at China Lake, was able to set up and use this method to analyze the BNC products. To support cross-validation, useful military grade NC materials with certified nitrogen content were collected from the China Lake energetics inventory for use as calibration standards. With initial proof of nitration the effort at China Lake then focused on processing the as supplied BC sheets to generate a reproducible powder BC feedstock and analysis of nitrated BC to determine the degree of substitution in-house using FTIR analysis. A series of five NC samples were gathered for FTIR analysis including two with certified N levels, as presented in Table 2.6-1.

| **Sample** | **%Nitrogen** |
|---|---|
| Hercules NC RS5 | 12.19 %N (Certified) |
| ATK NC RS30-35 | 12.57 %N (Certified) |
| Hagedorn NC H33 | 11.8-12.2 %N |
| Pelletized NC from Indian Head (Lot # H92H100-026) | Not available |
| TG-1927-38-1NC Coarse ground BNC | Not available |

**Table 2.6-1. Standard NC materials used for FTIR validation.**

The Attenuated Total Reflectance (ATR) FTIR spectrum of all wet samples, as received, were recorded (Figure 2.6-6) and are similar with the exception of two additional unidentified $NO_x$ peaks in the TG-1927-38-1NC sample around 1517 $cm^{-1}$ and 1345 $cm^{-1}$ (dashed box, bottom plot). These are likely from $NO_x$ peaks generated by the denitration/hydrolysis action of residual acid in the nitrated BC sample. Presumably the inability to completely remove all of the acid during the boiling water washing steps in the nitration procedure is due to the large particle size of the BC feedstock retaining substantial capillary structure that traps the nitrating acid mix.

After drying the samples at 80°C under vacuum overnight the additional peaks were no longer present in the ATR FTIR spectrum of TG-1927-38-1NC (Figure 2.6-7) as would be expected for volatile $NO_x$ species.

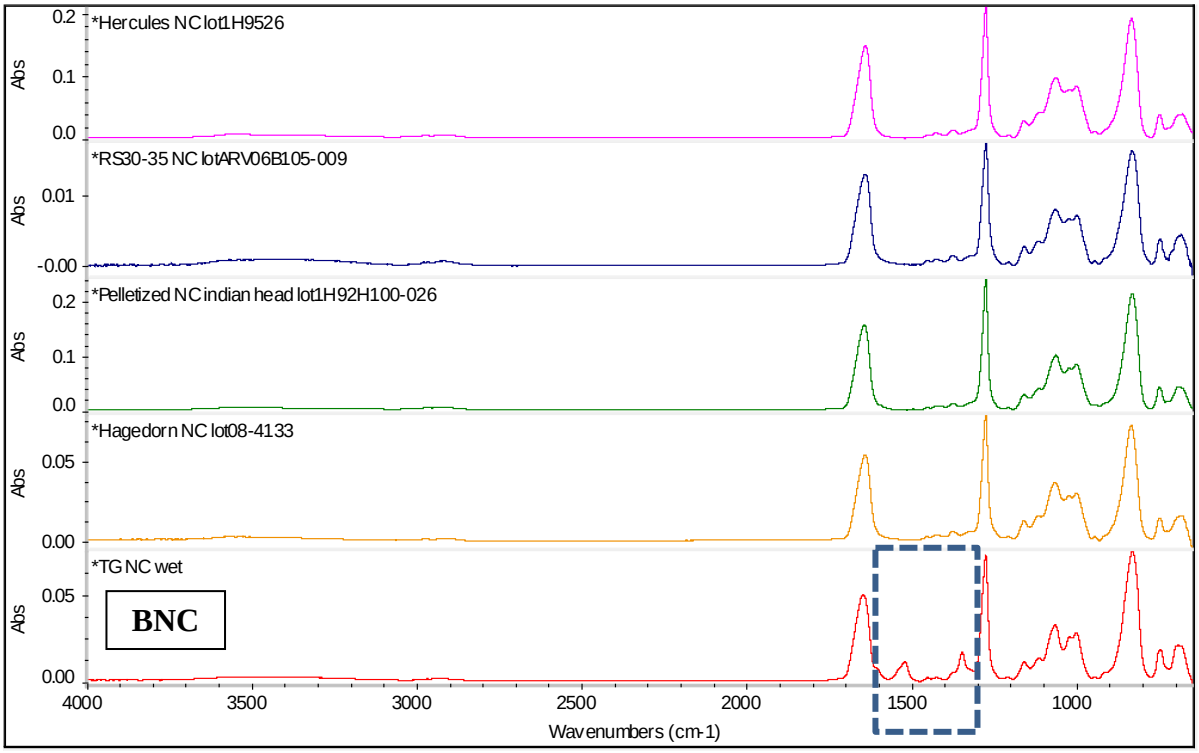



**Figure 2.6-6. ATR-FTIR ATR spectra of neat wet NC standard materials (top 4 spectra) and BNC (bottom spectra).**

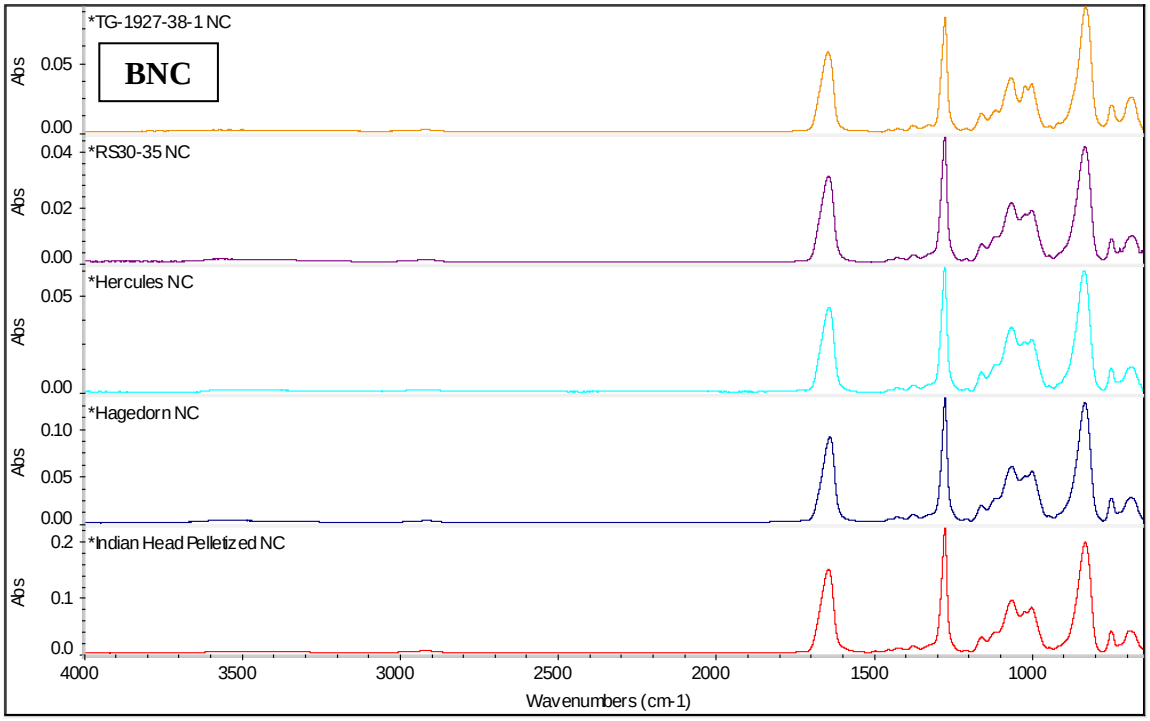



**Figure 2.6-7. ATR-FTIR ATR spectra of neat dry NC standard materials (bottom 4 spectra) and BNC (top spectra).**

The published FTIR method to determine the degree of substitution of NC was tested to decide if it could be validated and used to quantify the nitrogen content of nitrated BC samples in-house. All NC samples were thoroughly dried and dissolved in THF at a known concentration and the FTIR spectrum was recorded using a fixed path length solution cell. The total area was measured for the NO peak at about 6 μm, 1630 $cm^{-1}$ (Figure 2.6-8).

All of the samples with the exception of TG-1927-38-1NC completely dissolved in the THF solvent. There was clearly undissolved material in the TG-1927-38-1NC sample. This could be either a result of incomplete nitration of BC due to the large particle size of the feedstock, or partial denitration of the sample due to the presence of residual acid. As a result the FTIR analysis of the nitrated BC sample is likely to show a lower than expected nitrogen level and is not a reliable value due to the inhomogeneity of the large particle sample. Based on previously published work and the nitrating conditions used to generate the BNC sample TG-1927-38-1NC a 13.75 %N was expected and the measured level was 11.37 %N (Table 2.6-2 and Figure 2.6-9).

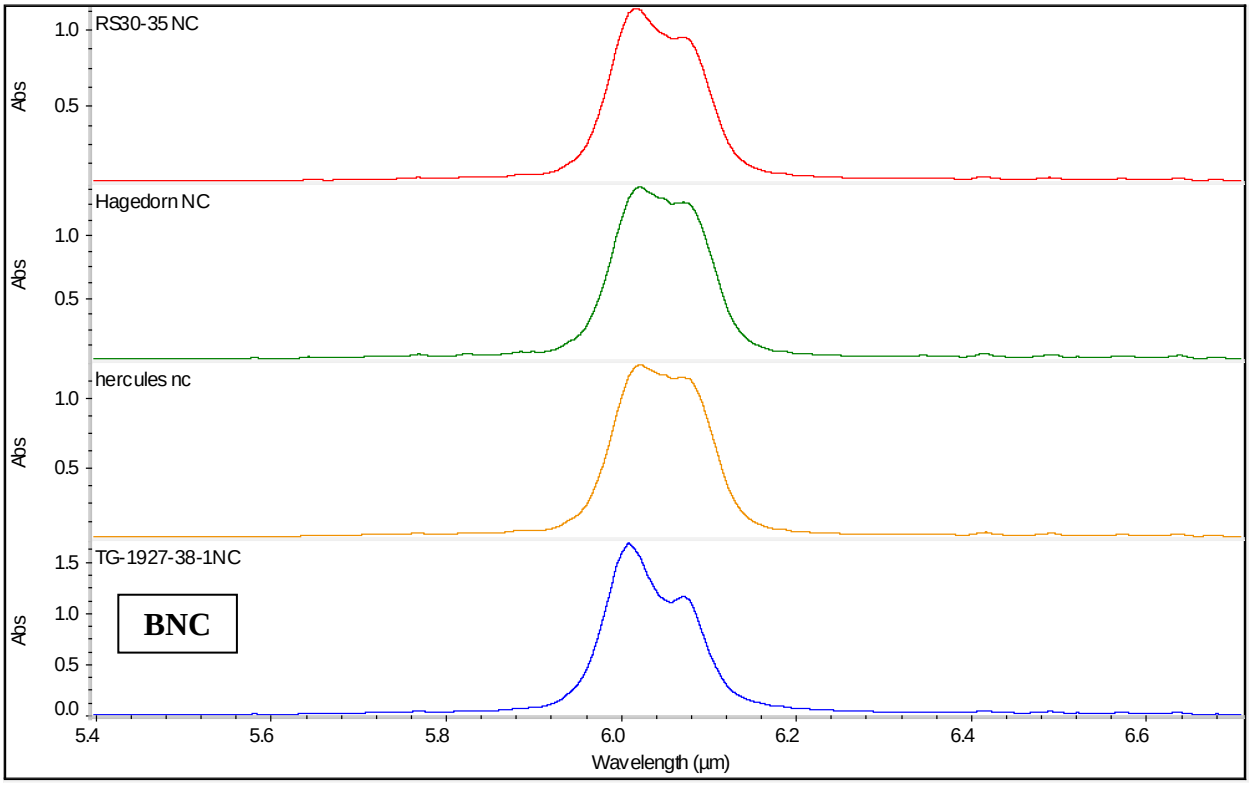



**Figure 2.6-8. FTIR NO peak of THF solutions of NC standard materials (top 3 spectra) and BNC (bottom spectra).**

| Sample | NO | Conc | Norm Conc | Normalized NO | Certified %N | Calculated %N |
|---|---|---|---|---|---|---|
| Hercules | 42.012 | 0.00788 | 1 | 42.012 | 12.19 | 12.19 |
| Hagedorn | 46.667 | 0.00762 | 0.966 | 48.291 | | 11.65 |
| RS30-35 | 36.574 | 0.00768 | 0.974 | 37.551 | 12.57 | 12.57 |
| TG BNC | 49.180 | 0.00751 | 0.953 | 51.603 | | 11.37 |

**Table 2.6-2. Calculated %N of standard NC materials based on FTIR analysis.**

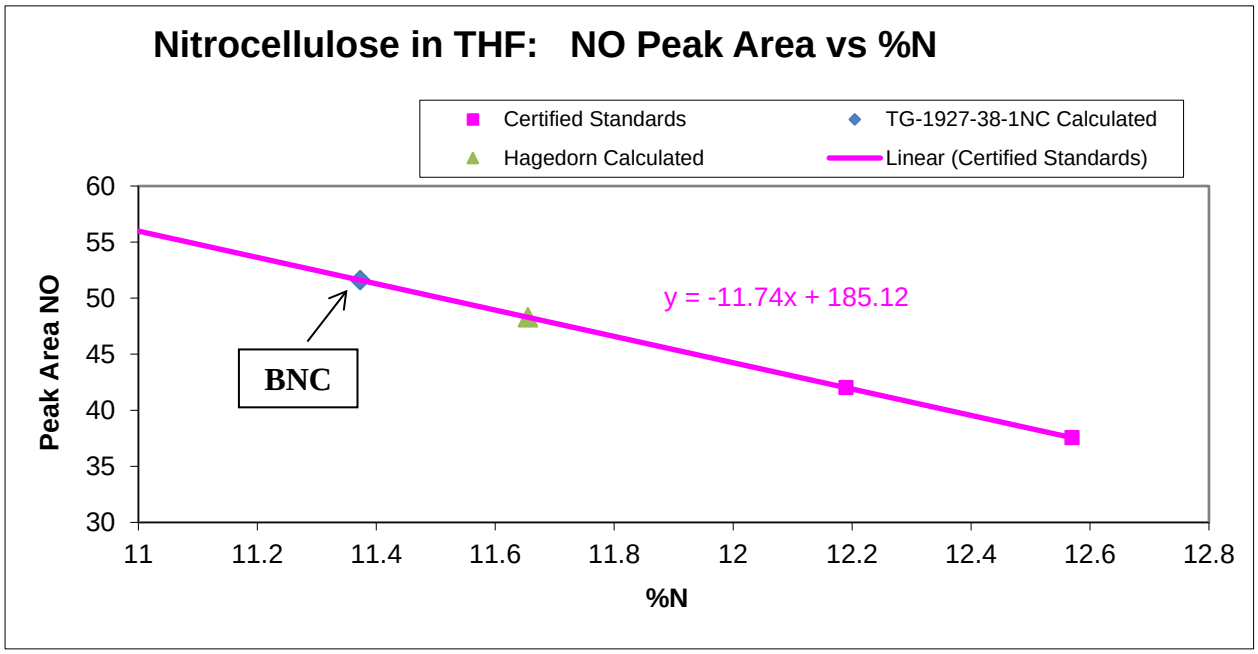



**Figure 2.6-9. NO Peak Area vs %N of neat dry NC standard materials and BNC in THF.**

Using the two certified NC standards, a calibration line was generated and the percent nitrogen values for other NC material were measured using FTIR spectra of THF solutions according to published methods. The sample from Hagedorn was found to be slightly less than the range provided by the manufacturer at the time of shipment. It should be noted that NC samples can drop in %N on storage. The value for the nitrated BC sample was determined to be unreliable due to the presence of insoluble material. To improve the accuracy of the method, additional standards would be required.

The FTIR analysis was useful in determining the presence of residual acid mix in the nitrated BC sample that must be removed to provide a stable NC product. However, at the present time the FTIR analysis has been found not to be suitable for in-house measurement of the relative degree of substitution of the BC during the nitration process.

NC digestion coupled with ion chromatography (IC) was also evaluated for determining %N. In this method, base hydrolysis was used to digest the NC and convert the nitrate ester to nitrate and nitrite ions that were quantitatively determined by IC. Christodoulatos et al. (12) determined that after alkaline hydrolysis consistently 92% was hydrolyzed to nitrite:nitrate in a 3:1 ratio while the remaining 8% was converted into other compounds or lost. More recently, Alinat et al. (1) extended this approach to determination of the nitrogen content of nitrocellulose by capillary electrophoresis after alkaline denitration.

The draft analysis procedure based on the published methods ((1, 12, 41)) was as follows:

Approximately 50 mg of wet nitrocellulose was placed in a pre-weighed, 50 ml polypropylene digestion vial and heated at 100 °C oven for 1 hour to dry the sample. After cooling to room temperature, the vial with dried NC sample was weighed to get a final sample weight. Five (5) mL of 5 M NaOH was added to the vial and sealed with a screw cap. The vial was placed in a heating block set to 75 °C, and incubated for 2.5 hours. The vial was then transferred to an ice bath for 15 minutes to quickly cool the reaction mixture and stop the hydrolysis reaction. After a return to room temperature, the solution was brought to a final volume of 50 mL using ultrapure deionized water.

Ion chromatography analysis was performed using a Dionex ICS-2500 High Performance Ion Chromatograph (with eluent generator) equipped with a Dionex AS18 analytical column coupled to a Dionex AG18 guard column. The flow rate was 0.8 mL/min with a mobile phase of 23 mM hydroxide solution. Samples were injected onto the column using a 25 µL sample loop. The background was suppressed electrolytically using an anionic self-regenerating suppressor (57 mA) and ultrappure deionized water at a flow rate of 4 mL/min. Analyte peaks (nitrite and nitrate) were detected using a conductivity detector. The chromatograph interfaced to a Dell GX260 Pentium 4 computer using Windows XP and Chromeleon software (Version 6.50) for data acquisition and manipulation. Appropriate dilutions were made to reduce any high concentration analytes down to within the range of the calibration standards.

Table 2.6-3 shows %N results obtained recently using the draft IC method to analyze two commercial samples whose %N is known, and one sample of BNC recently prepared. The TG-1927-52-1 was nitrated using optimal conditions as published by Sun et al. (63) which reported synthesis of a BNC with 13.75 %N.

Most of the data are reasonable, with the exceptions being the two ATK values of 14.58 %N and 14.16 %N, almost two percentage point higher than the actual value. Values lower than the certified values were expected, since digestion and conversion of all nitro esters in NC to nitrate and nitrite is not always complete. Furthermore, these values exceed the theoretical limit for fully nitrated NC, which is 14.14 %N.

While there are many reasons that a lower than expected %N value might be observed only a few possible explanations for higher than expected values have been identified so far. They are:

1) A weighing error resulted in a lower than actual starting mass. The weighing was repeated several times and consistent numbers were obtained, and the balance was checked and found to be accurate ruling this explanation out.

2) A nitrate salt impurity or additive that was not accounted for in the certified analysis value. This was minimized by thoroughly washing the NC twice in boiling water. The resulting %N value was still higher than the theoretical limit. If there is some other material present with higher nitrogen level it was apparently not water soluble.

The digestion ion chromatography method looks promising, provided care is taken to minimize all possible artifacts.

| Sample | Certified %N | Mass (mg) | %N $NO_3$ | %N $NO_2$ | $NO_2$:$NO_3$ Ratio | Measured %N | % Yield |
|---|---|---|---|---|---|---|---|
| RS30-35 ATK | 12.57 | 55.6 | 5.52 | 9.06 | 0.609 | 14.58 | 116 |
| RS30-35 ATK (washed) | 12.57 | 51.5 | 5.30 | 8.86 | 0.598 | 14.16 | 113 |
| RS30-35 ATK | 12.57 | 60.0 | 4.71 | 7.89 | 0.597 | 12.60 | 100 |
| | | | | | | | |
| RS55 (Hercules) | 12.19 | 59.4 | 4.35 | 6.58 | 0.661 | 10.93 | 90 |
| | | | | | | | |
| BNC TG-1927-52-1 | | 53.1 | 4.40 | 8.96 | 0.491 | 13.36 | 97* |
| BNC TG-1927-52-1 | | 44.1 | 4.47 | 9.21 | 0.485 | 13.68 | 99* |
| BNC TG-1927-52-1 | | 46.2 | 4.07 | 8.45 | 0.482 | 12.52 | 91* |

*Based on an expected 13.75% N

**Table 2.6-3. Nitrogen content of NC standards and BNC using ion chromatography.**

### *2.6.3 Conclusions*

This task clearly demonstrated that nitrocellulose can be generated from biocellulose. Initial adjustments to the standard nitration procedure were evaluated, but further optimization, including scale-up, need additional effort.

Additionally, less labor intensive, non-destructive methods for assessing nitrogen content in nitrocellulose were also examined. Again, further work is needed to optimize and validate these methods for routine use during nitrocellulose characterization.

## 3.0 CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

### 3.1 Key results

The key results of this project were as follows:

- One of the first reported inducible promoter systems for *Gluconacetobacter* spp. was identified and validated.
- Isoforms of the key biocellulose synthesis enzyme AcsA that were insensitive to the metabolic control molecule CDG were engineered.
- A method to use attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy to detect bacterial impurities (e.g., nucleic acids, lipids, protein) in biocellulose was developed.
- Detailed metabolic models of the cellulose producing strains used in this research were developed and used to identify key genes/pathways that are predicted to increase biocellulose production.
- Proof-of-concept synthesis of nitrocellulose from biocellulose was successfully performed, and preliminary characterization indicated that the material was comparable to military grade nitrocellulose with respect to FTIR spectra and nitrogen content.

The identification of the inducible promoter system opens the way for constructing a controllable genetic expression system, and gaining control over the synthesis of biocellulose by these strains. If and when the remaining hurdles are cleared, these genetic tools, combined with the CDG-insensitive AcsA isoforms and other genetic changes, are expected to allow measurable improvements in efficiency and yield.

### 3.2 Immediate implications

The results of this project have the following nearterm applications or implications:

- The development and publication of the genome sequences and metabolic models (Section 2.4) for the two biocellulose producing strains should prove useful to the broader biocellulose research community in efforts to improve biocellulose production.
- The development and validation of the FTIR BC impurity analysis (Section 2.5.3) should be applicable to the broader area of biocellulose research.

### 3.3 Suggested follow-on work

The results of this project do not support the immediate application of the BC to NC process at this time, nor any scale up of other biocellulose-related processes or products. The challenge faced by this project, as well as the larger research community, is the inability to improve biocellulose production by *G. xylinus* through genetic manipulation, and specifically, a controllable genetic expression system.

Most of the parts required to build controllable genetic expression system have been identified and validated, either through this project or other research efforts (e.g., reliable antibiotic markers, origins of replication, and reporter genes). The challenge flows from the lack of a simple, tightly regulated, strongly expressed promoter or suit of promoters. The absence of such prevents some of the most basic questions from being answered (e.g., what is the effect of gene X on biocellulose production when it is expressed or not expressed).

Such a promoter could be either an as of yet undiscovered or untried promoter system, or an improved version of vanillate operon discovered and validated during this project.

The steps recommended to move this concept forward would be (in order of priority):

- Use the genetic elements validated during this project to create a promoter probe system to interrogate known promoter systems and putative promoters identified in the genome of *G. xylinus*.
- Apply the tools of synthetic biology to create improved versions of the *G. xylinus* vanillate promoter that is more tightly regulated (less basal expression in the absence of an inducer).
- Generate and evaluate the biocellulose production of strains expressing the wild-type and CDG-insentive AcsA enzymes.
- Implement the genetic modifications identified by the metabolic modeling analysis and confirm the positive effects of these modifications on biocellulose production.
- Optimize fermentation to demonstrate BC production at the bench-scale.
- Specifically for BC to NC, further develop BC processing methods prior to nitration, including physical attrition combined with the use of ionic solvents as the nitration medium.

## 4.0 REFERENCES CITED


1. Alinat, E., N. Delaunay, C. Costanza, X. Archer, and P. Gareil. 2014. Determination of the nitrogen content of nitrocellulose by capillary electrophoresis after alkaline denitration. Talanta 125: 174-180.

2. Antoine, R., and C. Locht. 1992. Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from Gram-positive organisms. Molecular Microbiology 6(13): 1785-1799.

3. Baldwin, T. O., J. H. Devine, R. C. Heckel, J. W. Lin, and G. S. Shadel. 1989. The complete nucleotide sequence of the *lux* regulon of *Vibrio fischeri* and the *luxABN* region of *Photobacterium leiognathi* and the mechanism of control of bacterial bioluminescence. Journal of Bioluminescence and Chemiluminescence 4(1): 326-341.

4. Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, and P. A. Pevzner. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455-477.

5. Battad-Bernardo, E., S. L. McCrindle, I. Couperwhite, and B. A. Neilan. 2004. Insertion of an *E. coli lacZ* gene in *Acetobacter xylinus* for the production of cellulose in whey. FEMS Microbiology Letters 231(2): 253-260.

6. Bertalan, M., R. Albano, V. de Pádua, L. Rouws, C. Rojas, A. Hemerly, K. Teixeira, S. Schwab, J. Araujo, A. Oliveira, L. França, V. Magalhães, S. Alquéres, A. Cardoso, W. Almeida, M. M. Loureiro, E. Nogueira, D. Cidade, D. Oliveira, T. Simão, J. Macedo, A. Valadão, M. Dreschsel, F. Freitas, M. Vidal, H. Guedes, E. Rodrigues, C. Meneses, P. Brioso, L. Pozzer, D. Figueiredo, H. Montano, J. Junior, G. de Souza Filho, V. Martin Quintana Flores, B. Ferreira, A. Branco, P. Gonzalez, H. Guillobel, M. Lemos, L. Seibel, J. Macedo, M. Alves-Ferreira, G. Sachetto-Martins, A. Coelho, E. Santos, G. Amaral, A. Neves, A. B. Pacheco, D. Carvalho, L. Lery, P. Bisch, S. C. Rössle, T. Ürményi, A. Rael Pereira, R. Silva, E. Rondinelli, W. von Krüger, O. Martins, J. I. Baldani, and P. C. G. Ferreira. 2009. Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. BMC Genomics 10(1): 450.

7. Blatny, J. M., T. Brautaset, H. C. Winther-Larsen, K. Haugan, and S. Valla. 1997. Construction and use of a versatile set of broad-host-range cloning and expression vectors based on the RK2 replicon. Applied and Environmental Microbiology 63(2): 370-9.

8. Chakrabarty, A. M., G. Chou, and I. C. Gunsalus. 1973. Genetic regulation of octane dissimilation plasmid in *Pseudomonas*. Proceedings of the National Academy of Sciences of the United States of America 70(4): 1137-1140.

9. Chang, A. L., J. R. Tuckerman, G. Gonzalez, R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, and M.-A. Gilles-Gonzalez. 2001. Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor. Biochemistry 40(12): 3420-3426.

10. Cheng, K.-C., J. Catchmark, and A. Demirci. 2009. Enhanced production of bacterial cellulose by using a biofilm reactor and its material property analysis. Journal of Biological Engineering 3(1): 12.

11. Chien, L.-J., H.-T. Chen, P.-F. Yang, and C.-K. Lee. 2006. Enhancement of cellulose pellicle production by constitutively expressing *Vitreoscilla* hemoglobin in *Acetobacter xylinum*. Biotechnology Progress 22(6): 1598-1603.

12. Christodoulatos, C., T.-L. Su, and A. Koutsospyros. 2001. Kinetics of the alkaline hydrolysis of nitrocellulose. Water Environment Research 73(2): 185-191.

13. Clarkson, A., and C. M. Roberston. 1966. Refined calculation for determination of nitrogen in nitrocellulose by infrared spectrometry. Analytical Chemistry 38(3): 522-522.

14. de Lorenzo, V., L. Eltis, B. Kessler, and K. N. Timmis. 1993. Analysis of *Pseudomonas* gene products using lacIq/Ptrp-lac plasmids and transposons that confer conditional phenotypes. Gene 123(1): 17-24.

15. DiMarco, A. A., B. Averhoff, and L. N. Ornston. 1993. Identification of the transcriptional activator pobR and characterization of its role in the expression of *pobA*, the structural gene for p-hydroxybenzoate hydroxylase in *Acinetobacter calcoaceticus*. Journal of Bacteriology 175(14): 4499-4506.

16. Edwards, K. J., A. J. Jay, I. J. Colquhoun, V. J. Morris, M. J. Gasson, and A. M. Griffin. 1999. Generation of a novel polysaccharide by inactivation of the *aceP* gene from the acetan biosynthetic pathway in *Acetobacter xylinum*. Microbiology 145(6): 1499-1506.

17. Fedoroff, B. T., and O. E. Sheffield. 1962. Encyclopedia of Explosives And Related Items, PART 2700. U.S. Army Research and Development Command; TACOM, ARDEC; Warheads, Energetics and Combat Support Center, Picatinny Arsenal, NJ, USA.

18. Florea, M., H. Hagemann, G. Santosa, Abbott, C. N. Micklem, X. Spencer-Milnes, L. de Arroyo Garcia, D. Paschou, C. Lazenbatt, D. Kong, H. Chughtai, K. Jensen, P. S. Freemont, R. Kitney, B. Reeve, and T. Ellis. 2016. Engineering control of bacterial cellulose production using a genetic toolkit and a new cellulose-producing strain. Proceedings of the National Academy of Sciences 113(24): E3431-E344.

19. Florea, M., B. Reeve, J. Abbott, P. S. Freemont, and T. Ellis. 2016. Genome sequence and plasmid transformation of the model high-yield bacterial cellulose producer *Gluconacetobacter hansenii* ATCC 53582. Scientific Reports 6: 23635.

20. Gensh, K. V., P. V. Kolosov, and N. G. Bazarnova. 2011. Quantitative analysis of cellulose nitrates by Fourier transform infrared spectroscopy. Russian Journal of Bioorganic Chemistry 37(7): 814-816.

21. Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose $P_{BAD}$ promoter. Journal of Bacteriology 177(14): 4121-30.

22. Hall, P. E., S. M. Anderson, D. M. Johnston, and R. E. Cannon. 1992. Transformation of Acetobacter xylinum with plasmid DNA by electroporation. Plasmid 28(3): 194-200.

23. Hareland, W. A., R. L. Crawford, P. J. Chapman, and S. Dagley. 1975. Metabolic function and properties of 4-hydroxyphenylacetic acid 1-hydroxylase from *Pseudomonas acidovorans*. Journal of Bacteriology 121(1): 272-285.

24. Hungund, B. S., and S. G. Gupta. 2010. Improved production of bacterial cellulose from *Gluconacetobacter persimmonis* GH-2. Journal of Microbial & Biochemical Technology 2: 127-133.

25. Hwang, J. W., Y. K. Yang, J. K. Hwang, Y. R. Pyun, and Y. S. Kim. 1999. Effects of pH and dissolved oxygen on cellulose production by Acetobacter xylinum BRC5 in agitated culture. Journal of Bioscience and Bioengineering 88(2): 183-188.

26. Iniesta, J., M. D. Esclapez-Vicente, J. Heptinstall, D. J. Walton, I. R. Peterson, V. A. Mikhailov, and H. J. Cooper. 20110. Retention of enzyme activity with a boron-doped diamond electrode in the electro-oxidative nitration of lysozyme. Enzyme and Microbial Technology 46(6): 472-478.

27. Ishida, T., Y. Sugano, T. Nakai, and M. Shoda. 2002. Effects of acetan on production of bacterial cellulose by *Acetobacter xylinum*. Bioscience, Biotechnology, and Biochemistry 66(8): 1677-1681.

28. Iyer, P. R., S. M. Geib, J. M. Catchmark, T.-h. Kao, and M. Tien. 2010. Genome sequence of a cellulose-producing bacterium, *Gluconacetobacter hansenii* ATCC 23769. Journal of Bacteriology 192(16): 4256-4257.

29. Kaczmarczyk, A., J. A. Vorholt, and A. Francez-Charlot. 2014. Synthetic vanillate-regulated promoter for graded gene expression in *Sphingomonas*. Scientific Reports 4: 6453.

30. Kanehisa, M., and S. Goto. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1): 27-30.

31. Keshk, S. M. A. S., T. M. A. Razek, and K. Sameshima. 2006. Bacterial cellulose production from beet molasses. African Journal of Biotechnology 5(17): 1519-1523.

32. Keshk, S. M. A. S., and K. Sameshima. 2005. Evaluation of different carbon sources for bacterial cellulose production. African Journal of Biotechnology 4(6): 478-482.

33. Kim, J., J. L. Reed, and C. T. Maravelias. 2011. Large-scale bi-level strain design approaches and mixed-integer programming solution techniques. PLoS One 6(9): e24162.

34. Ko, J., K.-S. Ryu, H.-S. Kim, J.-S. Shin, J.-O. Lee, C. Cheong, and B.-S. Choi. 2010. Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by Pilz domain proteins. Journal of Molecular Biology 398(1): 97-110.

35. Krystynowicz, A., W. Czaja, A. Wiktorowska-Jezierska, M. Gonçalves-Miśkiewicz, M. Turkiewicz, and S. Bielecki. 2002. Factors affecting the yield and properties of bacterial cellulose. Journal of Industrial Microbiology & Biotechnology 29(4): 189-195.

36. Kubiak, K., M. Kurzawa, M. Jędrzejczak-Krzepkowska, K. Ludwicka, M. Krawczyk, A. Migdalski, M. M. Kacprzak, D. Loska, A. Krystynowicz, and S. Bielecki. 2014. Complete genome sequence of *Gluconacetobacter xylinus* E25 strain—Valuable and effective producer of bacterial nanocellulose. Journal of Biotechnology 176: 18-19.

37. Kuo, C.-H., H.-Y. Teng, and C.-K. Lee. 2015. Knock-out of glucose dehydrogenase gene in *Gluconacetobacter xylinus* for bacterial cellulose production enhancement. Biotechnology and Bioprocess Engineering 20(1): 18-25.

38. Kurosumi, A., C. Sasaki, Y. Yamashita, and Y. Nakamura. 2009. Utilization of various fruit juices as carbon source for production of bacterial cellulose by Acetobacter xylinum NBRC 13693. Carbohydrate Polymers 76(2): 333-335.

39. Labes, M., A. Puhler, and R. Simon. 1990. A new family of RSF1010-derived expression and lac-fusion broad-host-range vectors for gram-negative bacteria. Gene 89(1): 37-46.

40. Lesuisse, E., K. Schanck, and C. Colson. 1993. Purification and preliminary characterization of the extracellular lipase of *Bacillus subtilis* 168, an extremely basic pH-tolerant enzyme. European Journal of Biochemistry 216(1): 155-160.

41. López-López, M., J. M. R. Alegre, C. García-Ruiz, and M. Torre. 2011. Determination of the nitrogen content of nitrocellulose from smokeless gunpowders and collodions by alkaline hydrolysis and ion chromatography. Analytica Chimica Acta 685(2): 196-203.

42. Mermod, N., J. L. Ramos, P. R. Lehrbach, and K. N. Timmis. 1986. Vector for regulated expression of cloned genes in a wide range of gram-negative bacteria. Journal of Bacteriology 167(2): 447-454.

43. Meyer, R., D. Figurski, and H. D.R. 1975. Molecular vehicle properties of the broad host range plasmid RK2. Science 190(4220): 1226-1228.

44. Moosavi-Nasab, M., and A. Yousefi. 2011. Biotechnological production of cellulose by *Gluconacetobacter xylinus* from agricultural waste. Iranian Journal of Biotechnology 9(2)

45. Morgan, J. L. W., J. T. McNamara, and J. Zimmer. 2014. Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nature Structural & Molecular Biology 21(5): 489-496.

46. Nakai, T., N. Tonouchi, T. Konishi, Y. Kojima, T. Tsuchida, F. Yoshinaga, F. Sakai, and T. Hayashi. 1999. Enhancement of cellulose production by expression of sucrose synthase in *Acetobacter xylinum*. Proceedings of the National Academy of Sciences 96(1): 14-18.

47. Newman, J. R., and C. Fuqua. 1999. Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. Gene 227(2): 197-203.

48. Norwitz, G., and D. E. Chasan. January 1972. Infrared determination of nitrogen in raw nitrocellulose and nitrocellulose contained in propellants on a macro and semimicro scale. Department of the Army, Frankford Arsenal. Report# Test Report T72-3-1.

49. Ogino, H., Y. Azuma, A. Hosoyama, H. Nakazawa, M. Matsutani, A. Hasegawa, K.-i. Otsuyama, K. Matsushita, N. Fujita, and M. Shirai. 2011. Complete genome sequence of NBRC 3288, a unique cellulose-nonproducing strain of *Gluconacetobacter xylinus* isolated from vinegar. Journal of Bacteriology 193(24): 6997-6998.

50. Okiyama, A., H. Shirae, H. Kano, and S. Yamanaka. 1992. Bacterial cellulose I. Two-stage fermentation process for cellulose production by Acetobacter aceti. Food Hydrocolloids 6(5): 471-477.

51. Ormö, M., A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. 1996. Crystal structure of the *Aequorea victoria* green fluorescent protein. Science 273(5280): 1392-1395.

52. Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research

53. Pourramezan, G. Z., A. M. Roayaei, and Q. R. Qezelbash. 2009. Optimization of culture conditions for bacterial cellulose production by Acetobacter sp. 4B-2. Biotechnology Progress 8: 150-154.

54. Pultz, I. S., M. Christen, H. D. Kulasekara, A. Kennard, B. Kulasekara, and S. I. Miller. 2012. The response threshold of *Salmonella* PilZ domain proteins is determined by their binding affinities for c-di-GMP. Molecular Microbiology 86(6): 1424-1440.

55. Ross, P., R. Mayer, and M. Benziman. 1991. Cellulose biosynthesis and function in bacteria. Microbiology and Molecular Biology Reviews 55(1): 35-58.

56. Schäper, S., E. Krol, D. Skotnicka, V. Kaever, R. Hilker, L. Søgaard-Andersen, and A. Becker. 2016. Cyclic di-GMP regulates multiple cellular functions in the symbiotic Alphaproteobacterium *Sinorhizobium meliloti*. Journal of Bacteriology 198(3): 521-535.

57. Schmidt, A. J., D. A. Ryjenkov, and M. Gomelsky. 2005. The ubiquitous protein domain EAL Is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL Domains. Journal of Bacteriology 187(14): 4774-4781.

58. Schramm, M., and S. Hestrin. 1954. Factors affecting production of cellulose at the air/liquid interface of a culture of *Acetobacter xylinum*. Microbiology 11(1): 123-129.

59. Shigematsu, T., K. Takamine, M. Kitazato, T. Morita, T. Naritomi, S. Morimura, and K. Kida. 2005. Cellulose production from glucose using a glucose dehydrogenase gene (*gdh*)-deficient mutant of *Gluconacetobacter xylinus* and its use for bioconversion of sweet potato pulp. Journal of Bioscience and Bioengineering 99(4): 415-422.

60. Soh, C.-K., A.-K. Soh, and K.-Y. Lai. 1989. KBase: A customizable tool for building DBase-compatible knowledge-based systems. Advances in Engineering Software (1978) 11(3): 136-148.

61. Strand, T. A., R. Lale, K. F. Degnes, M. Lando, and S. Valla. 2014. A new and improved host-independent plasmid system for RK2-based conjugal transfer. PLoS One 9(3): e90372.

62. Sukchawalit, R., P. Vattanaviboon, R. Sallabhan, and S. Mongkolsuk. 1999. Construction and characterization of regulated L-arabinose-inducible broad host range expression vectors in *Xanthomonas*. FEMS Microbiology Letters 181(2): 217-223.

63. Sun, D.-P., B. Ma, C.-L. Zhu, C.-S. Liu, and J.-Z. Yang. 2010. Novel nitrocellulose made from bacterial cellulose. Journal of Energetic Materials 28(2): 85-97.

64. Tal, R., H. C. Wong, R. Calhoon, D. Gelfand, A. L. Fear, G. Volman, R. Mayer, P. Ross, D. Amikam, H. Weinhouse, A. Cohen, S. Sapir, P. Ohana, and M. Benziman. 1998. Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: Genetic organization and pccurrence of conserved domains in isoenzymes. Journal of Bacteriology 180(17): 4416-4425.

65. Thanbichler, M., A. A. Iniesta, and L. Shapiro. 2007. A comprehensive set of plasmids for vanillate- and xylose-inducible gene expression in *Caulobacter crescentus*. Nucleic Acids Research 35(20): e137.

66. Tsuchida, T., and F. Yoshinaga. 1997. Production of bacterial cellulose by agitation culture systems. Pure & Applied Chemistry 69(11): 2453-2458.

67. Verschuren, P. G., T. D. Cardona, M. J. R. Nout, K. D. De Gooijer, and J. C. Van den Heuvel. 2000. Location and limitation of cellulose production by *Acetobacter xylinum* established from oxygen profiles. Journal of Bioscience and Bioengineering 89(5): 414-419.

68. Whited, G. M., and D. T. Gibson. 1991. Toluene-4-monooxygenase, a three-component enzyme system that catalyzes the oxidation of toluene to p-cresol in *Pseudomonas mendocina* KR1. Journal of Bacteriology 173(9): 3010-3016.

69. Worsey, M. J., F. C. Franklin, and P. A. Williams. 1978. Regulation of the degradative pathway enzymes coded for by the TOL plasmid (pWWO) from *Pseudomonas putida* mt-2. Journal of Bacteriology 134(3): 757-764.

70. Wu, T. K. 1980. Carbon-13 and proton nuclear magnetic resonance studies of cellulose nitrates. Macromolecules 13(1): 74-79.

71. Zhang, Y., X. Shang, S. Lai, G. Zhang, Y. Liang, and T. Wen. 2012. Development and application of an arabinose-inducible expression system by facilitating inducer uptake in *Corynebacterium glutamicum*. Applied and Environmental Microbiology 78(16): 5831-5838.

72. Zhong, C., F. Li, M. Liu, X.-N. Yang, H.-X. Zhu, Y. Y. Jia, S.-R. Jia, and L. Piergiovanni. 2014. Revealing differences in metabolic flux distributions between a mutant strain and its parent strain *Gluconacetobacter xylinus* CGMCC 2955. PLoS One

73. Zhong, C., G.-C. Zhang, M. Liu, X.-T. Zheng, P.-P. Han, and S.-R. Jia. 2013. Metabolic flux analysis of *Gluconacetobacter xylinus* for bacterial cellulose production. Applied Microbiology and Biotechnology 97(14): 6189-6199.

## 5.0 APPENDICES

A. Other Supporting Data – N/A

B. List of Scientific/Technical Publications

Articles in peer-reviewed journals (planned)

Fuller, M.E., McClay, K.R., Andaya, C., Kumar P., and Reed, J.L. Development of an inducible genetic expression system for biocellulose producing *Gluconacetobacter xylinus* strain Gx 399 (in preparation; journal TBD).

Kumar P., McClay K., Fuller M.E., and Reed, J.L. Complete genome sequence of biocellulose producing bacteria *Gluconacetobacter xylinus* ATCC 53582 and 399 (in preparation for submission to Genome Announcements).

Kumar P., Pan S., McClay K., Zhang X., Fuller M.E., and Reed, J.L. Strategies to enhance the yield of biocellulose in *Gluconacetobacter xylinus* ATCC 53582 and 399 (in preparation for submission to Biotechnology & Bioengineering or BMC Systems Biology).

Conference or symposium abstracts

Kumar, P., K. McClay, M. Fuller, and J. Reed. 2016. Strategies to enhance the yield of cellulose in *Gluconacetobacter xylinus*. 2016 Synthetic Biology: Engineering, Evolution & Design (SEED). Chicago, IL, USA, July 18-21.

*Gluconacetobacter xylinu*s is known to produce high quality biocellulose. The biocellulose obtained from *Gluconacetobacter xylinus* is devoid of the cell wall components which makes it very pure. This highly pure cellulose can be used for many applications including medical and defense applications. This project seeks to enhance biocellulose yields from *Gluconacetobacter xylinus* strains. Optimizing metabolic and regulatory changes to enhance chemical production is a time-consuming and expensive process. However, this process can be expedited using a genome-scale metabolic model. Metabolic models can evaluate multiple strain design strategies to identify those with the greatest strain improvements. We sequenced the genome of *Gluconacetobacter xylinus* ATCC 53582 and used this information to build a genome-scale metabolic model of this strain. The growth predictions of the genome-scale metabolic model were consistent with the experimental growth observations performed using Biolog phenotype microarrays. We applied the SimOptStrain algorithm to the metabolic model to identify gene deletion and reaction addition strategies to enhance the yield of bio-cellulose from glucose in the minimal media. The proposed genetic changes are predicted to enhance the biocellulose yields by almost three fold compared to the yields reported for the wildtype strain. Future work will focus on implementing and validating the effects of the metabolic changes experimentally.