

Characterizing Uronate Dehydrogenase from *P. syringae* as a Step Towards Expressing a Retro-Biosynthetically Derived Glucaric Acid Pathway in *E. coli*

By

Pooya Iranpour

B.S. Chemical Engineering (2003)
The University of Texas at Austin

Submitted to the Department of Chemical Engineering in Partial Fulfillment of the Requirements for the Degree of

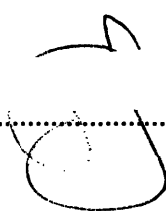
Master of Science in Chemical Engineering

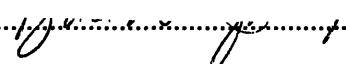
at the


Massachusetts Institute of Technology

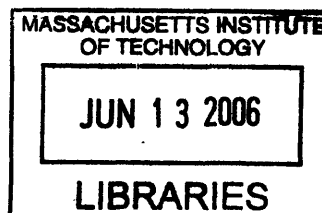
[June 2006]
May 2006

copyright 2006 Massachusetts Institute of Technology
All rights reserved

Signature of Author..........
Department of Chemical Engineering
May 12, 2006

Certified by..........
Joseph R. Mares (1924) Assistant Professor of Chemical Engineering
Thesis Supervisor
Kristala Jones Prather

Accepted by..........
William M. Deen
Chairman, Committee for Graduate Students
Department of Chemical Engineering



ARCHIVES

Characterizing Uronate Dehydrogenase from *P. syringae* as a Step Towards Expressing a Retro-Biosynthetically Derived Glucaric Acid Pathway in *E. coli*

By

Pooya Iranpour

Submitted to the Department of Chemical Engineering on May 12, 2006 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Chemical Engineering

Abstract

Retro-biosynthesis is a novel pathway design method that integrates biocatalysis (in particular, functional group specificity and reactivity of enzymes) and organic chemistry. Using this method, several novel pathways have been designed in this laboratory for the microbial production of glucaric acid. In addition to design, efforts have been made in order to express these pathways inside *E. coli*. For the production of glucaric acid in *E. coli*, toxicity of this organic acid as well as the uptake efficiency of glucose, the starting substrate, when in the presence of glucaric acid was assessed. Furthermore, a benchmark pathway for glucaric acid production that utilizes known reaction steps from different source organisms was designed. The design and expression of this pathway will be compared to the retro-biosynthetic pathway. The ease of design and expression will be analyzed in order to confirm the validity of retro-biosynthesis.

Toxicity results showed that glucaric acid concentrations, up to 40 mM, do not affect growth of *E. coli* DH10B cells. Furthermore, cells take up equal amounts of glucose even when glucaric acid is available, indicating that glucose is preferentially utilized by cells and that glucaric acid, once produced by cells, would not affect further production. Finally, efforts have been made in order to express the benchmark pathway in *E. coli*. The entire pathway requires expression of three enzymes. The final reaction step requires uronate dehydrogenase in order to convert glucuronic acid to the desired product, glucaric acid. Although the function and approximate molecular weight of uronate dehydrogenase is known, its DNA sequence is not. Isolating the gene encoding the enzyme can be facilitated by a partial digestion of the genomic DNA of the source organism, *P. syringae*. A screening method for uronate dehydrogenase activity has also been developed and optimized. Using this screening method, the library was constructed in *E. coli*. Once colonies that contain uronate dehydrogenase are selected, this enzyme's sequence can be determined and the benchmark pathway can be closer to becoming expressed.

Thesis Supervisor: Kristala Jones Prather

Title: Joseph R. Mares (1924) Assistant Professor of Chemical Engineering

TABLE OF CONTENTS

1. Introduction	8
1.1 The benchmark pathway	9
1.2 Expressing the final reaction step of the benchmark pathway in <i>E. coli</i>	13
2. Materials and Methods	15
2.1 Glucaric acid toxicity: specific growth rate determination for cells grown in varying concentrations of glucaric acid	15
2.2 Glucose uptake efficiency in the presence of glucaric acid.....	16
2.3 Preparing <i>P. syringae</i> lysate and testing for uronate dehydrogenase activity	17
2.4 Optimizing screening method	19
2.5 Testing screening method for DH10B-pGEM3Z and DH10B-pBluescript using SOC transfer method.....	20
2.6 Isolating genomic DNA of <i>P. syringae</i> DC3000 for partial digestion	20
2.7 Genomic DNA library construction.....	21
3. Results and Discussion	25
3.1 Glucaric acid toxicity	25
3.2 Testing <i>P. syringae</i> DC3000 for uronate dehydrogenase activity <i>in vitro</i>	29
3.3 Screening method for detecting uronate dehydrogenase activity in <i>E. coli</i>	31
3.4 Genomic DNA library construction (cloning uronate dehydrogenase)	39
4. Conclusions and Future Directions	45

LIST OF FIGURES

Figure 1: Benchmark Pathway for Glucaric Acid.....	11
Figure 2: Final Step of the Benchmark Pathway.....	13
Figure 3: Growth Curve and Specific Growth Rate Determination for DH10B Cells in Glucaric Acid	26
Figure 4: Uptake Efficiency of Glucose in the Presence of Glucaric Acid	28
Figure 5: Measuring Uronate Dehydrogenase Activity.....	30
Figure 6: Excision of a 1-5kb Range of Fragments from a Partial Digestion of the Genomic DNA of <i>P. syringae</i>.....	40
Figure 7: Obtaining Fragments from the Entire Partial Digestion of the Genomic DNA of <i>P. syringae</i>.....	43
Figure 8: Determining Insert Sizes from Successful Clones	44

LIST OF TABLES

Table 1: Growth of BL21(DE3) in Liquid Media Containing Organic Acids.....	33
Table 2: Growth of <i>E. coli</i> DH10B in Liquid and on Solid Media Containing Organic Acids.....	33
Table 3: Growth of DH5alpha and MG1655 in Liquid Media Containing Organic Acids.....	35
Table 4: Growth of DH10B-pGEM3Z in Liquid and on Solid Media Containing Organic Acids.....	36
Table 5: Growth of DH10B-pGEM3Z and DH10B-pBluescript on Solid Media Containing Organic Acids.....	38

1. Introduction

As more organic compounds are synthesized and identified as beneficial, new methods are being determined for their large scale and inexpensive production. Specifically, and for the interest of our laboratory, new pathway design tools and pathway expression/engineering (protein and metabolic engineering) techniques are being used for the production of these organic compounds in *E. coli*.

Some pathway design methods include using metabolic network and enzyme databases such as KEGG (1) and BRENDA (2), respectively. Computational tools that utilize graph theory integrated with group contribution theory can also identify alternative, thermodynamically favorable routes for the expression of a desired product (3). Pathway expression techniques include enzyme evolution and directed mutagenesis, gene/pathway insertion and deletion, as well as the modulation of enzyme expression or substrate type and availability. Our laboratory is applying a new pathway design method, retro-biosynthesis, in order to produce natural or un-natural organic compounds in *E. coli*. Once designed, the pathway will then be expressed in this microorganism by considering the many pathway engineering tools mentioned above.

This new design tool has the benefit of allowing one to control the starting substrate, making sure *E. coli* can either utilize or make it in large-scale, inexpensive quantities. Furthermore, unlike other pathway design methods, retro-biosynthesis is systematic and requires every step to be tested and controlled. As a result, desired products are made rather than random general structures like in previous pathway engineering efforts (the microbial production of isoprenoids, 4). Pathways are designed in

a more efficient and flexible manner as well. Enzymes are chosen for their functional group specificity and conversion rather than substrate specificity.

Retro-biosynthesis integrates knowledge of enzyme function, organic chemistry, and metabolism. Once a desired organic compound is identified, its structure is first compared to a substrate *E. coli* can make and/or utilize. Then, enzymes are analyzed for their ability to be selective towards functional groups and to perform the conversions needed in order to generate the product. Once the biosynthetic pathway is constructed in this way, expression optimization methods and viability considerations must be made in order to express the pathway inside *E. coli*. Retro-biosynthesis will be used in order to design a novel biosynthetic pathway for the microbial production of glucaric acid. However, a benchmark pathway was first designed using enzyme and reaction databases as discussed above.

1.1 The benchmark pathway

The benchmark pathway for the biosynthesis of glucaric acid was designed and will be compared to the pathways proposed via retro-biosynthesis. It was designed in order to express an already “proven” pathway inside *E. coli* and determine whether glucaric acid can be produced. The pathway is proven because every reaction step is known to occur even though all are not from the same source organism. While retro-biosynthesis considers general enzyme function (at the functional group level), the benchmark pathway considers specific enzyme reactions. Each step has already been identified in individual organisms. The required enzymes and their functions are displayed in enzyme/reaction databases. Step by step, they are retrieved and isolated from

the organisms in which they function and assembled to produce the desired product: glucaric acid.

Retro-biosynthetically derived pathways would consider enzymes' functional group specificity instead. The reaction steps are not necessarily natural steps already known to occur in organisms. Therefore, although easier to design, expressing the pathway inside *E. coli* can be more of a challenge. Enzymes may need to be modified in order to be substrate specific. Furthermore, new assays may need to be developed in order to prove enzyme function as well as measure and quantify product formation in *E. coli*.

The design method of retro-biosynthesis, and the work of having to identify/modify unproven reaction steps, can be compared to expressing naturally occurring reaction steps in *E. coli* by first expressing the benchmark pathway (Figure 1).

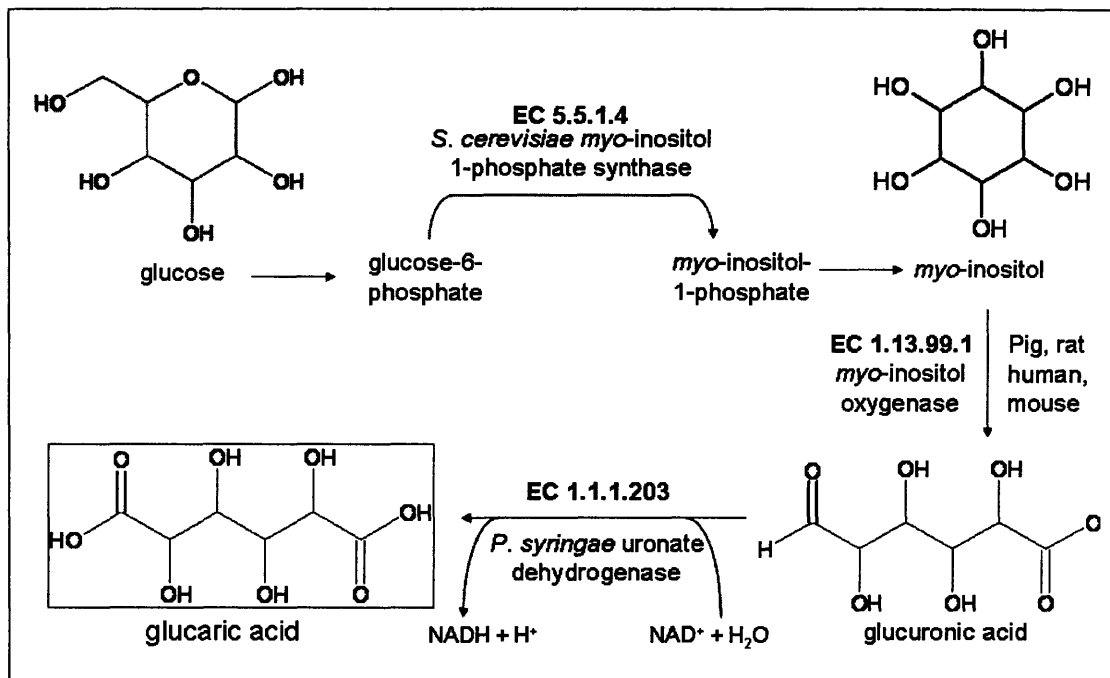


Figure 1: Benchmark Pathway for Glucaric Acid

Each step is already known to occur in specific organisms and each enzyme is substrate specific rather than functional group specific. The first three steps can occur by cloning just one enzyme (EC 5.5.1.4, 5). *Myo*-inositol oxygenase, found in pig, rat, human, and mouse, can convert *myo*-inositol to glucuronic acid (6). Uronate dehydrogenase, reportedly from *P. syringae*, has been shown to convert glucuronic acid to glucaric acid (7, 8, 9).

It may be possible to produce glucaric acid by expressing three already verified enzymes, obtained from separate organisms, inside *E. coli* that is grown in glucose (the starting substrate). The first three steps have already been demonstrated in *E. coli* (5). According to Hansen *et al.*, *myo*-inositol can be produced by *E. coli* after isolating and expressing *myo*-inositol 1-phosphate synthase from *S. cerevisiae*. Glucose is phosphorylated to glucose-6-phosphate (using *E. coli*'s phosphotransferase system and phosphoenolpyruvate as the source of phosphate). Glucose-6-phosphate can be converted to *myo*-inositol phosphate using *myo*-inositol 1-phosphate synthase. This substrate can then be dephosphorylated and has been shown to do so by "unknown

cytosolic or periplasmic phosphatase activity” (5). Then, *myo*-inositol oxygenase, found in humans, mice, rats, and pigs, can be used in order to convert *myo*-inositol to glucuronic acid (6). The final reaction step requires uronate dehydrogenase, which is known to occur in *P. syringae* (7, 8, 9). Once this enzyme functions properly in *E. coli* and has glucuronic acid as substrate via the other steps of the pathway, glucaric acid can then be produced.

Many considerations must be made before the benchmark pathway can be successfully expressed in *E. coli*. If product is available, toxicity and uptake efficiency studies can be performed in order to determine whether producing the desired organic compound will limit the cell’s viability or the starting substrate’s uptake, respectively. Furthermore, an *in vivo* and/or *in vitro* assay would need to be prepared for each step in order to prove, quantify, and ensure optimized conversion. Since each enzyme comes from a different source organism, isolation and compatibility issues must be considered. Substrates/products can be verified using previously known assays and/or structure determination techniques like HPLC, NMR, or mass spectroscopy.

My work includes performing these toxicity and uptake efficiency studies in order to determine whether glucaric acid affects cell viability and glucose uptake. Efforts have also been made in order to complete the final step of the pathway – obtaining uronate dehydrogenase, expressing it inside *E. coli*, and making sure glucaric acid is produced when glucuronic acid is available.

For expression of the other reaction steps of the benchmark pathway inside *E. coli*, the first two enzymes, *myo*-inositol 1-phosphate synthase and *myo*-inositol oxygenase have already been isolated from their respective source organisms by another

lab member. Methods for measuring glucuronic acid and *myo*-inositol have also been identified (10 and 11, respectively).

1.2 Expressing the final reaction step of the benchmark pathway in *E. coli*

The final reaction step, as mentioned before, is the conversion of glucuronic acid to the desired product, glucaric acid (Figure 2).

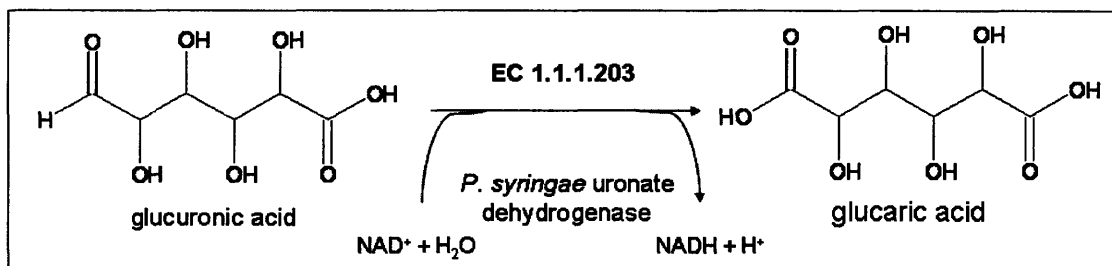


Figure 2: Final Step of the Benchmark Pathway

Glucuronic acid is converted to glucaric acid (with reduction of NAD^+ to $NADH$) by uronate dehydrogenase.

According to the reaction, uronate dehydrogenase, NAD^+ , and glucuronic acid are required for the production of glucaric acid (Figure 2). Since this reaction step occurs in *P. syringae*, the goal was to isolate uronate dehydrogenase from this organism, test the reaction *in vitro*, and express uronate dehydrogenase activity in *E. coli*.

First, an *in vitro* test was performed in order to confirm the presence of uronate dehydrogenase in *P. syringae* DC3000. Secondly, a screening method was determined and identified in order to be able to test uronate dehydrogenase's function in *E. coli*. Finally, uronate dehydrogenase will be isolated and expressed in *E. coli* in order to confirm it functions similarly in this strain.

Specific considerations needed to be made for isolating uronate dehydrogenase and testing its activity *in vivo*. Since the sequence of uronate dehydrogenase is not known, it was not possible to isolate the gene using PCR amplification. Furthermore, although the approximate size of the protein is known (reported to be 60 kilodaltons, 30 kd for each subunit, γ), it was not purified and sequenced. Instead, *P. syringae* cell lysates were prepared and used in order to test for uronate dehydrogenase activity *in vitro*. Then, once a screening method was identified for confirming the function of uronate dehydrogenase in *E. coli*, a partial digestion of *P. syringae* genomic DNA was performed and cloned into a plasmid. This library will be expressed in *E. coli*. The screening method will then be used in order to test for uronate dehydrogenase activity in *E. coli* (and to eventually isolate and sequence the gene). Once this enzyme is successfully identified and its function is confirmed, the final reaction step of the benchmark pathway will be complete.

The following includes procedures and current results for toxicity, uptake efficiency, and expressing the final step of the pathway in *E. coli*. I will conclude by discussing the future work necessary for completing expression of the benchmark pathway. Eventually, the retro-biosynthetically designed glucaric acid pathways can be expressed and compared to the benchmark pathway. The ease of retro-biosynthetic design and feasibility of expression can then be assessed.

2. Materials and Methods

2.1 Glucaric acid toxicity: specific growth rate determination for cells grown in varying concentrations of glucaric acid

Specific growth rates for *E. coli* DH10B cells were determined when grown in increasing concentrations of glucaric acid. DH10B was grown in minimal media (M9 salts, 0.8 mM leucine, 0.1 mM CaCl₂, and 0.2 mM MgSO₄, 12) supplemented with 22 mM glucose and increasing concentrations of glucaric acid. A 5X stock solution of M9 salts consists of 64 g Na₂HPO₅(7H₂O), 15 g KH₂PO₄, 2.5g NaCl, and 5.0 g NH₄Cl in one liter of deionized H₂O (12). Glucaric acid concentrations of 0 mM, 10 mM, 22 mM, and 40 mM were tested, adjusting media to physiological pH (7.4) as needed. Cultures were inoculated with 1% v/v of an overnight LB culture that reached an OD₆₀₀ of 2.3. Cell concentrations were measured at increasing time points. The specific growth rates for cultures grown in varying amounts of glucaric acid were determined using a general, first order rate expression for cell growth,

$$\frac{dX}{dt} = \mu X$$

where X is the cell concentration (OD₆₀₀) during the exponential phase, and μ is the specific growth rate (hr⁻¹). Upon integration, the following expression is obtained,

$$\ln \frac{X}{X_0} = \mu(t - t_0)$$

where X_0 is the initial cell concentration corresponding to the initial time point, t_0 , of the exponential phase. Accordingly, specific growth rates are the slopes when data points are plotted on a semi-log plot.

In order to determine the specific growth rates for *E. coli* DH10B cells when grown in increasing concentrations of glucaric acid, a growth curve was generated for each case. Since the rate expression is valid for cells grown in exponential phase, growth rate determination was based on data points from the exponential phase only (time points between 7 to 12 hours, according to the growth curves obtained). The rate expression was then evaluated for each of the four cell cultures in order to determine specific growth rates and whether glucaric acid is toxic to cells.

2.2 Glucose uptake efficiency in the presence of glucaric acid

Cells were grown in minimal media (same concentrations of M9, CaCl₂, MgSO₄, and leucine as above) supplemented with 22 mM glucose and increasing concentrations of glucaric acid (0 mM, 7.5 mM, and 15 mM). The media was adjusted to physiological pH. After DH10B cells were inoculated into 50 mL cultures, from a 2 mL LB overnight culture (that had reached an OD₆₀₀ of 1.3), 1 mL samples were removed at many time points throughout growth. Each sample was centrifuged at 5,000 rpm for 5 min. The supernatant was saved in the -20°C freezer until glucose concentration was measured. Once the culture reached stationary phase and the final sample was obtained, a YSI analyzer (YSI 2700 select) containing a glucose membrane (YSI 2365 glucose membrane kit; Yellow Springs Instruments Co., Inc.), was used in order to measure glucose concentrations for each time point. The membrane contains immobilized glucose oxidase, which completely oxidizes D-glucose to hydrogen peroxide and D-gluconic acid. The analyzer measures the amount of hydrogen peroxide produced, which reflects the concentration of D-glucose.

2.3 Preparing *P. syringae* lysate and testing for uronate dehydrogenase activity

P. syringae DC3000 was grown in a 25 mL solution of LB and 50 ug/mL of Rifampicin (shaken/incubated at room temperature). At an OD₆₀₀ of 1.1, 0.1% v/v of cell culture was inoculated into a 25 mL minimal media solution (M9 salts, 0.1 mM CaCl₂, 0.2 mM MgSO₄, and 50 ug/mL of Rifampicin) supplemented with 22 mM glucose. After 44 hours of growth, an OD₆₀₀ of 1.9 was reached and 0.1% v/v of this culture was inoculated into 25 mL of minimal media either supplemented with 22 mM glucose or glucuronic acid. At stationary phase, equal concentrations of cells were lysed in order to isolate comparable concentrations of protein. Before cells were lysed, each culture had an OD₆₀₀ of 1.75. This was achieved by diluting cells grown in glucuronic acid, which reached an OD₆₀₀ of 1.84, with minimal media (without antibiotic or carbon source).

Protein extracts were obtained using lysozyme and freeze/thaw cycles in liquid N₂ and 37°C water. A volume of 1 mL of equal concentrations of cell culture (*P. syringae* grown in glucose and glucuronic acid) was centrifuged at 5,000 rpm for 5 min (at 4°C). Pellets were resuspended in 100 uL of Phosphate Buffered Saline solution (1.37 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, and 2 mM KH₂PO₄, 13) and 1 mg/ml lysozyme. The tubes were then incubated in ice for 30 min. Five freeze/thaw cycles were performed with liquid N₂ and 37°C water. Once these cycles were complete, the samples were centrifuged at 14,000 rpm for 10 min (at 4°C). The supernatant was stored at -20°C until the reaction was performed.

These extracts were then reacted with glucuronic acid and NAD⁺. NADH concentrations were then determined by measuring absorbance at 340 nm. Equal protein concentrations were desired for this *in vitro* assay. The Bradford assay (14) was used in

order to measure protein concentrations. A concentrated Protein Assay reagent was purchased from Bio-Rad (Hercules, CA) and used according to manufacturer's protocol. First, a standard curve was obtained by measuring absorbances of known concentrations of BSA solution. The OD₅₉₅ corresponding to BSA concentrations of 0, 0.2, 0.4, 0.6, 0.8, and 1.0 mg/mL were measured. A plot of absorbance vs. protein concentration was generated and a line of best fit was identified.

Protein concentrations of 0.76 mg/mL and 0.68 mg/mL were obtained from cells grown in glucose and glucuronic acid, respectively. These cell extracts were not adjusted to have equal protein concentrations. Simulating the procedure performed by Wagner *et al.* (7), a 0.05 M Na₂HPO₄, 0.33 mM glucuronic acid, and 1 uM NAD⁺ solution was prepared for the reaction. 500 uL of this solution was transferred to a quartz cuvette in order to blank at 340 nm. 1/30th volume of protein was then added to the substrate mixture and absorbance was measured every 30 sec (for 10 min) in order to detect increasing concentrations of NADH. The concentration of NADH was determined for each time point using the Beer-Lambert Law:

$$A = \epsilon * C * l$$

where A is the absorbance (measured at 340 nm), ϵ is the extinction coefficient (which is a known value for NADH at 340 nm), C is the concentration of NADH (in mmol/L), and l is the path length (dependent on the size of cuvette used). The extinction coefficient of NADH, at 340 nm, is 6.22 L/(mmol*cm) (9) and the path length is 1 cm. The expression was evaluated in order to determine NADH concentration that corresponds to the absorbance measurements with time.

2.4 Optimizing screening method

BL21(DE3) was tested in liquid for its ability to grow in glucaric acid or glucuronic acid. Cells were grown in minimal media solution supplemented with 22 mM glucose. 0.1% v/v of a culture that reached an OD₆₀₀ of 2.5 (and then diluted to 0.4 using minimal media solution without carbon source) was added to a 25 mL solution of minimal media either supplemented with 22 mM glucose, glucaric acid, or glucuronic acid.

DH10B was also tested in liquid as well as on minimal media plates supplemented with glucose, glucaric acid, or glucuronic acid. Cells were grown in a 25 mL solution of minimal media (M9 salts, 0.8 mM Leucine, 0.1 mM CaCl₂, and 0.2 mM MgSO₄) supplemented with 22 mM glucose. 0.1% v/v of a culture that had reached an OD₆₀₀ of 2.3 was then inoculated into a 25 mL minimal media solution supplemented with 22 mM glucose, glucaric acid, or glucuronic acid (and pH adjusted as needed). Cells were plated as well. 100 uL of a 20% v/v solution of cell culture (that was at an OD₆₀₀ of 2.3) in water was spread on plates that contained minimal media supplemented with 22 mM glucose, glucaric acid, or glucuronic acid. All liquid and plates were incubated at 37°C. Liquid flasks were shaken at 200 rpm.

DH5alpha and MG1655 cells were tested and compared to the results obtained from BL21(DE3) and DH10B. Cells were grown in minimal media supplemented with 22 mM glucose. DH5alpha required 0.296 mM addition of thiamine in the minimal media solution. For DH5alpha, 0.1% v/v of a culture that reached an OD₆₀₀ 0.91 was added to a 25 mL solution of minimal media (with 0.296 mM thiamine) supplemented with 22 mM glucose, glucaric acid, or glucuronic acid. For MG1655, 0.1% v/v of a culture that

reached an OD₆₀₀ of 3.4 was added to a 25 mL solution with each of the organic acids as well. Cell concentrations were measured with time.

DH10B-pGEM3Z cells were grown in a 25 mL solution of minimal media supplemented with 22 mM glucose and 100 ug/mL of Ampicillin. 0.1% v/v of a culture that reached an OD₆₀₀ of 0.91 was added to a 25 mL solution of minimal media supplemented with 22 mM glucose, glucaric acid, or glucuronic acid (and 100 ug/mL of Ampicillin). Cell concentration was measured with time. Cells were also grown on plates. 100 uL of a 20% solution of cells (those that reached an OD₆₀₀ of 0.91) was spread on plates containing minimal media supplemented with 22 mM glucose, glucaric acid, and glucuronic acid.

2.5 Testing screening method for DH10B-pGEM3Z and DH10B-pBluescript using SOC transfer method

DH10B was transformed with either pGEM3Z or pBluescript using electroporation. Then, after a one hour incubation in SOC at 37°C, cells were spread on minimal media plates supplemented with 22 mM glucose used as a positive control, glucaric acid, or glucuronic acid (with 100 ug/mL Ampicillin and 0.1 mM IPTG). Transformations were diluted in LB to obtain about 100 colonies on plates (100 X 15mm Petri dishes).

2.6 Isolating genomic DNA of *P. syringae* DC3000 for partial digestion

P. syringae was grown in a 25 mL solution of LB and 50 ug/mL of Rifampicin. The Wizard Genomic DNA Purification Kit for gram-negative bacteria (Promega,

Madison, WI) was used in order to isolate the genomic DNA from 2 mL of cells that had reached an OD₆₀₀ of 1.9. A total volume of 400 uL of genomic DNA was obtained. The concentration of DNA was measured in order to determine the amount needed for the partial digestion. The absorbance of isolated genomic DNA, measured at 260 nm, was 31.1. It is known that an absorbance of 1 corresponds to a concentration of 50 ug/mL (15). Therefore, the absorbance obtained at 260 nm corresponds to a DNA concentration of 1,555 ug/mL. Furthermore, the OD₂₆₀:OD₂₈₀ was determined to be close to 2 (1.98) indicating a pure DNA solution. An equal volume of DNA rehydration solution was added in order to dilute this sample. A concentration of 777.5 ug/mL was obtained at a total volume of 800 uL.

2.7 Genomic DNA library construction

For the vector, pBluescript was isolated and digested with BamH1. A 65 ug/mL concentration of pBluescript was obtained. The digestion consisted of 3.64 ug of plasmid, 140 units of BamH1, and 10X BamH1 buffer. The final volume was 70 uL and this was incubated at 37°C for 3.5 hrs. The digest was run on a gel and purified to a concentration of 80 ug/mL (30 uL total volume). Alkaline phosphatase was then used in order to prevent vector re-ligation. For the dephosphorylation reaction, 2.088 ug of this newly digested plasmid was used with 9 units of Calf Intestinal Alkaline Phosphatase. This solution was incubated for 1 hr at 37°C and the DNA was purified using Qiagen's PCR purification kit (Qiagen, Inc., Valencia, CA). The final digested and dephosphorylated vector concentration was 45 ug/mL at a final volume of 30 uL.

In order to obtain the inserts that will hopefully carry uronate dehydrogenase, a partial digestion of the genomic DNA of *P. syringae* was performed. First, inserts within a 1-5 kb range were obtained and tested. These fragments were from several partial digestions. The conditions used for the partial digestions included 40 ug of DNA, 10X Sau3A1 buffer, varying concentrations of Sau3A1 enzyme in order to optimize, and enough ddH₂O in order to reach a final volume of 100 uL. The digest was incubated at 37°C for 2.5 hours. The 1-5 kb range was cut from the gel and purified. Enzyme quantity ranged from 1.5 to 2 units. All partial digestions were run on a gel (about 2 ug of each) and a 1-5 kb fragment range was excised from each and purified to 50 uL. All purified fragments were combined and used as insert. The purified product (1-5 kb range of fragments) was 15 ug/mL. The ligation reaction that resulted in the most successful clones consisted of 0.045 ug of digested, dephosphorylated vector, 0.06 ug of insert, 10X T4 ligase or 400 units (New England Biolabs, Beverly, MA), 10X T4 ligase buffer (New England Biolabs, Beverly, MA), and enough ddH₂O in order to reach a final ligation volume of 10 uL. This reaction was left at room temperature for 12 hrs before it was used for transforming DH10B cells. Cells were electroporated (1 uL of ligation into 20 uL of electrocompetent DH10B cells) and after a one hour incubation in SOC at 37°C, the entire transformation was spread on a single plate containing LB, 100 ug/mL Ampicillin, 0.1 mM IPTG, and 0.04g/L ready-to-use X-gal.

The volume size of the ligation reaction was increased in order to optimize library size. Using the same mass of DNA (insert and vector) and equivalent concentrations for ligase and ligase buffer, the total volume was increased to 25 uL, 50 uL, and 75 uL. These ligations were incubated at 16°C for ten hours. Ethanol precipitation (16) was then

performed in order to reduce the volume back down to 10 uL. The procedure includes adding 1/10 ligation volume of 3 M sodium acetate (at pH 5.5) and 1/100 ligation volume of 1% stock solution of blue dextran (1 g of blue dextran per 100 mL of water). After leaving 100% ethanol inside the -80°C freezer for five minutes, 2.5 times the ligation volume of this was added to the solution. This was centrifuged at 13,200 rpm for 15 min at 4°C. The supernatant was discarded using aspiration, and 50 uL of 70% ethanol was added to the pellet in order to wash. This was then centrifuged at 13,200 rpm for 3 min at 4°C. Finally, once the supernatant was removed via aspiration and all ethanol had been evaporated at room temperature or 37°C for 5-10 min, 10 uL of ddH₂O was added. This was then left in the 4°C refrigerator overnight.

DH10B cells were transformed, using these ligations, in a similar fashion as above. The entire transformation was then plated on LB, 100 ug/mL Ampicillin, 0.1 mM IPTG, and 0.04g/L ready-to-use X-gal.

For further optimization, whole partial digestions were purified using Qiagen's PCR purification method instead of excising size-selected bands and purifying them from a gel. For partial digestions, 40 ug of genomic DNA was used as before. The amount of enzyme was varied in order to try to optimize the size of fragments from the partial digestion. Four partial digestions were performed, using 1, 1.25, 1.5, and 2 units of Sau3A1. A 10X concentration of Sau3A1 buffer and enough volume to reach 100 uL for each reaction was used. The digests were incubated at 37°C for 2.5 hrs. Each partial digest was purified and eluted from the Qiagen column with 35 uL elution buffer. 10 ug was run on a gel to analyze the reactions. All four digests were combined to a total volume of 140 uL and concentration of 175 ug/mL.

The inserts were ligated with the original, BamH1-digested and dephosphorylated pBluescript vector. The optimal ligation, resulting in the most number of clones, consisted of 0.045 ug of digested, dephosphorylated vector, 0.35 ug of insert, 400 units of T4 ligase, 10X T4 ligase buffer, and enough water to reach 10 uL. This reaction was left at 16°C for 9 hrs before DH10B cells were transformed. The entire transformation (20 uL of electrocompetent DH10B cells were transformed with 1 uL of ligation as before) was incubated in SOC for 1 hr at 37°C then spread on a single plate that contained LB, 100 ug/mL Ampicillin, 0.1 mM IPTG, and 0.04g/L ready-to-use X-gal.

In order to determine the size of fragments that resulted in successful clones (to confirm the “real” library size), twenty white colonies were miniprepped and digested in order to check for insert. A blue, insert-less colony was also miniprepped and digested as control. Each colony was grown in a 2 mL solution of LB and 100 ug/mL of Ampicillin. Plasmids were then isolated from 1 mL cultures and eluted to 50 uL. Each plasmid was then digested with HindIII in order to determine the size of the inserts. The fragments are not able to be completely isolated since all cut sites are located at one end of these inserts. However, since the size of pBluescript is known (3 kb in length), the fragment size can be determined using a single digest. Each digestion reaction contained 20% v/v plasmid, 40 units of HindIII, 10X Buffer 2, and enough ddH₂O to reach a total volume of 25 uL. They were incubated at 37°C for 2.5 hrs.

3. Results and Discussion

3.1 Glucaric acid toxicity

A viable cell is required for glucaric acid production. Once the pathway is expressed and glucaric acid is generated (and released from the cell), understanding whether it affects a cell's growth is important for maximizing further production of the desired organic compound.

The specific growth rate defines the rate of cell production. Therefore, similar specific growth rates at each glucaric acid concentration would indicate an equal rate of production of DH10B cells at all concentrations tested. As a result, once this pathway is expressed in *E. coli* and the organic acid is produced, the cells' growth, and, therefore, glucaric acid production, would not be affected.

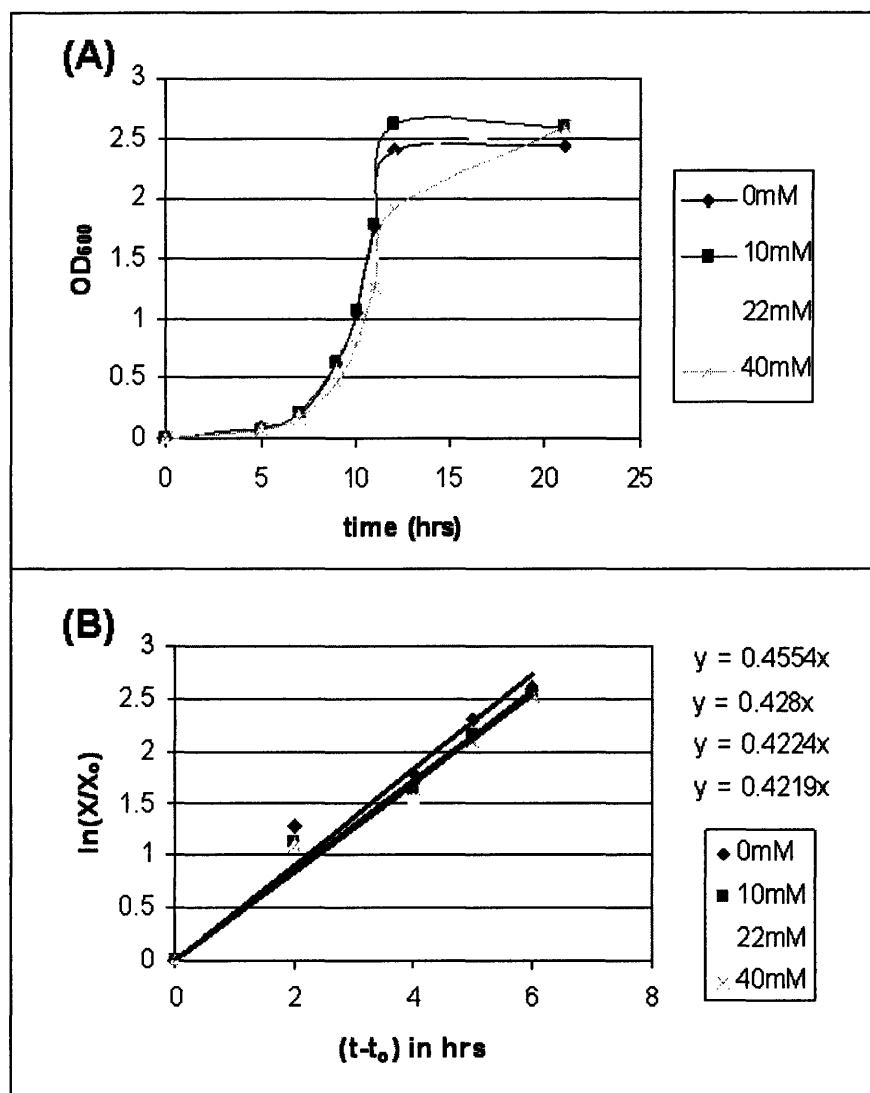


Figure 3: Growth Curve and Specific Growth Rate Determination for DH10B Cells in Glucaric Acid

(A) *E. coli* DH10B cells were grown in minimal media supplemented with 22 mM glucose and increasing concentrations of glucaric acid (pH adjusted). (B) Specific growth rates were only considered at exponential phase (data points were taken from 7 to 12 hrs).

Specific growth rates (the slope of each line) were 0.46, 0.43, 0.42, and 0.42 (hr^{-1}) for cultures grown in glucaric acid concentrations of 0 mM, 10 mM, 22 mM, and 40 mM, respectively (Figure 3B). Since these values are similar, glucaric acid, up to 40 mM tested, does not affect the growth of cells.

Toxicity studies confirmed that cells will remain viable when glucaric acid is produced. Uptake efficiency studies would indicate whether glucaric acid would affect the amount of glucose that can be transported (and, therefore, utilized) inside cells. The goal is to maximize the amount of glucose cells can utilize for energy, for other cellular processes, and for making glucaric acid (since glucose is the starting substrate for the benchmark pathway). Therefore, it is necessary to be confident that glucaric acid production, once the pathway is expressed inside cells, will not affect glucose uptake or utilization. This can be achieved by growing cells in glucose and varying concentrations of glucaric acid. As the cells grew, the amount of glucose still in the media was measured as a function of time. The addition of glucaric acid in the media did not affect glucose uptake since glucose concentrations (in g/L) were very close to equal at all glucaric acid concentrations for each time point (Figure 4).

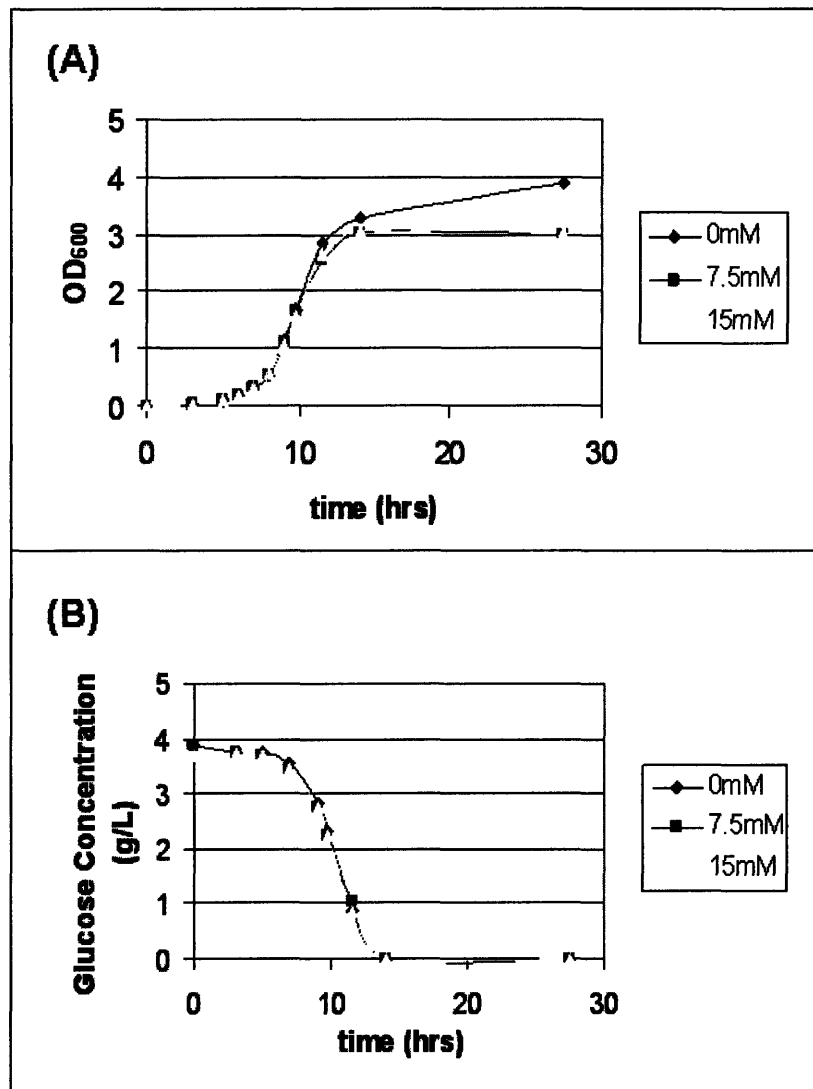


Figure 4: Uptake Efficiency of Glucose in the Presence of Glucaric Acid
 (A) Growth curve and (B) corresponding measurements of glucose concentration in the supernatant at specific time points.

For the growth curve, specific growth rates were calculated as before. The exponential phase was taken to be between 6 and 11.5 hrs. Specific growth rates for each culture were 0.56, 0.54, and 0.54 (hr^{-1}) for cells grown in 22 mM glucose and 0 mM, 7.5 mM, and 15 mM, respectively. The fact that specific growth rates were similar at increasing concentrations of glucaric acid confirms the *E. coli* DH10B toxicity results.

3.2 Testing *P. syringae* DC3000 for uronate dehydrogenase activity *in vitro*

After product considerations, efforts can now be made in order to express each reaction step of the benchmark pathway in *E. coli*. The function of uronate dehydrogenase was tested *in vitro* in order to confirm its presence in *P. syringae* as well as its ability to convert glucuronic acid to glucaric acid. This was possible because of the ability to repress or activate expression depending on the carbon source used. It is known that in *P. syringae*, glucose inhibits uronate dehydrogenase expression. However, glucuronic acid stimulates it (8). Once cells grow in glucose or glucuronic acid, their protein can then be isolated and reacted with glucuronic acid and NAD^+ . The amount of product can then be measured as an indication of whether uronate dehydrogenase is available or not. Since glucaric acid assays using the HPLC are in development, the enzyme's function was indirectly measured by determining the concentration of NADH. Cells expressing uronate dehydrogenase, those that were grown in glucuronic acid, should show an increase in NADH production due to this enzyme's ability to reduce NAD^+ . Conversely, low NADH concentration is expected for cells grown in glucose.

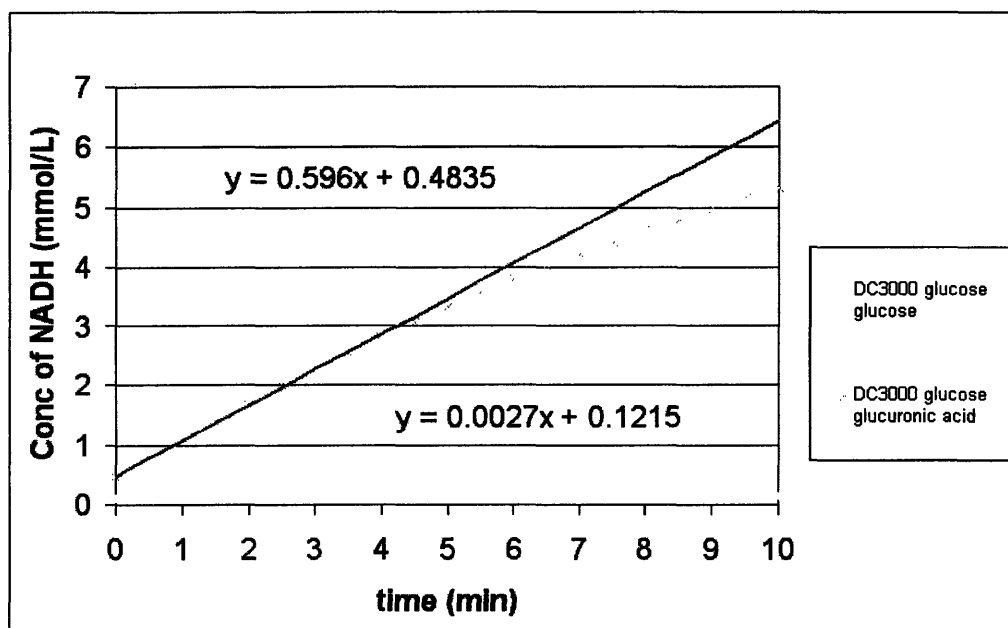


Figure 5: Measuring Uronate Dehydrogenase Activity

Slope is mmole of NADH/(L*min). The Beer-Lambert Law was used with an extinction coefficient of 6.22 L/(mmol*cm) and a path length of 1 cm.

The specific activities for protein extracts obtained from *P. syringae* grown in glucose and glucuronic acid were 0.1 U/g of protein and 26.4 U/g of protein, respectively. In this case, units (U) represents the amount of NADH produced (mmol) per unit time (min) in a total reaction volume of 500uL. Therefore, there is a significant increase in NADH production with protein extracts obtained from cells grown in glucuronic acid. Since it was reported that uronate dehydrogenase expression is stimulated with glucuronic acid and repressed when cells are grown in glucose, it is apparent the increase in NADH concentration via oxidation of glucuronic acid is due to the presence of uronate dehydrogenase in *P. syringae*. Future work will include using the HPLC (and a boronic acid gel in order to isolate glucuronic acid from other organic acids, 17) in order to measure glucuronic acid concentration. This will be invaluable for both proving uronate

dehydrogenase activity in *P. syringae* as well as measuring glucaric acid produced once the entire biosynthetic pathway is expressed in *E. coli*.

From previous articles concerning purification of uronate dehydrogenase and the increase in NADH concentration when the expression of this enzyme is activated (7, 8, 9), it is evident that uronate dehydrogenase functions as shown and can be found in *P. syringae*. Now, steps can be performed in order to identify a screening method that can be used in order to confirm uronate dehydrogenase's activity in *E. coli*.

3.3 Screening method for detecting uronate dehydrogenase activity in *E. coli*

If *E. coli* is unable, or just takes a very long time, to grow in glucuronic acid but can grow well in glucaric acid, this would offer a way to screen for cells that contain uronate dehydrogenase activity. This enzyme converts glucuronic acid to glucaric acid. As a result, *E. coli* containing uronate dehydrogenase would be able to grow in glucuronic acid since it would have the ability to make glucaric acid. Colonies that grow on both glucuronic acid and glucaric acid plates would therefore contain fragments, obtained from the partial digestion of the genomic DNA of *P. syringae*, that have uronate dehydrogenase activity.

Specifically, fragments that were obtained from the genomic DNA of *P. syringae* will be cloned into a vector and expressed in *E. coli*. Those cells that grow in glucuronic acid would be selected for since they would have a high probability of containing uronate dehydrogenase activity. The next step would be to purify protein after overexpressing these DNA fragments. Then, using similar *in vitro* tests as described above, NADH would be measured and compared to *E. coli* without *P. syringae* DNA fragments, as a

negative control, at 340 nm. Glucaric acid measurements would also need to be made once the HPLC procedure is developed. The screening method is dependent upon *E. coli*'s ability to grow in glucaric acid and not in glucuronic acid. The strain also depends on the vector that is chosen for the library.

It was preferred that a vector containing two promoters of opposite direction on either side of the multiple cloning site be chosen. That way, genes can be expressed from a plasmid-based promoter regardless of what direction they are placed inside the vector (especially important since inserts are obtained from a partial digestion of the entire genomic DNA and only one enzyme/cut site is used). The vector pGEM3Z was chosen because of this quality. It contains a T7 (up) and SP6 (down) promoter. However, in order to use both promoters, an *E. coli* strain with both the T7 and SP6 RNA polymerase had to be obtained. No strains were found with this capability. Therefore, the *E. coli* strain BL21(DE3) was used for its ability to use the T7 promoter. The lac promoter, which is in the opposite direction, was also available on the vector. Both promoters can be used (lac promoter, upon induction with IPTG as well as T7 since strain includes the T7 RNA polymerase) and there would be blue/white selection. Cell concentration as a function of carbon source was measured with time. However, results showed that BL21(DE3) does not grow in glucaric acid but does grow in glucuronic acid (Table 1). Because of this, BL21(DE3) cannot be used in order to test the library. The proposed screening method is not compatible with this strain.

time (hrs)	glucose	glucaric acid	glucuronic acid
13		0.003	
18		0.003	
24		0.006	
35		0.031	

Table 1: Growth of BL21(DE3) in Liquid Media Containing Organic Acids
 BL21(DE3) liquid results in 22 mM glucose, glucaric acid, and glucuronic acid. No cell growth was apparent in glucaric acid. However, cells did grow in glucuronic acid.

Liquid Results			
time (hrs)	glucose	glucaric acid	glucuronic acid
7	0.023	0.003	0.003
22		0.009	0.012
47		0.048	0.006
53			0.005
59			0.006
77			0.007

Plate Results			
time (hrs)	glucose	glucaric acid	glucuronic acid
21	very tiny	none	none
29		none	none
47		tiny	none
60			none

Table 2: Growth of *E. coli* DH10B in Liquid and on Solid Media Containing Organic Acids

Liquid and Plate results for DH10B grown in 22 mM glucose, glucaric acid, and glucuronic acid.

E. coli DH10B cells were tested as well. Cells were grown in minimal media supplemented with glucose, glucaric acid, and glucuronic acid. Cells were also spread on plates containing these organic acids. DH10B liquid results show that these cells can grow in glucaric acid, after 53 hrs, but not in glucuronic acid. Furthermore, colonies appear on glucaric acid plates, after 60 hrs in this case, but not on plates supplemented

with glucuronic acid (Table 2). Therefore, the proposed screening method seems as though it would work with these cells. The next steps would be to confirm that cells containing pGEM3Z act similarly in glucose, glucaric acid, and glucuronic acid as above, and, finally, ligate fragments in the vector, express fragments in cells, and select for colonies that grow on glucuronic acid.

The difference in growth between BL21(DE3) and DH10B may be due to a genetic difference in strains. Because of this great difference in *E. coli* strain growth in liquid cultures, it was interesting to test other strains, like MG1655 and DH5alpha, in order to determine what strains show similar results with DH10B. Furthermore, if a different strain is found to grow in glucaric acid but not glucuronic acid or in glucaric acid at a faster rate, this may be more advantageous when the library is screened (although it would not be possible to use the T7 promoter in either case). Faster growth in glucaric acid would assume faster growth in glucuronic acid with uronate dehydrogenase, depending on the rate of conversion from glucuronic acid to glucaric acid. This would result in a faster observation of colonies containing uronate dehydrogenase activity.

DH5alpha Results			
time (hrs)	glucose	glucaric acid	glucuronic acid
21	0.024	0.007	
27		0.003	
32		0.016	

MG1655 Results			
time (hrs)	glucose	glucaric acid	glucuronic acid
15		0.019	
29			

Table 3: Growth of DH5alpha and MG1655 in Liquid Media Containing Organic Acids

DH5alpha grew in glucuronic acid but not in glucaric acid during the time cells were measured. MG1655 grew both in glucaric acid as well as in glucuronic acid.

DH5alpha grew in glucuronic acid but not in glucaric acid (Table 3). However, since cell concentrations were not measured at longer times, it is not evident whether the cultures just needed more time. MG1655, however, was able to grow both in glucaric acid as well as in glucuronic acid within the time that was measured. Due to growth in glucuronic acid, neither strain can be used for the library. Therefore, it is necessary to revert back to DH10B.

Liquid Results			
time (hrs)	glucose	glucaric acid	glucuronic acid
18		0.009	0.002
59		0.027	0.003
64		0.026	
72		0.067	
83			
87			0.002
94			0.004
112			0.010

Plate Results			
time (hrs)	glucose	glucaric acid	glucuronic acid
17	none	none	none
34	none	none	none
44	none	none	none
58		none	none
63		none	none
94		none	none
106		none	none

Table 4: Growth of DH10B-pGEM3Z in Liquid and on Solid Media Containing Organic Acids

Comparing Table 4 to Table 2, although different starting cell concentrations were used, it seems the vector causes cells to grow slower in liquid and to not grow at all on plates when supplemented with 22 mM glucaric acid.

DH10B containing pGEM3Z was tested for growth on the organic acids. This strain, containing the plasmid of choice for the library, was tested both in liquid and on plates in order to ensure similar results to DH10B without plasmid. DH10B-pGEM3Z was not able to grow on plates supplemented with either glucaric acid or glucuronic acid (Table 4). Furthermore, comparing liquid results of Table 4 to Table 2, although at different starting cell concentrations, it seems pGEM3Z caused cells to grow slower in glucaric acid. As expected, cells did not grow in glucuronic acid though. Since DH10B-pGEM3Z was not able to grow in glucaric acid, pGEM3Z cannot be the vector of choice

for the library. The lack of growth in glucaric acid may be due to the high copy number of pGEM3Z, possibly resulting in a metabolic burden. Instead, a lower copy number plasmid, pBluescript, was tested both in liquid and on plates. Furthermore, a new method of transfer (from liquid to plates and spreading technique) was used in order to ensure it was not the method but the type of vector that caused the changes that were seen.

This time, a procedure was used that simulated how the actual library would be tested. Once fragments obtained from the partial digest of *P. syringae* are ligated into the vector of choice and library size has been optimized, *E. coli* would then be transformed with these newly constructed, insert-containing, plasmids. Once transformed, cells would then be spread on minimal media plates supplemented with either glucaric acid or glucuronic acid. Therefore, whole plasmids (without fragments) were tested this way.

For purposes of determining whether the colonies grow on glucaric acid plates and not on glucuronic acid as well as ensuring that equal numbers of colonies grow on glucose and glucaric acid, efforts were made in order to have a countable number of colonies on each plate. Furthermore, colonies were spread out in order to ensure that neighboring cells will not affect how cells grow on their respective plates.

DH10B-pGEM3Z Results				
time (hrs)	LB	glucose	glucaric acid	glucuronic acid
24			none	none
120			~60 (tiny and sickly)	none

DH10B-pBluescript Results				
time (hrs)	LB	glucose	glucaric acid	glucuronic acid
18			none	none
40				very tiny
90				tiny/small

Table 5: Growth of DH10B-pGEM3Z and DH10B-pBluescript on Solid Media Containing Organic Acids

Cells containing pBluescript grew well on glucaric acid. The small colonies that grew on glucuronic acid are not an issue considering how small they were even after 90 hrs. It may merely be cells that grew in the LB media transformations were diluted with before plating.

The method of transfer and plating may or may not have affected how cells grew according to plate results seen in Table 4. There were 60 DH10B-pGEM3Z colonies on glucaric acid plates after 120 hrs (Table 5). However, the small colonies can either be due to the length of time in incubation (plate conditions may have changed after so long) or poor growth due to the vector used. DH10B-pBluescript grew well on glucaric acid using this method. However, small colonies were present on glucuronic acid as well. This will be overlooked due to the length of time it took these colonies to grow and the possibility that these cells grew merely from the LB the transformations were diluted and spread with. Comparing DH10B-pBluescript on glucose and glucaric acid plates, equal colonies grew on glucose and glucaric acid. Furthermore, a similar number of colonies were present on LB and glucose plates (76 and 86 colonies, respectively). This ensures that a

maximum number of colonies can grow on minimal media plates. This is advantageous since this amount would reflect the amount of colonies that would be tested on glucuronic acid when the library is finally expressed.

Since DH10B-pBluescript, and the method used to transfer and plate cells, appeared to work well, it was not worth determining the cause for the results seen with DH10B-pGEM3Z. The strain that will be used is DH10B. The vector of choice for my library is pBluescript. Now, the library can be constructed and expressed in DH10B. Then, those colonies that grow on glucuronic acid plates would be selected as having fragments with uronate dehydrogenase activity.

3.4 Genomic DNA library construction (cloning uronate dehydrogenase)

In order to isolate uronate dehydrogenase, a partial digest of the genomic DNA of *P. syringae* was performed. The goal is to clone these fragments into pBluescript and then express them in *E. coli* DH01B. Finally, using the proposed screening method already discussed, select for colonies that exhibit uronate dehydrogenase activity.

BamH1 was used in order to digest pBluescript. For inserts, the genomic DNA of *P. syringae* DC3000 was isolated and a partial digestion was performed. The Sau3A1 enzyme, which has a four base pair recognition site, was used (rather than a 6 base pair site) in order to maximize the number of cuts that can be made. This site is compatible with the BamH1 site and since pBluescript contains this restriction site, Sau3A1 can be used. Partial digestion would then result in a wide range of fragment sizes and a larger number of fragments that may carry uronate dehydrogenase activity. The optimal

condition would result in a partial digest that gives the most fragments that are within the target gene's size range.

In order to ensure this range of fragments is obtained, partial digestions were run on a gel and a 1-5 kb range was excised and gel-purified. Uronate dehydrogenase is estimated to be about 1,300 bp, based on the molecular weight of the protein. Figure 6 is a gel that includes two partial digests (left side of ladder). Lane 1 is from a partial digest that includes 0.5 units of Sau3A1. The second lane contains 0.75 units of enzyme. A 1-5 kb piece of gel that was cut out from the partial digest that contained 0.75 units of enzyme is also shown as an example (each partial digest contained 40 ug of DNA, 10XSau3A1 buffer, enough ddH₂O to reach 100 uL once varying amounts of Sau3A1 was added, and incubated at 37°C for 2.5 hrs).

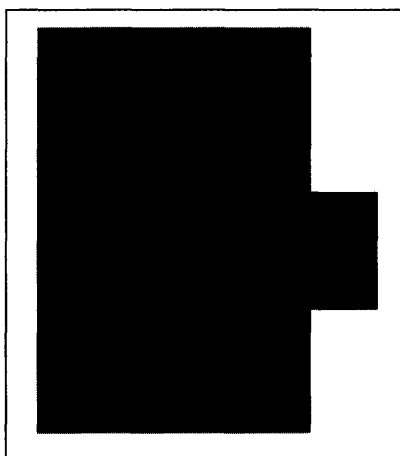


Figure 6: Excision of a 1-5kb Range of Fragments from a Partial Digestion of the Genomic DNA of *P. syringae*

Partial digest of the genomic DNA of *P. syringae* at two different digestion conditions. Lane 1: 2ug of a partial digest containing 40ug of DNA, 0.5 units of Sau3A1, and incubated for 2.5 hrs at 37°C. Lane 2: similar conditions but with 0.75 units of Sau3A1 instead. Lane 4: same as lane 2. The 1-5 kb range of fragments was cut from the gel in order to show as an example. Lane 3 contains 0.25 ug of a 1.0 kb DNA ladder (New England Biolabs, Beverly, MA).

The partial digests resulted in fragment sizes that covered a range between ~0.1 kb and ~15 kb (Figure 6). Instead of considering the entire range of fragments, a 1-5 kb range of digest was cut out of the gel and purified. Isolating this range will limit those fragments that are of impractical size (either too small or too large to be uronate dehydrogenase). However, it is important to note that fragments too large still have the opportunity of expressing uronate dehydrogenase, either from its own promoter, or the gene of interest may be within expression limits according to the vector's promoter and strain's RNA polymerase.

As mentioned in the Materials and Methods section, several partial digests were actually performed, containing enzyme concentrations from 1.5 to 2.0 units (Figure not shown). Partial digests were run on a gel and a 1-5 kb range was cut from each lane similar to how it is shown in Figure 6. Each excised gel was purified and combined. Now that the inserts and vector (digested pBluescript) are available, a ligation reaction was performed, DH10B was transformed (via electroporation), and the library size was counted.

If uronate dehydrogenase is estimated to be 1,300 bp (assumed from the measured protein size), the number of clones required in order to obtain the gene inside fragments of this exact length can be determined. Although, due to my partial digestion, there is a wide range of fragments, this estimation can still be used in order to set the required library size. The following equation can be used in order to determine the number of clones needed (18).

$$N = \frac{\ln(1 - P)}{\ln(1 - f)}$$

From the equation, N is the number of clones needed, P is the desired probability that the sequence will be represented (in the clone), and f is the ratio of gene size to the size of the chromosome. The genomic DNA of *P. syringae* DC3000 was used for my partial digestion. The size of this strain's chromosome is about 6.4 million base pairs. If the desired probability is 99% and all the fragments were about 1,300 bp, a library size of about 22,700 would be needed.

In a total of 100 colonies, half contained clones while the other half were insert-less vector even when alkaline phosphatase was added. Attempts were made to increase the library size by diluting the ligation and then concentrating this down to 10 μ L using ethanol precipitation. Diluting the ligation may result in more successful clones since inserts would have more room to ligate with vector. Then, after the reaction is concentrated in order to regain high DNA concentration, cells would be transformed and the number of clones would be identified. After size is optimized by varying dilution, the total volume would be increased and concentrated in order to transform cells with even more DNA. This way, the transformation will be more efficient and more cells would contain clones. However, results showed that the dilution step was not successful.

Plate results showed no improvement to the number of colonies. There were about 100 colonies per transformation on each plate. Half were successful clones while the other half were insert-less vector. A ligation that results in 50 successful clones would not be able to reach needed library size even after increasing the total volume (increasing the amount of DNA).

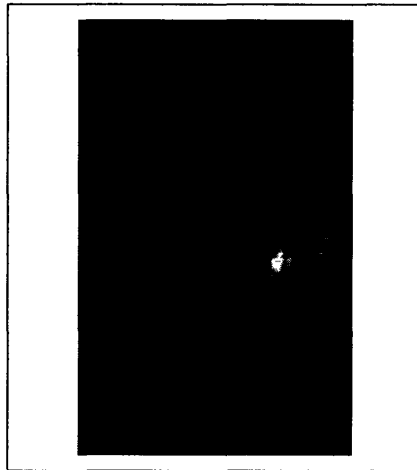


Figure 7: Obtaining Fragments from the Entire Partial Digestion of the Genomic DNA of *P. syringae*

Four different partial digestions. 10ug of each was run on a gel. Each reaction included 40 ug of genomic DNA, 10XSau3A1 buffer, 1, 1.25, 1.5, or 2 units of Sau3A1 (lanes 2 to 5, respectively), and enough water to reach 100 uL. Lanes 1 and 6 contain 0.25 ug of a 1.0 kb DNA ladder (New England Biolabs, Beverly, MA).

After a large enough library size could not be obtained, a new method for insert isolation was performed. Instead of purifying a 1-5kb range of fragments from a gel, the entire partial digest was purified using Qiagen's PCR purification method. The partial digestions spanned an average of ~0.1 kb to ~10 kb (Figure 7). After each partial digest was purified, all four digests were combined as insert. These inserts were ligated into pBluescript and *E. coli* DH10B was transformed. Plate results showed there to be about 1,000 successful clones and only ~50 (1/20th) colonies containing insert-less vector. This is a 20 fold improvement over ligations consisting of gel-purified inserts.

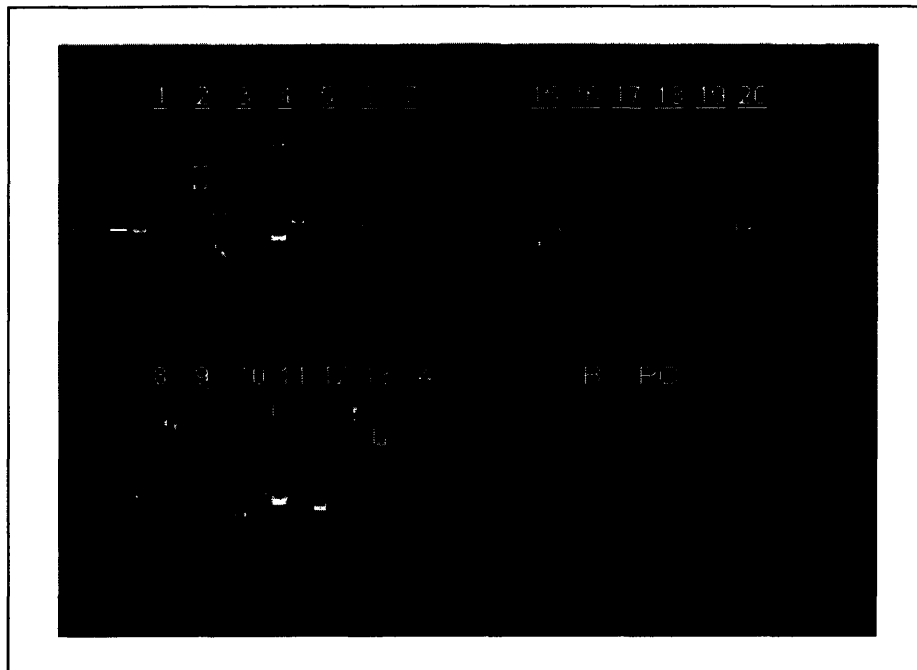


Figure 8: Determining Insert Sizes from Successful Clones

Twenty colonies were isolated and miniprepmed. Each uncut and cut plasmid is shown. The (B) lanes include uncut and cut plasmids from a single blue, insert-less, colony. The (PC) lanes include uncut and cut pBluescript. The first lane of each half of the gels contains 0.25 ug of a 1.0 kb DNA ladder (New England Biolabs, Beverly, MA). The 3 kb mark (size of pBluescript) on ladder is indicated in order to show where vector, without insert, should be.

Each set of two lanes, corresponding to each colony number or type, includes uncut and cut plasmid (Figure 8). Out of twenty colonies checked, seventeen appeared to be successful clones (unless fragments from colonies 10, 16, and 18 were just too small to detect). There is a wide range of insert sizes. As an example, set 3 (corresponding to colony #3) contains uncut and cut plasmid in its first and second lane, respectively. According to the lane that includes the digest, the fragment size is ~7 kb. Clone #4 contained an insert that was too small (less than 0.5 kb). Since uronate dehydrogenase is estimated to be 1,300 bp, this fragment would have a low chance of having this enzyme's activity. Furthermore, inserts that are too large may not express fully unless their

promoters are included inside the fragment itself. Sets 5, 6, 7, 9, 11, 14, 17, and 20 contained fragments of practical size. Therefore, 8/20 clones have a chance of exhibiting uronate dehydrogenase activity. According to these results, for my library size of 1,000, and out of 950 successful clones, it is estimated that 380 colonies could have uronate dehydrogenase activity. The current library size is, therefore, about 380. The plan is to increase the volume 10 fold. Previous experience with ethanol precipitation showed a 100 fold increase in colonies when a 10 fold volume increase of ligation is performed (data not shown). Future work will include optimizing ligation volume in order to ensure there is a 100 fold increase in library size.

4. Conclusions and Future Directions

DH10B cells were grown in varying concentrations of glucaric acid in order to test whether glucaric acid, once produced by cells, is toxic to them. Then, glucose uptake was analyzed in the presence of glucaric acid in order to ensure maximum substrate utilization once glucaric acid is produced. The vector for the library was chosen and tested on glucose, glucaric acid, and glucuronic acid in order to verify that the screening method will work. Finally, steps have been made in order to optimize library size in order to express the final step of the benchmark pathway in *E. coli*.

According to the results, glucaric acid is not toxic to DH10B cells. Specific growth rates were determined for cells grown in 22 mM glucose and increasing concentrations of glucaric acid. The values were very similar, regardless of the concentration of glucaric acid (up to 40 mM) cells were grown in. Furthermore, glucaric acid does not affect glucose uptake and so substrate utilization would not be hindered

either. This was shown by measuring the concentration of glucose in the media as cells were grown in varying amounts of glucaric acid (with an initial concentration of 22 mM glucose in each culture). There was no apparent difference in glucose concentrations, at any time point, regardless of the amount of glucaric acid present in the media.

The screening method was based on selecting for cells that can grow on glucuronic acid plates. This was dependent on whether *E. coli* strains can grow on glucaric acid and not on glucuronic acid. Then, because of the function of uronate dehydrogenase, once cells contain this enzyme (after expressing fragments obtained from the partial digestion), they will be able to grow on glucuronic acid – converting glucuronic acid to glucaric acid, which can then be converted to biomass. DH10B cells harbored the ability to grow on glucaric acid and not on glucuronic acid plates. Unlike DH10B-pGEM3Z, DH10B-pBluescript was also clearly able to grow on glucaric acid and not on glucuronic acid. Furthermore, the maximum number of colonies was able to grow on glucaric acid as compared to the number of colonies on LB and glucose plates. This would, therefore, maximize the number of clones tested on glucuronic acid plates. From these results, pBluescript is the vector of choice. The library is now being constructed and optimized in order to express the final step of the pathway and select for cells that contain uronate dehydrogenase activity.

It was evident that inserts purified from a gel resulted in lower library size than those obtained straight from the partial digestion (using the PCR purification method). Steps are now being taken in order to obtain the necessary number of successful clones. Then, cells will be transformed and spread on glucuronic acid plates. Colonies that grow will be further analyzed for their ability to convert glucuronic acid to glucaric acid. Then,

protein will be isolated and tested *in vitro* for uronate dehydrogenase activity (ability to produce NADH and glucaric acid). The HPLC (and a proposed boronic acid gel) will be used in order to identify, quantify, and purify the glucaric acid that is produced.

Future work will also include isolating the enzymes necessary in order to complete the other steps of the benchmark pathway. An assay for each reaction step will be required in order to prove that the enzyme was successfully isolated from the host organism and functions properly in *E. coli*. Once the entire pathway is expressed and glucaric acid is produced, designed pathways using retro-biosynthesis will then be expressed. The ease of design, expression, and ability to produce glucaric acid will then be assessed by comparing it to the methods taken in designing and expressing the benchmark pathway.

REFERENCES

1. <http://www.genome.jp/kegg/>
2. <http://www.brenda.uni-koeln.de/>
3. Li, C; Henry, CS; Jankowsky, MD; Ionita, JA; Hatzimanikatis, V; Broadbelt, LJ. 2004. Computational Discovery of Biochemical Routes to Specialty Chemicals. *Chemical Engineering Science*. 59:5051-5060.
4. Lee, PC; Momen, AZ; Mijts, BN; Schmidt-Dannert, D. 2003. Biosynthesis of Structurally Novel Carotenoids in *Eschericia coli*. *Chemistry and Biology*. 10:453-462.
5. Hansen, Chad A; Dean, Amy B; Draths, KM; Frost, JW. 1999. Synthesis of 1,2,3,4-Tetrahydroxybenzene from D-Glucose: Exploiting *myo*-Inositol as a Precursor to Aromatic Chemicals. *Journal of American Chemical Society*. 121:3799-3800.
6. Moskala, R; Reddy, CC; Minard, RD; Hamilton, GA. 1981. An Oxygen-18 Tracer Investigation of the Mechanism of *myo*-Inositol Oxygenase. *Biochemical and Biophysical Research Communications*. 99:107-113.
7. Wagner, G; Hollmann, Siegfried. 1976. Uronic Acid Dehydrogenase from *Pseudomonas syringae*: Purification and Properties. *European Journal of Biochemistry*. 61:589-596.
8. Kilgore, WW; Starr, MP. 1959. Uronate Oxidation by Phytopathogenic Pseudomonads. *Nature*. 183:1412-1413.
9. Bateman, DF; Kosuge, T; Kilgore, WW. 1970. Purification and Properties of Uronate Dehydrogenase from *Pseudomonas syringae*. *Archives of Biochemistry and Biophysics*. 136:97-105.

10. Charalampous, FC; Chryssoula, L. 1957. Biochemical Studies on Inositol: IV. Conversion of Inositol to Glucuronic Acid by Rat Liver Extracts. *Journal of Biological Chemistry*. 228:1-13.
11. MacGregor, LC; Matschinsky, FM. 1984. An Enzymatic Fluorimetric Assay for *myo*-Inositol. *Analytical Biochemistry*. 141:382-389.
12. Sambrook, J; Russel, DW. 2001. Molecular Cloning Laboratory Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York. Volume 3, section A2.2, third edition.
13. Sambrook, J; Russel, DW. 2001. Molecular Cloning Laboratory Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York. Volume 3, section A1.7, third edition.
14. Bradford, MM. 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. *Analytical Biochemistry*. 72:248-254.
15. Sambrook, J; Russel, DW. 2001. Molecular Cloning Laboratory Manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York. Volume 1, section 6.11, third edition.
16. <http://www.research.umbc.edu/~jwolf/m5.htm>
17. Poon, R; Villeneuve, DC; Chu I; Kinach, R. 1993. HPLC Determination of D-Glucaric Acid in Human Urine. *Journal of Analytical Toxicology*. 17:146-150.
18. <http://www.promega.com/faq/genclone.html#q05>