Biosensor Directed Protein Engineering

Presented by: Alice Tong

In partial fulfillment of the requirement for graduation with the Dean's Scholars Honor's degree in Biochemistry.
Under the supervision of John Leavitt and Dr. Hal Alper.
Spring 2015

_____ _____

Hal Alper — Date

Supervising Professor

_____ _____

Jeffrey Barrick — Date

Honors Advisor in Biochemistry

Table of Contents

Abstract

Microbial species can be used to convert renewable feed-stocks into valuable chemical compounds in a cost-effective and environmentally friendly manner. Through the expression of a heterologous enzymatic pathway in *Saccharomyces cerevisiae*, the polymer precursor muconic acid was synthesized drawing off flux from the Shikimate pathway. With the application of metabolic engineering of the host yeast strain, production of muconic acid was improved 40-fold.

To facilitate protein engineering of pathway elements using high throughput screening, a biosensor targeting the downstream products of the Shikimate pathway was developed utilizing an inducible hybrid-promoter based on the ARO9 promoter. First, the limits of rational engineering were reached on the target strain. Once this point was reached, the biosensor was integrated into the engineered strain. Utilizing this biosensor, an initial round of mutagenesis and selection was performed on the TKL1 gene, mutant proteins isolated and their improvements quantified.

Protein engineering of TKL1 and ARO1 will allow the production of a yeast strain which functions as a strong host for production of muconic acid and other Shikimate pathway products. The process provides insight into the regulatory landscape of amino acid biosynthesis and principles of enzymatic catalysis.

Introduction

Muconic Acid is a molecule of interest for biological production due to its properties as a chemical precursor for plastics such as polyethylene terephthalate, nylon-6,6 and polyurethane (Curran *et al.*, 2013). Currently, these molecules are derived from Benzene and Para-Xylene which are products of oil. As a result of the global push to reduce the petroleum footprint, a number of methods have been employed to produce this and other such specialty chemicals from renewable carbon sources.

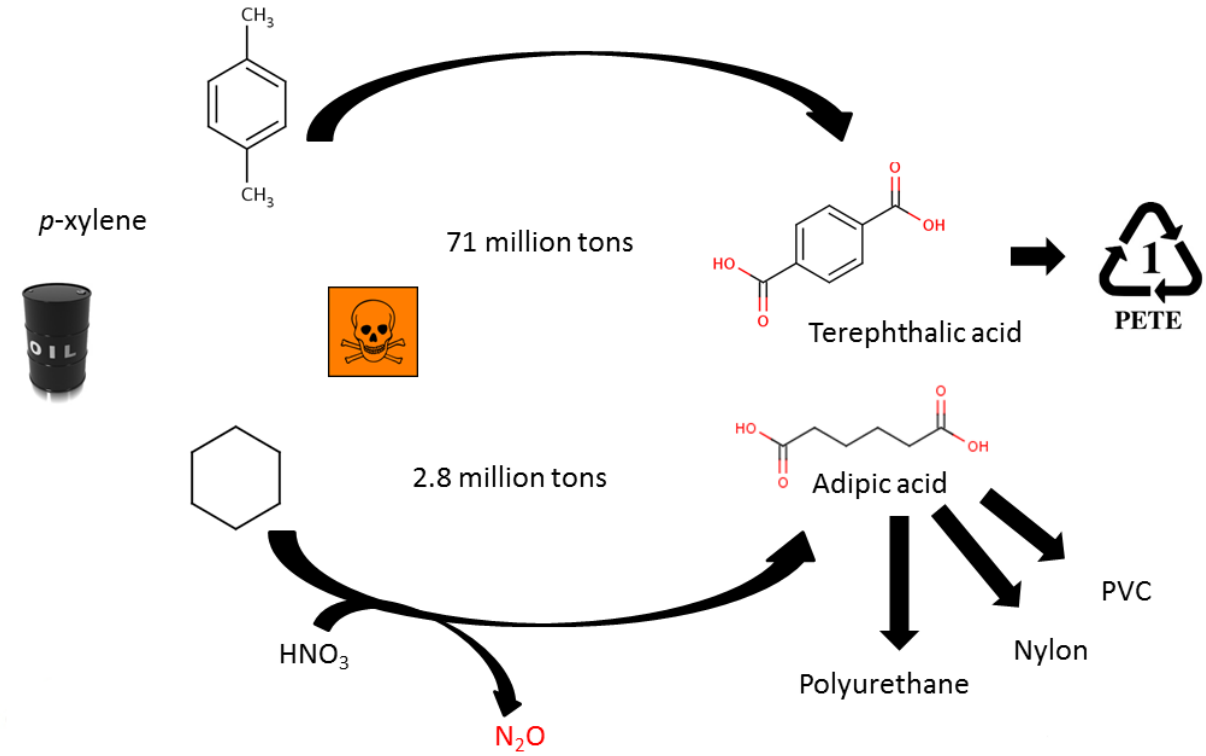


Figure 1. The reaction schematic for production of plastics and polymers from oil, through p-xylene and benzene.

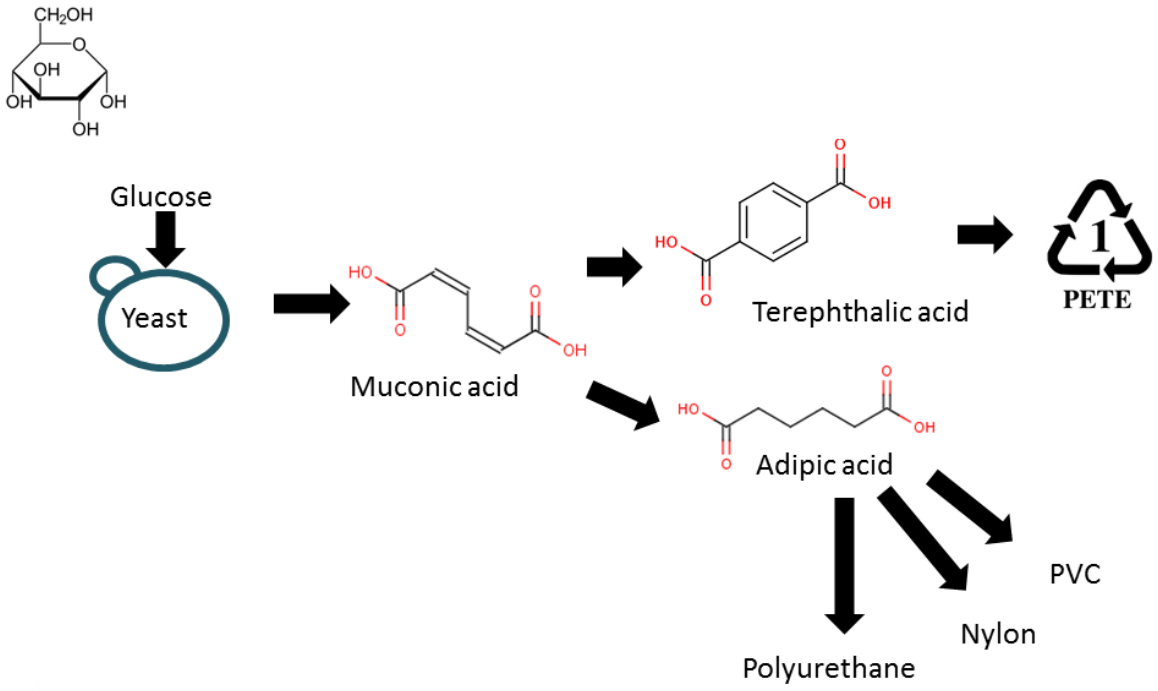


Figure 2. Reaction pathway featuring bioproduction of muconic acid leading to the same plastics and polymer products as figure 1, from glucose.

Muconic acid was previously produced in recombinant *Escherichia coli* cells through a heterologous pathway. In order to produce it in yeast, a synthetic production pathway had to be created. Translating metabolic pathways from one organism into a non-native organism can often be difficult. The inherent complexities of regulation and degree of interconnectedness of cellular metabolism can compound this problem (Curran *et al.*, 2012). This is particularly important when attempting to develop a novel pathway for an unnatural product.

Specific factors such as a lower growth pH, lower optimal growth temperature, sustainability when grown in large fermentation-style cultures, and lack of susceptibility to phage contamination make the yeast *Saccharomyces cerevisiae* the optimal host organism for industrial scale production of muconic acid. Most notably of these factors, the lower pH requirements make this yeast particularly well suited to produce the acid, as production of the acid should not adversely affect the growth of the yeast. For these reasons it was decided to pursue further optimization of the muconic acid production pathway in yeast, rather than continuing work in *E. coli*.

Based on previous work from the construction of the *E. coli* pathway, it was known that a 3-dehydroshikimate (DHS) dehydrase, a protocatechuic acid (PCA) decarboxylase, and a 1,2 catechol-dioxygenase were needed (Draths and Frost, 1994; Niu et al., 2002; Curran *et al.*, 2013). Various candidates for these enzymes were first individually expressed in *Saccharomyces*

cerevisiae and tested for activity using an *in vitro* enzyme assay. Strong candidates were codon and expression-optimized and cloned into plasmids. This represented the first time muconic acid was produced in yeast (Curran *et al.*, 2013).

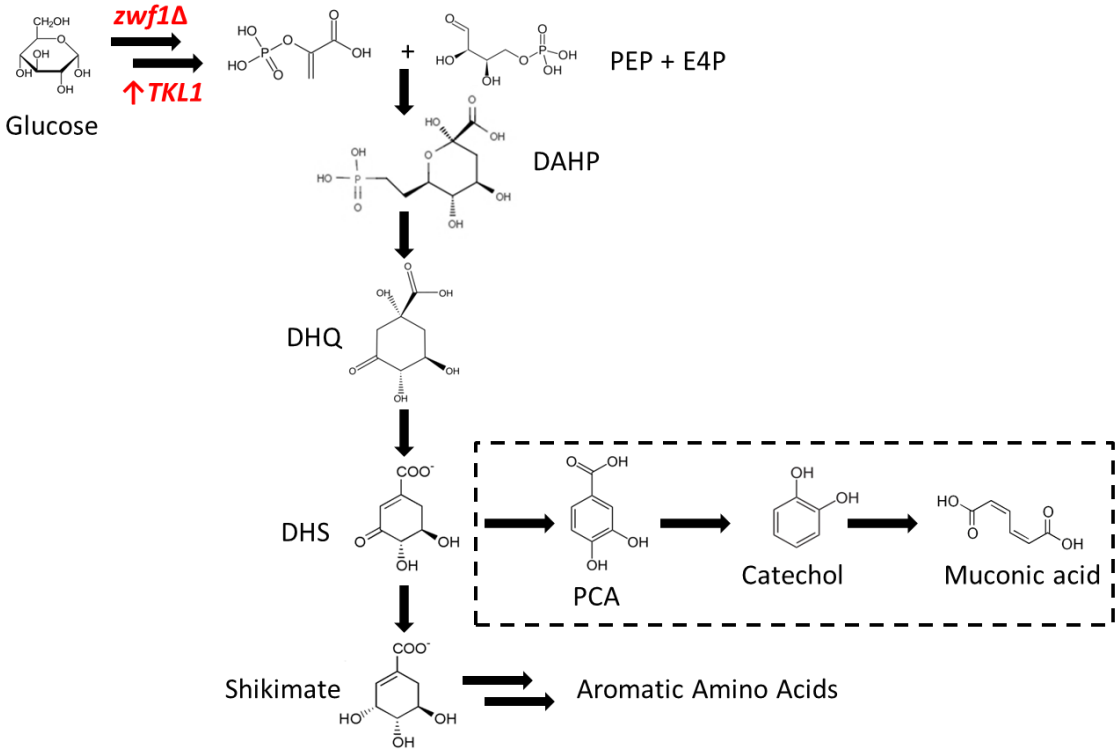


Figure 3. The synthetic pathway leading to muconic acid production.

A target protein for directed engineering is the TKL1 protein. This protein is involved in rerouting carbon flux from the glycolytic pathway into the pentose phosphate pathway (PPP) and Shikimate pathways when the Δzwf1 mutation prevents glucose-6-phosphate from flowing naturally into the PPP (Curran *et al.*, 2013). The native route of carbon into the PPP is through glucose-6-phosphate. With ZWF1 knocked out, muconic acid production was increased by 50% (Leavitt, 2013). TKL1 increases flux through the Shikimate pathway by maximizing carbon flux through the reverse transketolase reaction. This takes carbon from fructose-6-phosphate and glyceraldhyde-3-phosphate in the glycolytic pathway to erythrose-4-phosphate and xylulose-5-phosphate in the PPP. These molecules can then enter the Shikimate pathway. When the Δzwf1 mutation is present, adding extra copies of the TKL1 gene on a plasmid increased the production of muconic acid by 200% (Leavitt, 2013). Based on this experimental result, it can be hypothesized that additional mutations improving the efficiency of the TKL1 protein to further favor the reverse transketolase reaction could feasibly cause even larger gains in the production of muconic acid. Metabolic engineering protocols and methods are used to redirect the metabolism of yeast cells to produce the desired product. Directed evolution is one

of these tools in which we select mutations which provide improvements at the protein and genome level which lend themselves towards increasing the production of our desired product. Work on this project centers around developing a strain of *S. cerevisiae* that has high flux through the Shikimate pathway and would function as a good platform for production of muconic acid. After initial success producing muconic acid in yeast and subsequent modification which resulted in the highest reported yield through the Shikimate pathway, the limit for rational genetic manipulation was reached at an output of 140 mg/L muconic acid (Curran *et al.*, 2013).

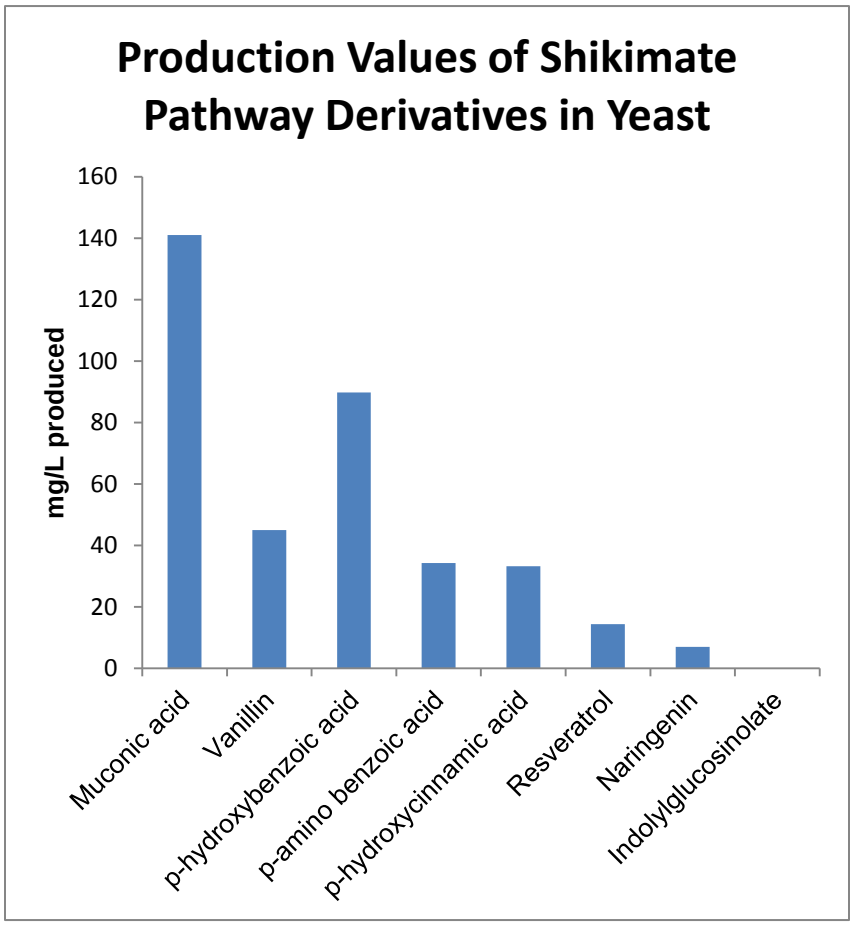


Figure 4. Titers of chemicals produced heterologously in yeast utilizing the Shikimate pathway. Obtained from Curran, Leavitt et al. Metabolic Engineering 2012

In order to further improve the production, non-rational techniques such as directed evolution and adaptive laboratory engineering must be used. To screen for strain improvements with the greatest muconic acid production, either improved growth rate or some other measurement of fitness that can be measured using a high throughput method must be employed (Leavitt, 2013). The high throughput method under development uses an inducible promoter to engineer a pathway. The promoter of ARO9 acts as an inducible

promoter in response to the amount of aromatic amino acids present, leading to increased gene expression through the transcription factor ARO80 (Iraqui, I., et al., 1998). As these aromatic amino acids are the downstream products of the Shikimate pathway, an increase in their presence will likely correlate with capacity to produce more muconic acid.

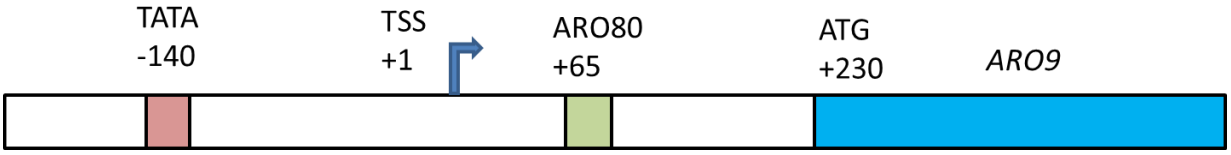


Figure 5. A detailed view of the ARO9 promoter sequence.

Since the engineering of more efficient pathways to synthesize muconic acid essentially depends on the development of a working inducible promoter system, the first step in optimizing this pathway is to develop an aromatic amino acid biosensor. This will allow mutations that increase flux through the Shikimate pathway to be more easily detected. Thus these mutations will be able to be accumulated and compounded as the directed evolution of proteins participating in this pathway is implemented.

To this end, examination of the ARO9 promoter will allow identification of the Upstream Activating Elements, essential for induction in response to aromatic amino acids. When these portions are combined with minimal promoter elements to drive the transcription of a fluorescent protein in response to aromatic amino acids, we will create an effective biosensor for measuring aromatic amino acid concentration. The sensor can then be used to detect positive mutations in pathway proteins using flow cytometry.

Results

Biosensor

It was hypothesized that within the upstream promoter region there was a specific upstream activating sequence (UAS) responsible for the transcriptional activation of ARO9 in response to aromatic amino acids. It was the goal of this experiment to cut down the full length sequence until the much smaller UASaro sequence was found, and then use this to drive induction to be used for the biosensor. Following a hybrid promoter approach, we generated an inducible promoter by combining the UAS element with a minimal promoter containing the core promoter elements including a TATA box. First, the 755 base pair sequence was cut down to the shorter 355 base pair sequence. This sequence was then further trimmed down until the final UASaro sequence was isolated. The UASaro sequence was placed in front of a minimal promoter and the Yellow Fluorescent protein, YEcitrine, to build the biosensor.

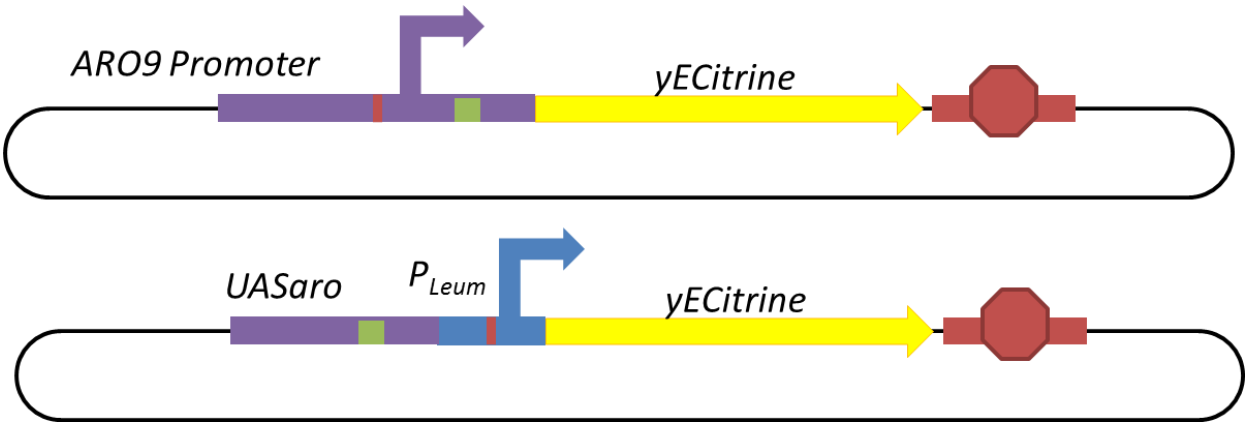


Figure 6. Modified hybrid promoter featuring only the UAS sequence pictured side-by-side with the original promoter for comparison. As noted above, the hybrid promoter sequence features the trimmed down sequence.

In order to further ramp up expression of the fluorescent protein, the hybrid promoter was further modified by stacking multiple UASaro sequences in front of the original. The new promoter was given the name 4xUASaro as a result of the 4 copies of the UASaro promoter it carried, as opposed to the original 1xUASaro hybrid promoter or the previous 355bp and 755bp ARO9 promoters. These promoters were cloned into a yeast expression vector and transformed into the BY4741 wild type yeast strain and maintained using an auxotrophic marker. YFP expression was assayed with various concentrations of spiked in tryptophan. It was discovered that despite being functional in other *S. cerevisiae* strains, the 755bp full length promoter was non-functional in BY4741.

The hybrid promoters did not initially appear to provide any additional benefit over the 355bp promoter. However after evaluation, it was determined that this was due to saturation of the ARO80 transcription factor. There simply weren't enough copies of this protein to bind to all of the available UASaro sites. Thus, more copies of the ARO80 gene were introduced into these strains on plasmids. Once the yeast strains were provided with additional copies of the ARO80 transcription factor, flow cytometry showed remarkable improvements in fluorescence between each of the 355bp, 1xUASaro, and 4xUASaro promoters over each other and over the original 755bp promoter. This demonstrated the sensitivity to aromatic amino acids of the promoter and shows that it was ready to be used as a selection factor for the engineering of proteins within the muconic acid pathway.

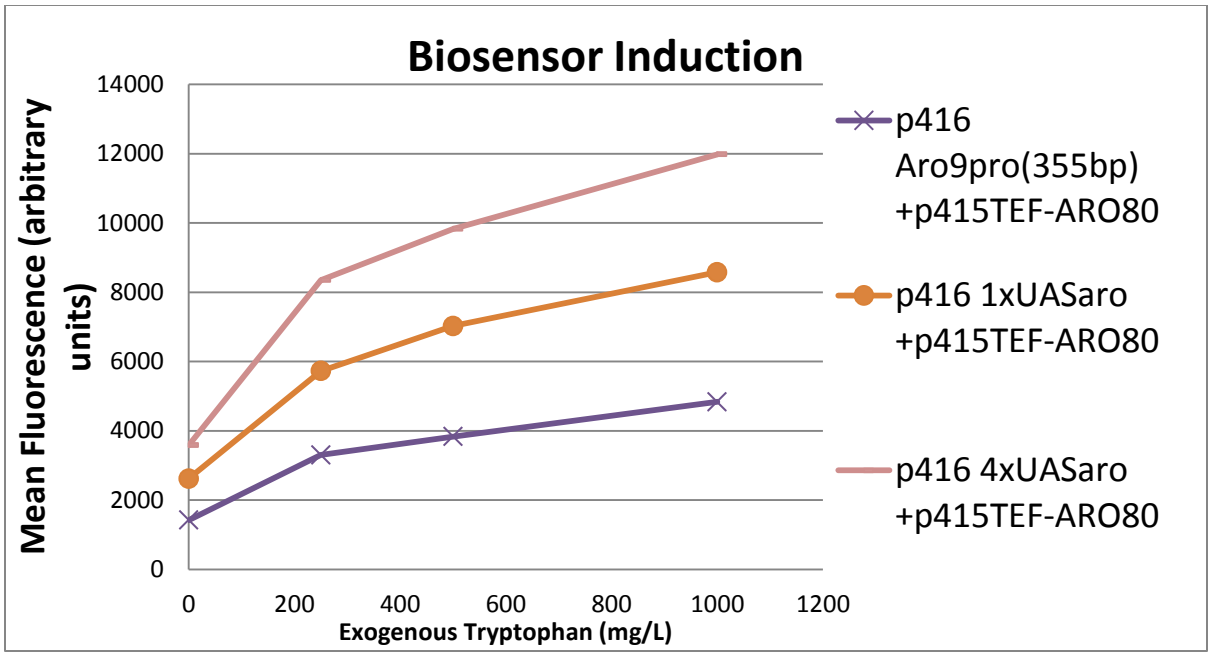


Figure 7. Biosensor induction of the engineered hybrid promoter can be seen. After addition of sufficient amounts of ARO80 transcription factor, the 4XUASaro hybrid promoter saw a threefold increase in induction as compared to the original unmodified promoter, when turned on by the presence of exogenous Tryptophan.

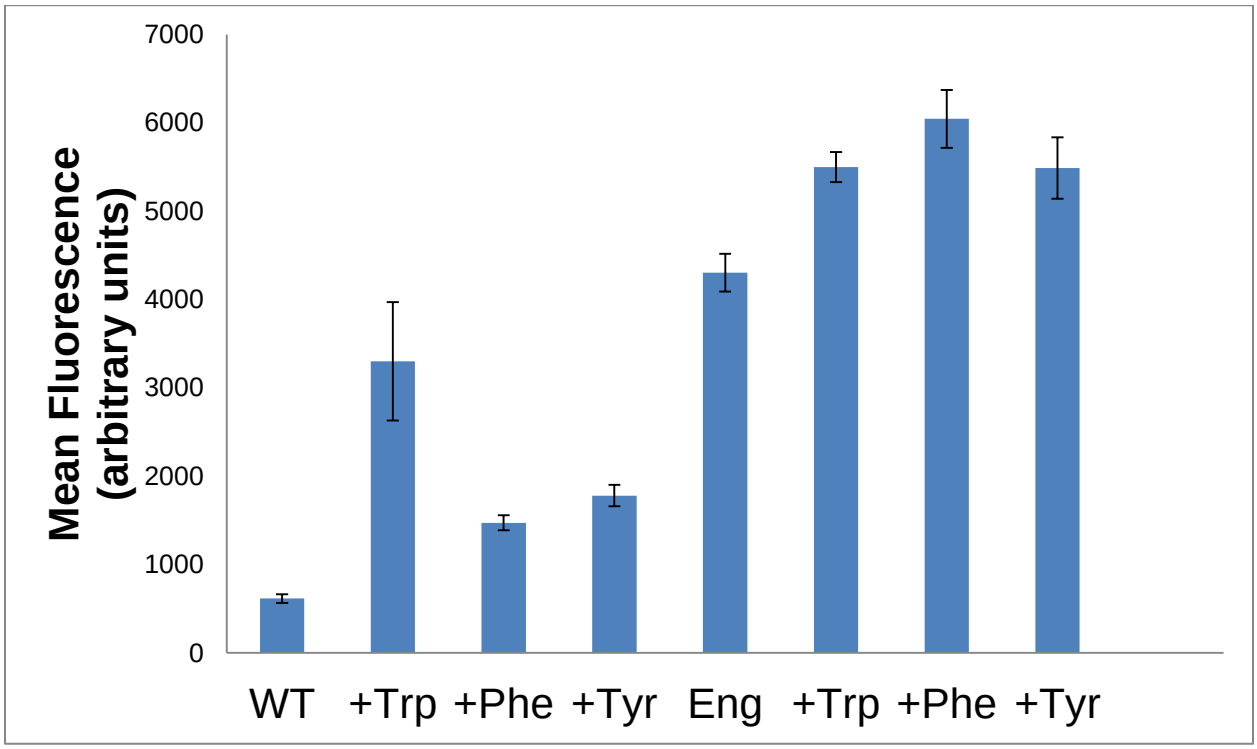


Figure 8. The biosensor was tested with previously engineered strain of yeast and proven to detect improvements in flux through pathway over the wild type yeast when induced with each of three different aromatic amino acids.

TKL1 Engineering

The TKL1 sequence was subjected to error prone PCR in order to generate a variety of versions with different mutations. These PCR products were cloned into yeast expression vectors to produce plasmid libraries of the mutated TKL1 gene. The TKL1 library DNA was then transformed into yeast and assayed. However, in order to properly screen these mutant libraries, they must be transformed into strains of yeast lacking the native ZWF1 and TKL1. The Δzwf1/Δtkl1 genetic background allows flux into the Shikimate pathway to be routed through the mutant TKL1 such that improvements will be easily detected.

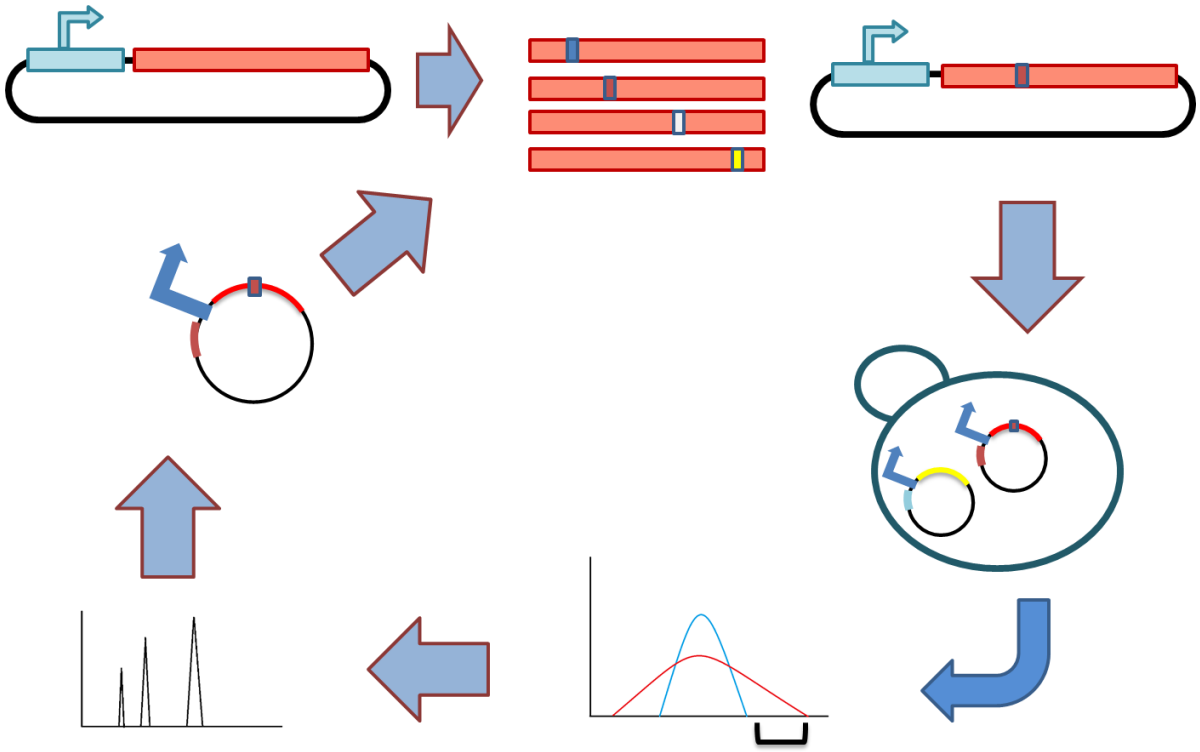


Figure 9. Visual schematic of the general method for protein engineering employing the biosensor. First, diversity in the target gene is generated using error-prone PCR. These libraries of modified copies of the gene are then transformed into yeast screening strains lacking the endogenous version of the gene. The cells are then screened via flow cytometry and improvements are quantified. This process was repeated iteratively as necessary.

These genes were knocked out using the "delete and repeat" method (Curran *et al.*, 2012). Cassettes containing the KanMX selectable marker blanked on both sides by LoxP sites were created via PCR so that they contained 40 homologous base pairs on either side of the target integration site. Following transformation of these DNA sequences, the strain was put under selection pressure for the KanMX marker. Each successive knockout could only be

created by removing the selectable marker, before a new knockout strain could be attained. This was achieved with the help of Cre-Recombinase, which worked alongside the LoxP sites to remove the selectable marker.

These libraries will be screened for improvements over the wild type using flow cytometry. Individual high fluorescence mutants were isolated with Fluorescently Activated Cell Sorting (FACS), sequenced, and differences in mutants were identified in relation to the wild type. This sequence of events beginning from the error prone PCR was repeated as necessary.

As can be seen from figure 11, the first completed round of sorted TKL1 mutants yielded a wide spread of results. Using spiked in tryptophan as an approximate measurement for improvements provided by the mutant TKL1 proteins, the fluorescence of the biosensor was measured. The fluorescence measured from those mutants on the far right of the graph suggest that pathway yields could be boosted by over five-fold with the best mutants from just this first round of mutagenesis.

Discussion and Conclusion

Engineering TKL1 is an extremely important step in the metabolic engineering of *Saccharomyces cerevisiae* to produce more muconic acid. TKL1, in its current state, creates a bottleneck in the enzymatic pathway from glucose towards Erythrose-4-phosphate, and thus is a kinetic barrier towards the production of muconic acid. Previous work has been done on engineering TKL1 for it's role in the PPP, however no one has studied it's role in the Shikimate pathway, nor was there a high throughput method to detect flux through this pathway. The native version of this protein directs flux in the wrong direction for our purposes. With the help of directed evolution techniques, the hope is to change specificity of this enzyme, modifying it to favor flux in the direction of the Shikimate pathway.

Erythrose-4-phosphate is a PPP product that is known to be the rate limiting step for flux through the Shikimate pathway. In order for carbon to be funneled into the PPP from glycolysis, it can utilize either the Glucose-6-phosphate dehydrogenase bridge catalyzed by the gene product of ZWF1, or the reversible reaction catalyzed by TKL1. It was predicted that the optimal route to increase flux towards the Shikimate pathway would be to knock out ZWF1 and optimize the TKL1 reverse reaction. Current kinetics reveal that wild type TKL1 has a 100-fold higher Km for the glycolytic intermediates than those of the PPP, indicating that this protein would make an excellent target for directed evolution.

Following general principles of kinetics, a build-up or excess of any intermediates along the pathway should push the equilibria towards the ultimate product. Thus, the more efficient the engineered TKL1 enzyme is, the greater the concentration of Erythrose-4-phosphate in the

cell will be, leading towards increased flux through the downstream pathway all the way towards muconic acid.

Current experimentation has led to development of a biosensor capable of detecting aromatic amino acids which are one of the downstream products of the engineered pathway. Mutant TKL1 genes have also been engineered, which when screened with the biosensor reveal an increase in induction potentially representative of a five-fold improvement in Shikimate pathway flux. This is the first step towards further engineering yielding even greater improvements in the metabolic pathway towards muconic Acid production.

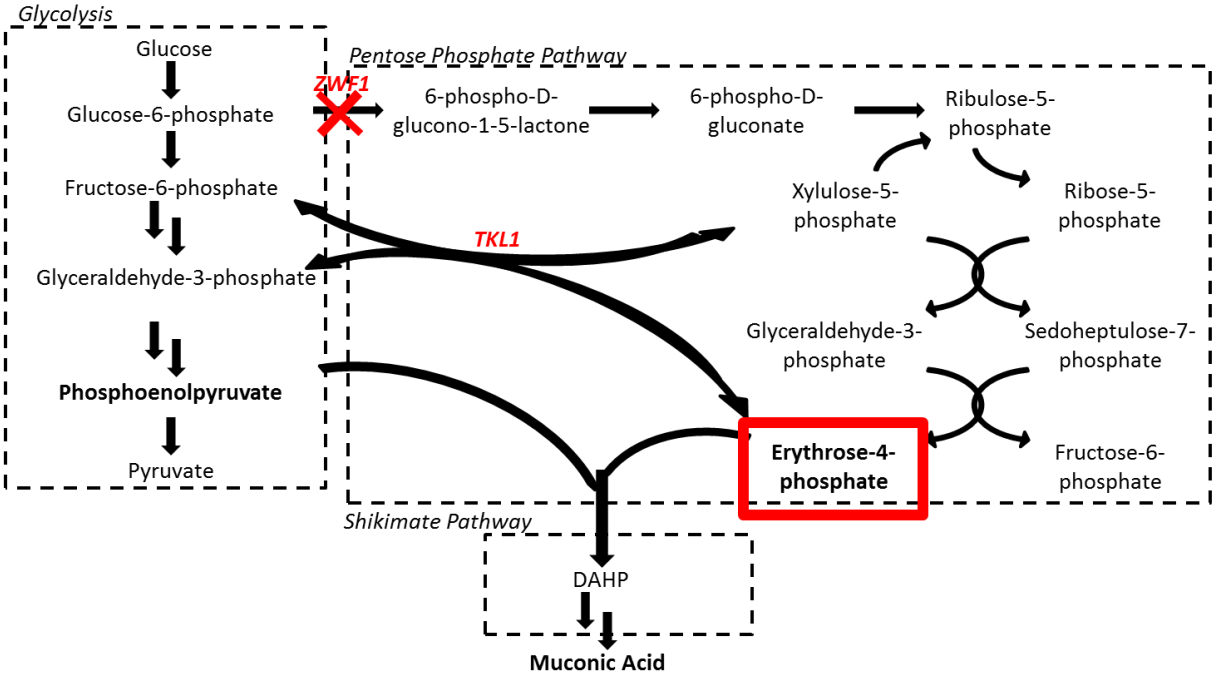


Figure 10. This schematic shows a more detailed view of the intermediate steps of figure 2, from Glucose to DAHP. With the increased funneling of intermediates towards E4P and DAHP, kinetics principles dictate this should push the pathway towards greater production of muconic acid.

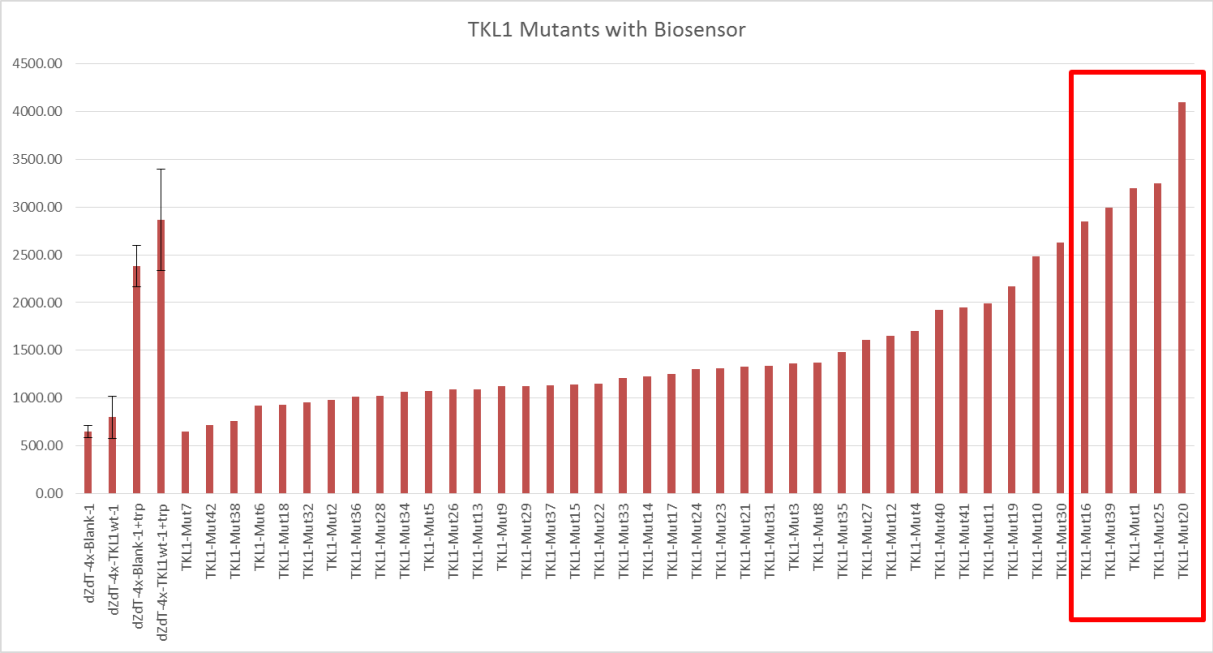


Figure 11. First round results from screening of TKL1 mutants. There is a non-linear relationship between induction and endogenous concentration. A 50% improvement in induction could represent a much higher improvement in flux through the pathway. Thus it will take further testing to determine the potential for muconic acid production with these first round mutants, and further rounds of mutagenesis to produce even more efficient mutant strains.

Proposals for Future Work

TKL1 mutant possessing improvements generated as outlined earlier will be confirmed using muconic acid fermentations. They will then be used for further rounds of mutagenesis and protein efficiency improvement. Another target protein for directed protein engineering is ARO1. The ARO1 gene encodes a protein responsible for five enzymatic steps in the biosynthesis of aromatic amino acids (Larimer, 1983). ARO1 engineering will be performed following the same concepts as that of TKL1, with one notable exception. Prior to subjection to mutagenesis, the ARO1 gene will be truncated. Only the portions of the enzyme which catalyze reactions leading up to DHS dehydrogenase will remain intact. In this way, once the wild type ARO1 gene has been knocked out and replaced, cells will have no way of utilizing DHS dehydrogenase to make aromatic amino acids. Instead, all of the built up DHS dehydrogenase will be funneled through the pathway pictured in figure 2, through PCA decarboyxlase towards muconic acid production.

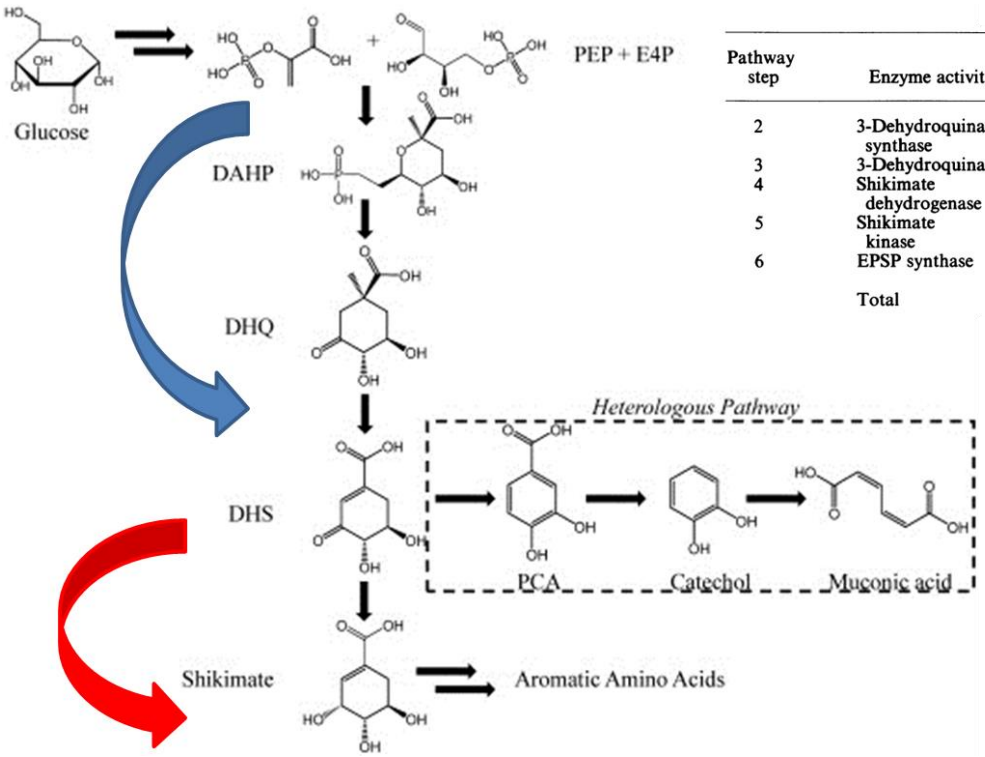
TABLE 1. Genes and enzymes for the biosynthesis of chorismate in *S. cerevisiae*

Reaction	Compound	Enzyme designation	Gene	Chromosome	mRNA length (kb)	Polypeptide size (kDa)[a]
	PEP + E4P					
1	↓	DAHP synthase (EC 4.1.2.15)	*ARO3*	IV	1.2	41 (370 aa)
			ARO4	II	1.4	39 (367 aa)
	DAHP					
2	↓	DHQ synthase (EC 4.6.1.3)	*ARO1C*	IV		175 (1,588 aa): aa 1–392
	DHQ					
3	↓	DHQ dehydratase (EC 4.2.1.10)	*ARO1E*	IV		aa 1059–1293
	DHS					
4	↓	DHS dehydrogenase (EC 1.1.1.25)	*ARO1D*	IV		aa 1306–1588
	Shikimate					
5	↓	Shikimate kinase (EC 2.7.1.71)	*ARO1B*	IV		aa 886–1059
	Shikimate 3-phosphate					
6	↓	EPSP synthase (EC 2.5.1.19)	*ARO1A*	IV		aa 404–886
	EPSP					
7	↓	Chorismate synthase (EC 4.6.1.4)	*ARO2*	VII	1.4	
	Chorismate					
	↓					
	Phe, Tyr, Trp, ubiquinone, *p*-aminobenzoate, vitamin K					

Figure 12. This figure from Braus 1991 shows the reactions catalyzed by the wild type pentafunctional ARO1 enzyme. The point at which the gene will be truncated is pictured highlighted in red. The first three reactions will be left intact, only the last two will be trimmed.

Once a functional truncation has been generated, it will enter the mutagenesis cycle proposed by figure 9. For screening confirmation purposes, the truncation will be screened first with a strain possessing a functional AROE gene from *E. coli*. If the cells are able to successfully grow on nutrient restricted medium lacking aromatic amino acids, this will confirm that the truncated gene is functional.

It is anticipated that once mutants of both the ARO1 and TKL1 genes are generated, they will be placed together in the same strain which will have a markedly improved efficiency in muconic acid production. Together with the other modifications of the strain, increased flux through the pathway should flow towards the ultimate product.



Pathway step	Enzyme activity	*E. coli* gene
2	3-Dehydroquinate synthase	*aroB*
3	3-Dehydroquinase	*aroD*
4	Shikimate dehydrogenase	*aroE*
5	Shikimate kinase	*aroL*
6	EPSP synthase	*aroA*
	Total	

Figure 13. Concepts to be employed from ARO1 truncation cloning and screening are pictured in this map created with the aid of Duncan 1987. The blue arrow signifies functions that the functional truncation will have the ability to catalyze while the red arrow represents pathways that the *E. coli*. AROE will catalyze. If a strain expressing both of these genes is able to grow on nutrient restricted media lacking aromatic amino acids, it will confirm that the ARO1 truncation is functional.

Acknowledgements

I would like to start by thanking John Leavitt for all of his hard work in mentoring, advising, and supervising my work in the Alper Laboratory. His patience and endless willingness to teach and answer questions inspired me to fully dedicate myself to the project at hand. In addition, his tireless editing and encouragement were very much appreciated.

Very special thanks also goes out to Dr. Hal Alper, not only for allowing me to join the lab as a freshman, but also for serving as a fantastic advisor and mentor. Without his help I would not have had the means or access to pursue the incredible opportunities that were presented to me. In addition, I would like to express my gratitude to all the members of the Alper lab, for their endless patience, guidance, and support.

Lastly, I would like to thank my parents, friends, and past supervisors. Without your support and care I would not have become the person I am today, for that I am grateful.

Works Cited

Braus, G H. “Aromatic Amino Acid Biosynthesis in the Yeast Saccharomyces Cerevisiae: A Model System for the Regulation of a Eukaryotic Biosynthetic Pathway.” *Microbiological Reviews* 55.3 (1991): 349–370. Print.

Burridge, E., 2011. Adipic acid. *ICIS Chem. Bus.* 279 43-43.

Curran, Kathleen A., and Hal S. Alper. "Expanding the chemical palate of cells by combining systems biology and Metabolic engineering. "*Metabolic Engineering.* 14 (2012): 289-97. Print.

Curran, Kathleen A., John M. Leavitt, Ashty S. Karim, and Hal S. Alper. "Metabolic Engineering of muconic acid production in *Saccharomyces cerevisiae*." *Metabolic Engineering.* 15 (2013): 55-66. Print.

Draths, K.M., Frost, J.W. “Environmentally compatible synthesis of adipic acid from D-glucose.” *J. Am. Chem. Soc.* 116(1994): 399–400.

Duncan, K, R M Edwards, and J R Coggins. “The Pentafunctional Arom Enzyme of Saccharomyces Cerevisiae Is a Mosaic of Monofunctional Domains.” *Biochemical Journal* 246.2 (1987): 375–386. Print.

Ikeda, M. “Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering.” *Applied microbiology and biotechnology*. 69(2006): 615-26.

Iraqui, I. , et al*. Molecular and cellular biology.* 1999.

Iraqui, I., et al. “Characterization of Saccharomyces cerevisiae ARO8 and ARO9 genes encoding aromatic aminotransferases I and II reveals a new aminotransferase subfamily.” *Molecular & general genetics : MGG.* 257(1998): 238-48.

Larimer, F.W., Morse, C.C., Beck, A.K., Cole, K.W., Gaertner, F.H. “Isolation of the ARO1 Cluster Gene of Saccharomyces cerevisiae.” *Molecular and Cellular Biology.* 3(1983): 1609-14.

Leavitt, John M. "Aromatic Amino Acid Biosensor Development and Utilization in S. Cerevisiae." Thesis Proposal. University of Texas at Austin, 2013. Print.

Mirasol, F., 2011. PTA. *ICIS Chem. Bus.* 279 43-43.

Niu, W., Draths, K.M., Frost, J.W. “Benzene-free synthesis of adipic acid.” *Biotechnol. Prog.* 18(2002): 201–211.