Carbon Catabolite Repression Relaxation: Approaches for Sugar Co-Utilization in *Escherichia coli*

By

Kevin J. Fox

Bachelor of Chemical Engineering
University of Minnesota-Twin Cities, 2017
Master of Science in Chemical Engineering Practice
Massachusetts Institute of Technology, 2019

Submitted to the Department of Chemical Engineering
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemical Engineering Practice
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 2022
© 2022 Massachusetts Institute of Technology.
All rights reserved

Signature of Author ....

Kevin J. Fox
Department of Chemical Engineering
May 13, 2022

Certified by ....

Kristala L. J. Prather
Professor of Chemical Engineering
Thesis Supervisor

Accepted by ....

Patrick S. Doyle
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Carbon Catabolite Repression Relaxation: Approaches for Sugar Co-Utilization in *Escherichia coli*

By
Kevin J. Fox

Submitted to the Department of Chemical Engineering
on May 13, 2022, in Partial Fulfillment of the
requirements for the Degree of Doctor of Philosophy in
Chemical Engineering Practice

Abstract

Bioprocessing provides a sustainable, renewable, and green alternative to petroleum-based methods for the production of chemicals. A primary reason for the benefit of bioprocessing is its ability to utilize waste materials containing sugars as a feedstock, such as those from agriculture. This advantage is complicated with *Escherichia coli* due to carbon catabolite repression (CCR), an intrinsic sugar preference system. The relaxation of the effects of CCR has the possibility to increase bioprocessing's economic viability and sustainability through better feedstock utilization. In this work we examine reported strategies for the relaxation of CCR and describe novel methods for the utilization of sugar mixtures for the production of an industrially relevant chemical.

A microbial production platform was developed to synthesize enantio-pure D-glyceric acid, a chemical with potential use in the materials industry, from D-galacturonate. The expression of *udh* from *Pseudomonas syringae* and *gli* from *Agrobacterium fabrum*, along with the inactivation of *garK*, encoding for glycerate kinase, enables D-glyceric acid accumulation by utilizing the endogenous expression of *garD*, *garL*, and *garR*. Optimization of carbon flux through the elimination of competing metabolic pathways led to the development of a *ΔgarKΔhyiΔglxKΔuxaC* mutant strain that produced 4.8 g/l of D-glyceric acid from D-galacturonate, with an 83% molar yield. Additionally, a substrate-based induction platform was developed that enabled the expression of *udh* and *gli* upon the addition of D-galacturonate by utilizing the transcription factor ExuR from *Bacillus subtilis*, eliminating the need for chemical induction.

Two strategies for CCR relaxation were investigated; one employing a global alleviation strategy and the other a sugar-specific strategy. A mutation in $EIIA^{glc}$, an essential part of the phosphoenolpyruvate transferase system (PTS), was investigated as a global CCR relaxation strategy due to its ability to lock the protein in its phosphorylated state, mimicking a lack of glucose. While this engineered strain did co-utilize sugar mixtures, this phenotype was not stable. To enable sugar-specific CCR relaxation, the galacturonate-specific permease ExuT was engineered to lessen the effects of inducer exclusion. A galacturonate-specific biosensor was utilized to perform high-throughput screening of an *exuT* mutant library to search for mutants that enabled higher levels of intracellular galacturonate. Utilizing the synthetic pathway to produce D-glyceric acid from galacturonate, a S391R mutant of ExuT increased titer by 20% when a co-feed of galacturonate and glucose was used.

An analysis of the opportunity for synthetic biology to disrupt the specialty chemicals industry was performed. Synthetic biology firms should leverage their capabilities to utilize sustainable feedstock and synthesize novel products to differentiate and limit commoditization. Competitive landscape analysis displayed the relative success of the differing strategies of current players.

Acknowledgements

Firstly, I would like to thank Kris for her support and mentorship throughout my time here at MIT. Thank you for many exciting, invigorating, and helpful discussions, my success as a researcher has relied heavily on your advice and experience. Additionally, thank you for supporting my career goals and ensuring that my experience as part of the PhD CEP program was a positive and enriching one. I would also like to thank my thesis committee members, Professors Gregory Stephanopoulos and Anthony Sinskey, for their support.

I would like to extend my thanks and appreciation to all members of the Prather Lab, past and present. Thank you to Jason for always being willing to hear my crazy ideas during the first year of research and allowing me to tap into your vast experience whenever I had a question. Kristina and Wesley, thank you for bringing additional expertise and perspective to the lab, a quick conversation with either of you always gave me new hope and inspiration. Lisa and Stephanie, thank you for bringing a sense of comradery to the lab and greeting me warmly as a member of community from day one. Thank you Christina for all your help on random cloning questions, tagging along to the Vikings bar, and sharing great UMN ChemE stories. Thank you Cynthia and Jennifer for your light-hearted attitudes and consistent advice and friendship which brightened my days even when my cloning didn't work for what seemed to be the millionth time. K'yal, thank you for the great conversations we had in the office about the world and our career goals. Bernie and Vincent, thank you for your incredibly helpful comments and ideas during group meeting and bringing cool and interesting data for us to enjoy every time.

Thank you to all my friends and classmates that I have met during my time at MIT. Your support helped get me through some of the toughest parts of this program. Additionally, thank you to my Practice School cohort, working with all of you closely was an incredible experience.

To all my mentors during all of my undergraduate research experiences, thank you for culturing my love for science and showing me how cool the field of metabolic engineering is. Professor Zhang, thank you for accepting me into your lab and trusting me enough to both mentor other students as well as spearhead my own projects. Yi-Shu and Qiuge, thank you for being supportive, intelligent, and friendly supervisors. I owe much of my current work ethic, lab protocol, and experimental mindset to your interest in me growing as a researcher.

I would like to extend a special thanks to members of the ChemE Department, Boston at large, and MIT Sloan who played squash with me during my PhD research. Having the honor to teach and learn from so many athletes from diverse backgrounds during the past few years was an incredible experience.

Last, but not least, I would like to thank my family for their love and support. Thank you to my parents for acting interested in my research when I explain it to them and always pushing me to pursue my goals, no matter the obstacles in the way. Thank you to my brother, Kyle, for providing fun insight on the random conversations that we have when I am at home. To my amazing girlfriend Jillian, thank you so much for always being by my side and being genuinely interested in my research. Our excursions around Boston and trips to random parts of New England provided me so much joy and couldn't have asked for a better adventure buddy.

-Kevin Fox

Contents

1. Introduction 17

1.1 Metabolic Engineering for Value-Added Product Formation 18

1.2 Carbon Utilization in *Escherichia coli* 19

1.3 Carbon Catabolite Repression in *Escherichia coli* 20

1.4 Global Carbon Catabolite Repression Relaxation 22

1.4.1 PTS Modifications 23

1.4.2 Co-Culture Applications 24

1.4.3 CRP Modifications 25

1.4.4 Other Genetic Modifications 26

1.5 Sugar-Specific Carbon Catabolite Repression Relaxation 28

1.5.1 Xylose-Glucose Co-Utilization 28

1.5.2 Other Sugar Co-Utilizations 29

1.6 Biosensors for High-Throughput Screening of Mutant Libraries 31

1.6.1 Biosensors for Metabolic Engineering 31

1.6.2 Application to High-Throughput Screening 33

1.7 Thesis Objectives 35

1.8 Thesis Organization 36

2. Production of D-Glyceric Acid from D-Galacturonate in *Escherichia coli* 37

2.1 Introduction 39

2.2 Materials and Methods 43
2.2.1 E. coli Strain and Plasmid Construction 43
2.2.2 Culture and Analysis Conditions for D-glyceric acid Production 44
2.2.3 Quantification of mRNA Levels for Expression Platform 45
2.2.4 D-glyceric acid Purification 46
2.3 Results and Discussion 50
2.3.1 D-glyceric acid Production using Novel Pathway and Strain Optimization for Titer and Molar Yield from D-galacturonate 50
2.3.2 Carbon Source Co-feed Analysis 54
2.3.3 Feedstock-based Expression of D-Glyceric Acid Production Pathway 57
2.3.4 Purification of D-Glyceric Acid from Spent Fermentation Medium 62
2.4 Conclusions 65
3. CCR Relaxation Strategies for Global and Sugar-Specific Alleviation 67
3.1 Introduction 69
3.1.1 Global Alleviation through the Mutation of $EIIA^{glc}$ 71
3.1.2 Sugar-Specific Alleviation through the Directed Evolution of a Galacturonate Permease 71
3.2 Materials and Methods 76
3.2.1 E. coli Strain and Plasmid Construction 76
3.2.2 Culture Conditions for Mutant $EIIA^{glc}$ Strain Fermentation of Sugar Mixtures 77

3.2.3 Directed Evolution Protocol for ExuT 82

3.2.4 Culture Conditions for D-Glyceric Acid Production 83

3.2.5 Analysis Conditions for Sugar, D-Glyceric Acid, and Fluorescence Detection 83

3.3 Results and Discussion 85

3.3.1 Global Alleviation through the Mutation of $EIIA^{glc}$ 85

3.3.2 Sugar-Specific Alleviation through the Directed Evolution of a Galacturonate Permease 95

3.4 Conclusions 109

4. Market Opportunity of Synthetic Biology in the Specialty Chemicals Industry 111

4.1 Introduction 113

4.2 Market Attractiveness of the Specialty Chemicals Industry 116

4.2.1 Profitability and Volatility 117

4.2.2 Industry Size and Growth 118

4.2.3 Total Addressable Market 119

4.3 Market Dynamics of the Specialty Chemicals Industry 121

4.3.1 Buyer Power 122

4.3.2 Supplier Power 122

4.3.3 Industry Rivalry 123

4.3.4 Substitutes 124

4.3.5 New Entrants 124

4.4 Opportunity for Synthetic Biology to Disrupt the Specialty Chemicals Industry 126

4.4.1 Benefits of Synthetic Biology Application 128

4.4.2 Challenges Still Needing to be Addressed for Synthetic Biology 131

4.5 Competitive Landscape of Synthetic Biology Firms 134

4.6 Conclusions 138

5. Conclusions and Outlook 139

5.1 Summary and Future Directions 140

5.2 Outlook 143

5.2.1 Metabolic Engineering for the Production of Sustainable Monomers 143

5.2.2 Utilization of Sustainable Feedstock for Chemical Production 144

5.2.3 Novel Engineering Approaches to Carbon Catabolite Repression Alleviation 145

5.2.4 Biosensors as Tools for High-Throughput Screening 146

References 147

List of Figures

Figure 1-1. Schematic of PTS-facilitated glucose transport and its effect on the catabolism of secondary sugars 21

Figure 1-2. Definition of global and sugar-specific CCR relaxation engineering strategies 22

Figure 1-3. Number of publications containing keywords 'biosensor' and 'metabolic engineering' from a Google Scholar search 33

Figure 1-4. Throughput of mutants tested depending on screening method utilized 34

Figure 2-1. Designed pathway for the production of D-glyceric acid from D-galacturonate in E. coli 41

Figure 2-2. Full D-glyceric acid pathway from galacturonate 51

Figure 2-3. Production of D-glyceric acid by different MG1655(DE3) mutant strains 52

Figure 2-4. Production of D-glyceric acid from MG1655(DE3)Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC* (pRSFD_*udh_gli*) with various secondary sugar co-feeds 56

Figure 2-5. Optical density (at 600 nm) measurements for the production of D-glyceric acid in minimal medium from 5 g/l of galacturonate co-fed with 5 g/l of each specified secondary carbon source. 57

Figure 2-6. Expression platform for the activation of the D-glyceric acid pathway by D-galacturonate 58

Figure 2-7. Fermentation data from the production of D-glyceric acid from D-galacturonate using different platforms for the expression of gli and udh 60

Figure 2-8. Extraction procedure for the enrichment of D-glyceric acid from spent fermentation medium. 63

Figure 2-9. Analysis of D-glyceric acid solution after purification procedure 64

Figure 3-1. Effect of glucose on the action of a galacturonate-specific biosensor 73

Figure 3-2. Schematic of the procedure for the directed evolution of exuT to increase galacturonate transport in the presence of glucose. Random mutagenesis is performed on the wildtype exuT gene to generate a mutant library of exuT genes 75

Figure 3-3. Sugar consumption trends of co-feeds of glucose and glycerol with galacturonate and lactose by wildtype E. coli 86

Figure 3-4. Fermentation of glucose-galacturonate co-feed by strain expressing H75Q mutant $EIIA^{glc}$ 87

Figure 3-5. Fermentation of glucose and galacturonate co-feed by adapted strain expressing H75Q mutant $EIIA^{glc}$ 88

Figure 3-6. Fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing H75Q mutant $EIIA^{glc}$. 90

Figure 3-7. Fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing H75Q mutant $EIIA^{glc}$ that was plated either on LB or MOPS agar 92

Figure 3-8. Fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing H75Q mutant $EIIA^{glc}$ plated on MOPS agar and expanded in LB or MOPS medium. 93

Figure 3-9. Clonal variation shown in fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing mutant $EIIA^{glc}$ plated on MOPS agar and then expanded in MOPS medium 94

Figure 3-10. GFP response upon the addition of galacturonate to different strains transformed with the galacturonate biosensor..................... 96

Figure 3-11. GFP distribution profiles for first ExuT mutant library sorting procedure, using flow cytometry 98

Figure 3-12. GFP analysis of isolated clones done in a BioLector from the sorting experiment of the first mutant ExuT library........ 99

Figure 3-13. Relief of effect transcriptional inhibition on the molar yield of D-glyceric acid when utilizing plasmid-based expression of exuT and co-feeding glucose and galacturonate by overexpressing endogenous pathway genes........ 101

Figure 3-14. D-glyceric acid titer from the plasmid-based expression of different ExuT mutants with the feed of only galacturonate......... 102

Figure 3-15. D-glyceric acid titer and molar yield from the plasmid-based expression of different ExuT mutants isolated from the first mutant library with a feed of glucose and galacturonate.. 103

Figure 3-16. D-glyceric acid titer from the genome-based expression of wildtype and S391R ExuT with a feed of glucose and galacturonate........ 104

Figure 3-17. GFP distribution profiles for second ExuT mutant library sorting procedure, detected using flow cytometry 106

Figure 3-18. GFP analysis of isolated clones done in a BioLector from the sorting experiment of the second mutant ExuT library........ 107

Figure 3-19. D-glyceric acid titer from the plasmid-based expression of different ExuT mutants isolated from the second mutant library with a feed of glucose and galacturonate........ 108

Figure 4-1. Graphical representation of the differentiating factors between commodity and specialty chemicals 113

Figure 4-2. Make-up of the specialty chemical industry in 2019 and its positioning in relation to the global chemical industry 115

Figure 4-3. Global market for biorefinery products by bioproduct category, 2019-2025 127

List of Tables

Table 1-1. Summary of global carbon catabolite repression relaxation strategies discussed. 27

Table 1-2. Summary of sugar-specific carbon catabolite repression relaxation strategies discussed. 30

Table 2-1. Primer sequences used in vector preparation and gene mutation procedures............ 47

Table 2-2. Sequence of gblock encoding for gli from Agrobacterium fabrum that was used in the creation of the pRSFD_udh_gli vector. 48

Table 2-3. Plasmids used in this study. 48

Table 2-4. Strains used in this study. 49

Table 2-5. Fraction of galacturonate and secondary carbon source consumed, titer and molar yield of D-glyceric acid in minimal medium experiment 57

Table 3-1. Primer sequences used in genetic manipulation procedures..................................... 79

Table 3-2. Plasmids used in this study. 81

Table 3-3. Strains used in this study. 81

Table 4-1. Summary of analysis of the attractiveness of the specialty chemicals industry. 116

Table 4-2. Summary of analysis of Porter's five forces on the specialty chemicals industry. .. 121

Table 4-3. Selection of synthetic biology companies in the specialty chemicals industry. 135

1. Introduction

Portions of this chapter have been published in:

Fox, K.J., and Prather, K.L.J. Carbon Catabolite Repression Relaxation in *Escherichia coli*: Global and Sugar-Specific Methods for Glucose and Secondary Sugar Co-Utilization. *Curr. Opin. Chem. Eng*. 2020. **30**: p. 9-16.

1.1 Metabolic Engineering for Value-Added Product Formation

For thousands of years, the ability of microbes to produce chemicals of interest has been harnessed through the fermentation process to generate beer and other food products. Pioneers in this area did not fully understand why their food or drink tasted better, but they continued to promote microbial processes, like the conversion of glucose to ethanol by *Saccharomyces cerevisiae,* through primitive cell-culture techniques [1]. In the centuries to follow, the mysteries behind these phenomena were explained through the discovery of microbes and the study of their central metabolic pathways. More recently, the advent of recombinant DNA technology [2] has enabled genetic engineers to express heterologous proteins using a biological host. The ability to control cellular machinery at an unprecedented level led to the development of the field of metabolic engineering, which aims to restructure metabolic networks to improve the production of molecules or proteins of interest [3]. This vision is accomplished by identifying and improving rate-limiting steps in metabolic pathways, expressing non-native enzymes to enable novel metabolic channels, and modifying endogenous regulatory networks [4, 5].

The modern metabolic engineer has many tools available for the development of synthetic expression systems. These include databases of known genes (NCBI[6]), pathways and reactions (KEGG[7]), and proteins (Uniprot[8]). Additionally, computational methods for the effective design of parts are also available (RBS calculator[9]). These resources enable the creation of multi-level control systems on the transcriptional, translational and post-translational levels to optimize production of compounds of interest [10]. Metabolic engineering strategies have enabled the development of bioprocessing as a platform to produce numerous commodity chemicals and monomers that would otherwise have been synthesized using petroleum-based chemistry [11]. Impressively, tools like these have also led to success in ‘pressure test’ situations where platforms

for the production of 6 out of 10 desired chemicals, not previously synthesized via bioprocessing, were developed in a span of 90 days [12]. This drill-like procedure displays the incredible power of the modern toolkit available to metabolic engineers for the quick development of bioprocesses to produce desired chemicals.

Further optimization of bioprocessing techniques for the production of biofuels [13] and specialty chemicals [14] provides the opportunity for greener and more sustainable industrial practices. This is due to bio-based methods frequently requiring less energy input in comparison to chemical processes as well as the ability for biological systems to utilize renewable resources, such as sugars from agricultural waste, as a feedstock. However, many challenges are still present that prevent the adoption of bioprocessing including low petroleum prices, lower effectiveness compared to chemical processes, and limited technologies currently available for the utilization of low cost raw cellulose [15].

1.2 Carbon Utilization in *Escherichia coli*

The economic efficacy and sustainability of bioprocessing as a mode for chemical synthesis is partially dependent on the ability to utilize waste materials containing sugars and other carbon sources, such as those from agriculture, as a feedstock. *Escherichia coli (E. coli)*, a model organism commonly used for bioprocessing, has endogenous uptake and catabolism channels to utilize these sources and can be genetically altered using metabolic engineering techniques to produce value-added products from them. Current production pathways have predominately been designed and optimized for a pure single-sugar feedstock. Realistically, however, raw biomass is composed of cellulose, hemicellulose and pectin which themselves have subunits of multiple different sugars and sugar acids including glucose, xylose, arabinose, and galacturonic acid among others [16]. Technologies for the breakdown of these complex polymers into their monomeric subunits have

been studied intensively and allow for the isolation of free sugar mixtures from biomass [17]. Having a diverse feed, however, presents feedstock utilization challenges in *E. coli* due to carbon catabolite repression (CCR), an intrinsic substrate preference mechanism found in many organisms. Specifically in *E. coli*, CCR is commonly materialized in the preference of consumption of glucose over other 'secondary' sugars, such as xylose and arabinose. Metabolic engineers have a vested interest in developing methods to alleviate the effects of CCR and allow for non-discriminatory sugar utilization due to the large potential of engineered systems with mixed sugar feedstocks [18].

1.3 Carbon Catabolite Repression in *Escherichia coli*

In *E. coli*, CCR is primarily realized through the phosphoenolpyruvate transferase system (PTS). The PTS enables the transport of glucose into the cell via a membrane protein complex, EIIB/C, encoded by *ptsG* (Figure 1.1). Upon uptake, a phosphate is transferred from an $EIIA^{glc}$ protein, encoded by *crr*, to glucose to promote sugar sequestration within the cell. The $EIIA^{glc}$ protein itself receives a phosphate from a signal cascade that originates at phosphoenolpyruvate. Extensive studies have been done on the interaction of the PTS with non-glucose sugar transport and catabolism systems [19]. The PTS influences catabolism protein transcription through cyclic AMP receptor protein (CRP) mediated activation. Phosphorylated $EIIA^{glc}$ (P-EIIA) activates adenylyl cyclase, initiating the cyclization of ATP into cyclic AMP (cAMP). This enables the formation of a cAMP-CRP complex which activates the transcription of sugar catabolism genes [20]. Non-PTS (secondary) sugar transport is affected by the PTS through the inducer exclusion phenomenon [21]. Non-phosphorylated $EIIA^{glc}$ (nP-EIIA) interacts with non-PTS sugar transporters, inhibiting their activity. In general, the ratio of P-EIIA to nP-EIIA is large when glucose is not present and small when glucose is present due to its phosphate being transferred to the transported glucose. In

reality, CCR is much more complicated than this simple description, with regulatory interactions between transcription factors associated with the genes of non-PTS sugars, which creates an intrinsic hierarchy of substrate preference in *E.coli* [22-24]. Due to this complexity, studies into understanding CCR through the lenses of cellular resource allocation and computation have been suggested [25, 26]. These modes of analysis have enabled a better understanding of the sugar preference strategy of *E.coli*, but the mechanistic functions of the PTS are still being resolved. This is apparent in recent discoveries such as the functionality of the glucose-based repression of the mannitol operon [27].

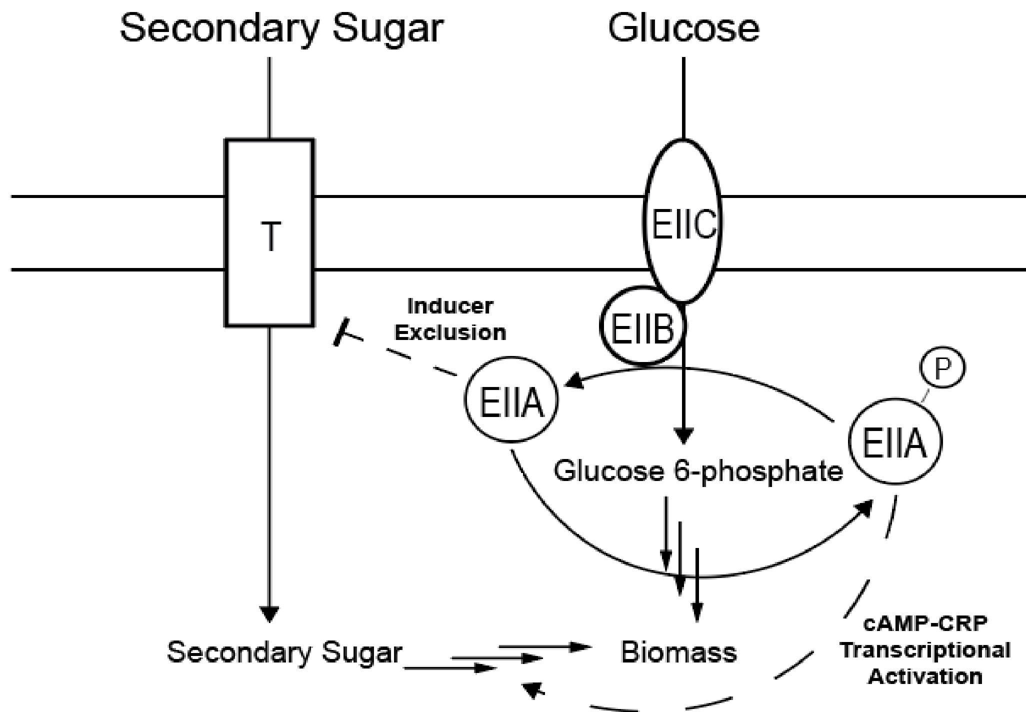

Figure 1-1. Schematic of PTS-facilitated glucose transport and its effect on the catabolism of secondary sugars. Glucose is transported into the cell through the EIIB/C membrane protein complex. Upon being transported, a phosphate is translocated to the glucose from a P-EIIA protein. nP-EIIA is phosphorylated by a signal cascade starting at phosphoenolpyruvate, a metabolite in glycolysis. Secondary sugar transport through membrane proteins (T) is inhibited through the inducer exclusion interaction of nP-EIIA with T. Secondary sugar catabolism is activated through cAMP-CRP-mediated transcriptional activation, initiated by the presence of P-EIIA.

Due to the evolving understanding of CCR and the large potential benefit for bioprocessing by enabling mixed sugar feedstock co-utilization, novel interventions to alleviate its effect are constantly being sought after. Often, these solutions are prescribed to release the cell from the transcriptional regulation mode of repression, from inducer exclusion or from both. This thesis will differentiate approaches towards alleviating CCR as targeting either cell-wide signaling systems (global CCR relaxation) or sugar-specific signal cascades (sugar-specific CCR relaxation) (Figure 1.2).

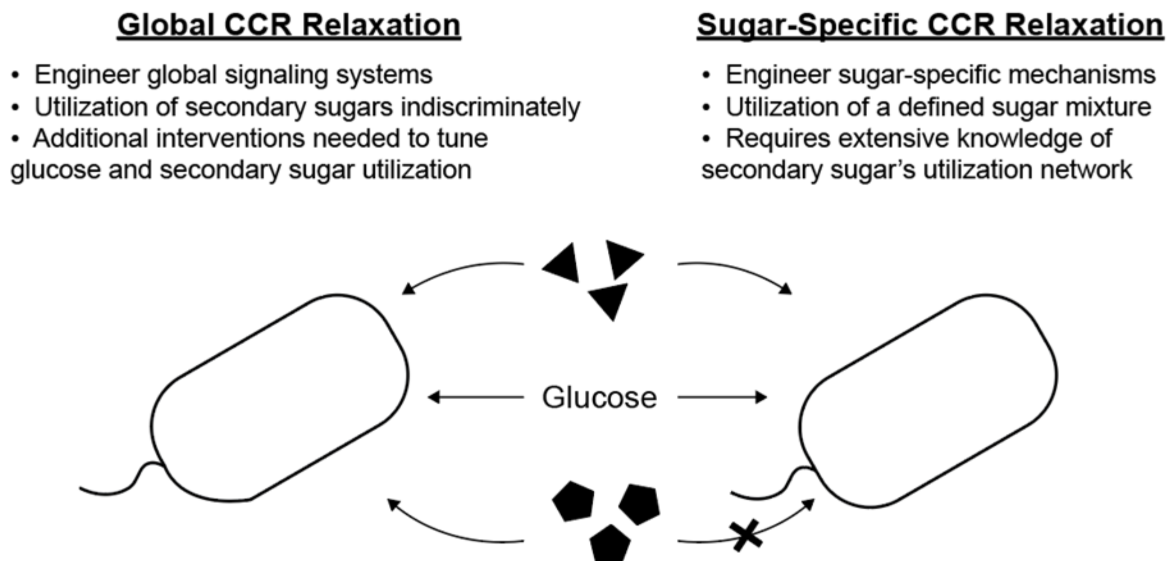

Figure 1-2. Definition of global and sugar-specific CCR relaxation engineering strategies. Global strategies necessitate the engineering of cell-wide signaling systems and often require additional interventions to allow glucose utilization and regain high growth rates. Sugar-specific strategies require extensive knowledge of orthogonal sugar utilization networks to engineer them and enable utilization of a defined sugar mixture.

1.4 Global Carbon Catabolite Repression Relaxation

The majority of CCR alleviation strategies reported in the literature are classified as global relaxation solutions. These leverage changes in global signaling mechanisms, such as modifications to the PTS or CRP, to allow for the uptake and simultaneous utilization of non-PTS

sugars in the presence of glucose. Often, it is the case that further engineering interventions are necessary to optimize the strain for a specific sugar mixture or to regain the ability to consume glucose satisfactorily due to deficiencies introduced by modifying global systems. A summary of the global strategies for CCR relaxation discussed is provided in Table 1.1.

1.4.1 PTS Modifications

The most common modification made to the PTS is to generate a strain with a Δ*ptsG* mutation which prevents the expression of the EIIB/C complex that facilitates the transport and phosphorylation of glucose. The Δ*ptsG* mutation requires glucose to enter through an alternative membrane transporter and causes $EIIA^{glc}$ to remain primarily in its phosphorylated state, both reducing inducer exclusion and enabling activation of catabolic genes. This strategy has been shown to increase titers of hydrogen, ribose, xylonate and many other products from feedstocks of mixed sugars due to increased secondary sugar uptake [28-30]. However, these studies showed both lower glucose utilization and lower growth rates in comparison to wildtype strains. Since the PTS is the primary glucose transport mechanism for *E.coli*, its mutation will lead to a slow growth and low glucose utilization phenotype [31]. Besides allowing sugar co-utilization, Δ*ptsG* mutants have also been postulated to increase phosphoenolpyruvate availability in the cell, which has been utilized to increase shikimate titer from a glycerol-glucose co-feed [32]. The Δ*ptsG* mutation approach has been shown to be successful in situations where the glucose content of the feedstock is low as well as when cellular growth rate is not of concern; however, these two qualifiers are often not true.

Due to the negative effects that PTS modifications have on strain performance, additional engineering designs are often required to improve strain utility. One such intervention is to induce adaptive evolution by applying a selective pressure and screening for growth to find strains with

beneficial mutations. Proteomic analysis showed that a PTS-deficient mutant subjected to glucose-enriched minimal medium preferentially increased concentrations of the transporters LamB, MglB and ManX to replace lost glucose transport activity [33]. These results provide a basis for the rational design of strains for the uptake of glucose in the presence of a PTS mutation. This information of novel preferred glucose transport alternatives is important as recent studies have shown that the overexpression of GalP, a commonly used replacement for glucose transport in engineered PTS mutant strains, is repressive to xylose utilization [34]. Alternatively, genetic switch mechanisms that either induce the expression of *ptsG* when the temperature is increased or degrade a subunit of the PTS using an activated protease-based inverter have enabled biphasic sugar consumption control and led to more efficient utilization of sugar mixtures [35, 36]. Computational approaches to predict consumption using whole genome models have also shown success in suggesting modifications that would increase co-utilization of sugars [37]. While often requiring large interventions, the re-introduction of wildtype-like glucose consumption through strategies like these makes a Δ*ptsG* mutation a more attractive avenue for CCR alleviation.

1.4.2 Co-Culture Applications

Another strategy employed in the literature to efficiently utilize mixed sugar feedstock is to co-culture strains with limited glucose consumption due to a PTS mutation with strains that can efficiently utilize glucose but cannot consume the secondary sugar due to CCR. This has been employed to consume a mixture of glucose, mannose, and galactose relatively successfully, with the co-culture utilizing galactose more slowly due the longer lag of the galactose-specific strain [38]. The construction of a consortium of xylose-preferring *E.coli* with yeast to utilize a mixture of glucose and xylose and generate ethanol has also been successful [39]. Innovative pathway engineering has led to ideas like the co-utilization of glucose and xylose for the efficient production

of butanol in a consortium of *E.coli* operating as a 'Y-shaped' unit [40]. In this design, two butanol producing strains, one glucose-preferring and the other xylose-preferring, were co-cultured. The production pathway within each strain differs until the formation of pyruvate, at which node the same pathway is utilized within each strain, forming the 'Y-shape'. Additionally, a strategy of separating the expression of the production pathway into different strains in a co-culture was used to increase both butanol and muconic acid titer from a glucose-xylose co-feed [41, 42]. This allowed for each strain to be engineered for either glucose or xylose preference and limited pathway down regulation caused by intermediate metabolites. Due to the relative simplicity of their design, the primary engineering hurdle for engineering effective co-culture systems is the maintenance of culture diversity during the fermentation, an area of research for any consortium design.

1.4.3 CRP Modifications

The modification of the CRP is another point of engineering that has been investigated for the ability to relax CCR globally. Multiple CRP* proteins were developed to have a decreased dependence on cAMP for the activation of genes by conformationally resembling cAMP-bound proteins [43]. These decrease the effects of CCR in *E.coli*; however, transcriptomic analysis later revealed that off target genes were differentially expressed in *crp** strains relative to wild-type in the absence of glucose, suggesting incomplete understanding of its global effects [44]. The coupling of a Δ*ptsG* mutation with expression of a *crp** enabled the co-utilization of glucose, xylose and arabinose from corncob hydrolysate to produce xylitol [45]. Additionally, the modulation of the ribosome binding site strength of the endogenous *crp* gene has been used to adjust the CRP content to better match the levels of cAMP present when other CCR relaxation strategies are undertaken, leading to increases in methyl ketone titer from a glucose-xylose co-feed

[46]. Moving forward, novel computational tools for estimating CRP binding affinities and new *crp* mutants discovered through deep sequencing of evolved mutants provide avenues for better understanding of the genome-wide interactions of CRP and could lead to further innovation in CRP engineering [47, 48].

1.4.4 Other Genetic Modifications

While mutation of the PTS is the most common genetic modification for the global alleviation of CCR, other genetic engineering strategies have shown success. Redirection of glycolytic flux through supplementary channels, such as the Entner Doudoroff pathway, while significantly decreasing growth rates, shows alleviation of CCR by circumventing the PTS [49-51]. Additionally, characterization of signaling proteins Mlc and SgrT that interact with the PTS could lead to further application for co-consumption purposes [52, 53]. Finally, strategies utilizing oligosaccharides containing glucose, such as cellobiose or sucrose, have shown promise in sugar co-utilization through the bypass of the PTS system by conversion of the oligomeric precursor to its free sugar constituents inside the cell [54, 55].

Table 1-1. Summary of global carbon catabolite repression relaxation strategies discussed.

Strategy	Engineering Intervention	Outcome	Reference(s)
PTS Modification	Δ*ptsG*	Increase in secondary sugar uptake coupled with lowered growth rates and glucose utilization	[28-30]
	Δ*ptsG* + temperature-inducible *ptsG* expression cassette	Arabinose, galactose, and xylose used for growth initially, glucose used for production after temperature increase	[35]
	protease-based inverter for degradation of PTSI	Increased titer of xylitol from a glucose-xylose feed when [Glucose]:[Xylose] > 1 compared to wildtype	[36]
Co-Culture	Co-culture of Δ*ptsG* strains with engineered sugar preferences	Enabled co-consumption of glucose, mannose and galactose with lag observed from galactose-preferring strain	[38]
	Co-culture of Δ*ptsG* xylose-preferring *E. coli* strain with *Saccharomyces cerevisiae*	Fermented all sugars in a 10% solids sugar cane bagasse slurry to ethanol in 30 hr	[39]
	'Y-shaped' microbial consortium	Enabled co-utilization of glucose and xylose for the production of butanol	[40]
	Separation of pathway components into different strains	Increased product titer by decreasing intermediate repression and splitting sugar usage	[41]
CRP Modification	*crp**	Facilitated activation of catabolic genes without the need for cAMP-CRP	[43]
PTS+CRP Modification	Δ*ptsG* + *crp**	Combination of two strategies; led to more efficient co-utilization of glucose and xylose than each individually	[45]
Other	Direction of flux through other catabolic channels using Δ*pfk*, Δ*zwf,* and/or Δ*pgi* mutations	Observed an increase in co-utilization of sugars by avoiding the PTS, but decreased growth	[49-51]
	Engineering for utilization of oligosaccharides containing glucose	Utilized cellobiose and sucrose to provide glucose without using the PTS	[54, 55]

1.5 Sugar-Specific Carbon Catabolite Repression Relaxation

Another, less investigated strategy to alleviate the effects of CCR in *E. coli* is to target sugar-specific parts of the signaling cascade. This engineering approach leads to mutants that co-consume pre-determined dual sugar mixtures containing glucose. The primary distinction between these solutions and global strategies are that cell-wide signaling systems, especially the PTS, are not modified and therefore the growth inhibition and glucose utilization problems that are normally observed with those modifications are avoided. A summary of the sugar-specific CCR relaxation strategies discussed is provided in Table 1.2.

1.5.1 Xylose-Glucose Co-Utilization

Due to the high xylose content in lignocellulosic biomass, significant research has been focused on its co-utilization with glucose for the production of value-added products. A complexity in the consumption of xylose is its regulation by the arabinose operon, creating a secondary repression system that leads to arabinose being preferentially consumed over xylose by *E.coli* [56]. The extensive knowledge of the regulation of the xylose operon has allowed for specific mutations to be made in signaling systems that are not known to affect the cell globally. A mutant deficient in *araC*, a gene encoding a transcription factor for arabinose and xylose catabolism genes, was able to co-utilize glucose and xylose after adaptive evolution that upregulated *xylA* as well as introducing an *araE*^{S91I} mutation [57]. Mutations introduced into transcription factors can allow for modification of their binding affinity and activity. A mutant XylR containing amino acid substitutions R121C and P363S, isolated through laboratory evolution, was found to increase the activation of the xylose operon and led to co-consumption of glucose and xylose without the need to overexpress any endogenous catabolic genes [58, 59]. Alternatively, the overexpression of *xylA, xylB,* and *araE* was found to enable the co-utilization of glucose and xylose to a limited degree for

the production of poly(3-hydroxybutyrate) [60]. To overcome inducer exclusion for xylose, successes in directed evolution of the promiscuous glucose transporter Glf towards increased specificity for xylose has also led to the co-utilization of glucose and xylose mixtures [61]. To predict interactions that could alleviate CCR for xylose, a computational method was developed using the condition that cells will die unless xylose is used as a carbon source in addition to glucose [62]. This study produced strains that grew well on mixtures of glucose and xylose but could not grow when glucose or xylose was fed individually. While this phenotype is not optimal, as single sugar feeds are still important to be accounted for in an industrial setting, the study showed the power of the simulation software to find co-utilization networks that do not involve PTS modifications.

1.5.2 Other Sugar Co-Utilizations

Sugar-specific interventions for other substrates are limited in the literature. A general solution for alleviating transcriptional regulation was shown through the expression of galactose metabolic genes under the control of a new promoter and synthetic UTR, which allowed for the co-utilization of glucose and galactose [63]. This strategy is powerful due to its ability to be engineered for application to multiple non-PTS sugars but is limited by its lack of confronting inducer exclusion, which could prove to be a larger problem for other systems. A more specific approach for acetate and glucose co-utilization has been shown through the introduction of a mutant in acetyl-CoA synthase that eliminates its sensitivity to acetylation, an effect caused by high glucose concentrations [64]. Due to the complexity of orthogonally engineering regulatory networks, there is a lack of literature on sugar-specific CCR alleviation strategies for non-xylose secondary sugars. However, there are many opportunities for the engineering and discovery of novel transport mechanisms that could lead to better sugar-specific CCR relaxation solutions [65].

Table 1-2. Summary of sugar-specific carbon catabolite repression relaxation strategies discussed.

Sugar Mixture	Engineering Intervention	Outcome	Reference(s)
Xylose + Glucose	Δ*araC*, overexpress *xylA*, *araE*^{S91I}, and constitutive expression of pentose metabolism genes	Allowed sugar co-consumption and utilization better than Δ*ptsG* mutants of the same strain, enabled production of xylitol	[57]
	xylR$^{R121C, P363S}$	Mutant has a higher DNA binding affinity that allows for CCR-independent gene activation	[58, 59]
	Overexpression of *xylA, xylB,* and *araE*	Mutant strain produced 2.09 times more poly(3-hydroxybutyrate) in comparison to wildtype due to increased co-utilization	[60]
	Engineered Glf protein expression	Engineered protein had a 10.8-fold improvement in xylose analog transport in the presence of glucose. No change was observed for co-utilization of the sugars due to transcriptional regulation	[61]
Galactose + Glucose	Replacement of endogenous UTR for galactose catabolism genes with a synthetic one	Enabled galactose and glucose co-consumption when fed low concentrations of both sugars	[63]
Acetate + Glucose	Overexpression of *acs*^{L641P}	Enabled 2.7-fold increase in acetate uptake in the presence of glucose in comparison to wildtype	[64]

1.6 Biosensors for High-Throughput Screening of Mutant Libraries

1.6.1 Biosensors for Metabolic Engineering

Through evolution, microbes possess the ability to sense various stimuli and react based on those conditions. Scientists have re-purposed these intrinsic regulatory systems to generate biosensors that can detect specific conditions and result in a desired phenotypic-conferring reporter, such as fluorescent protein expression [66]. The general classes of endogenous biosensors include two-component systems [67], riboswitches [68], and transcription factor-based systems [69].

Two-component systems combine sensing and regulatory mechanisms by utilizing both a sensory domain that binds an extracellular ligand and a kinase domain that elicits a cellular response through a signaling cascade [67]. These constructs have been successful in the detection and removal of copper [70] and zinc [71] through the engineering of the CusSR and ZraSR two-component systems, respectively, as well as enabling *E. coli* to 'see' light by grafting the EnvZ-OmpR system to the photoreceptor Cph1 [72].

Riboswitches bind to specific ligands and then control gene expression through mRNA degradation, and transcription or translation inhibition [68]. Many natural examples of riboswitches exist in biology including those for lysine [73], flavin mononucleotide [74], and thiamine pyrophosphate [75]. Synthetic constructs can also be engineered to enable a response to non-native ligands by modifying the aptamer region of the biosensor. Rational design of these regions can be attempted by using design rules developed using physics-based models [76, 77]. Additionally, non-rational approaches utilizing chemical genetics and genetic selection have also been shown to be effective [78]. These methods have led to the development of aptamers that respond to theophylline [79], ammeline [78], and azacytosine [78].

Transcription factor-based biosensors utilize regulatory proteins that possess both a ligand and DNA binding domain to modulate the transcription of a gene [69]. This structure lends itself to numerous engineering possibilities due to the vast number of transcription factors present in nature. Synthetic biologists have developed complex genetic circuits that act as logic gates by combining multiple orthogonal regulation systems [80]. Metabolic engineers often construct transcription factor-based biosensors by introducing an operator specific for a metabolite-responsive regulatory protein in front of a reporter gene and tuning expression to elicit a dose-response curve. This has been done for compounds such as butanol [81] and acyl-CoAs [82] to enable the identification of high-producing clones. Additionally, the modification of the ligand-binding domain of the transcription factor can elicit a response to a non-native ligand. The random mutation of 5 codons in the effector binding pocket of AraC, an arabinose-responsive transcription factor, enabled the isolation of a mevalonate-responsive system after high-throughput screening [83]. Rosetta, a ligand docking model, simplifies the development of transcription factor-based biosensors to respond to non-native ligands by optimizing the placement of the ligand in the binding pocket, leading to rational protein engineering [84].

The response triggered by the selected stimuli has also been engineered to increase the productivity of metabolic systems. The most common response is that of fluorescence or growth modification to be used for high-throughput screening, which is discussed in the following section. Using biosensors to introduce dynamic pathway regulation has also been shown to increase productivity. By using two biosensors to control both *pfkA* and *Miox* expression, a system was developed for the production of glucaric acid that transferred from a growth to a production phase and resulted in the highest titers to date [85]. Additionally, the combination of two quorum-sensing circuits led to increases in the titer of naringenin and salicylic acid by tuning the timing of the expression of

pathway constituents [86]. With their tunable nature and high utility, it is not a surprise that the amount of literature that is analyzing the application of biosensors to metabolic engineering has grown exponentially over time (Figure 1.3).

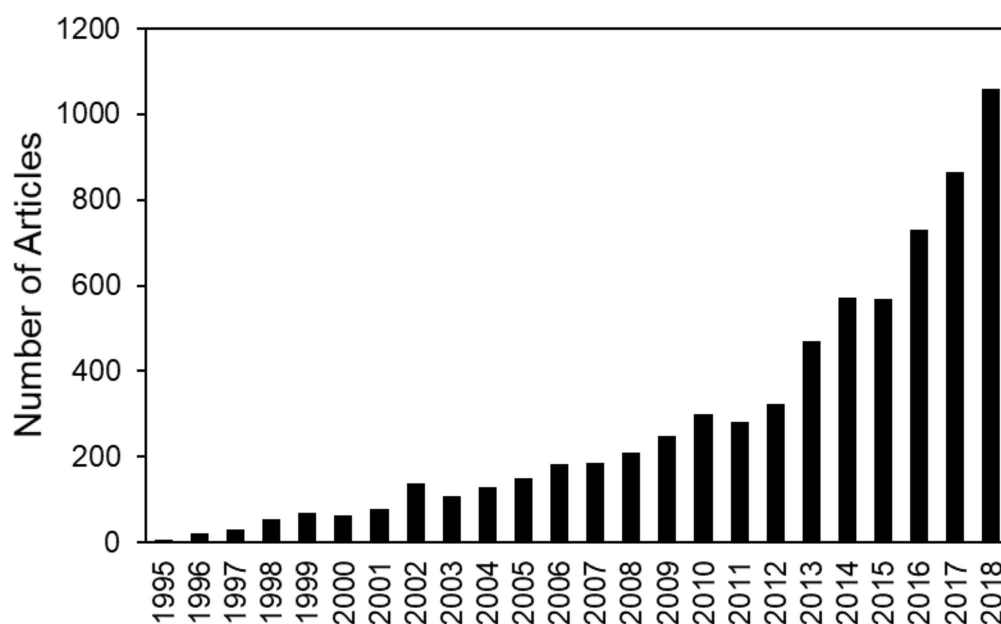

Figure 1-3. Number of publications containing keywords 'biosensor' and 'metabolic engineering' from a Google Scholar search. Figure was co-created with Stephanie Doong and is also included in https://hdl.handle.net/1721.1/123072.

1.6.2 Application to High-Throughput Screening

The recent development of the directed evolution framework has led to an increased interest in the investigation of high-throughput screening methods to increase its efficacy. Directed evolution works by mimicking nature, subjecting a diverse mutant population to a stimuli and selecting for top performers based on a designed screen [87]. The best individual from the first screen is then subjected to a procedure to re-introduce diversity and the process is repeated until a locally optimal individual is isolated. For metabolic engineers, this generally involves the introduction of random base pair mutations, with procedures like error-prone PCR, to a gene of interest and then screening

to determine if any mutant confers the desired activity characteristics [88]. Due to the large number of resources available for the production of diverse libraries, discussed in Section 1.1, the efficacy of a directed evolution experiment is often dictated by the screening strategy utilized. Figure 1.4 displays how the throughput of mutants tested varies immensely depending on the screening method.

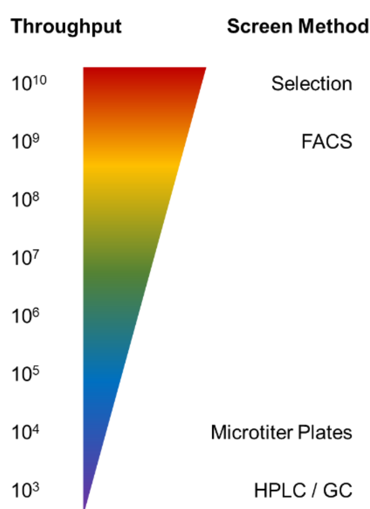

Figure 1-4. Throughput of mutants tested depending on screening method utilized. Adapted from Wang, Q., et. al. 2017 [89]. Figure was co-adapted with Stephanie Doong and is also included in https://hdl.handle.net/1721.1/123072.

While the accuracy of screening mutant performance using analytical methods, such as high-performance liquid chromatography (HPLC), is much greater than most high-throughput methods, the time and effort required to screen large populations diminishes its utility for this use. Colorimetric assays using microtiter plates increase the throughput of screening by tying the performance of a mutant to a color-based reporter that can be measured with a plate reader. This strategy has been implemented to develop screens for mutagenesis studies of terpene synthases [90] and to analyze L-tyrosine production of different bacterial strains [91]. Selection-based screening methods rely on increased survivability of mutants that possess the desired phenotype.

An example of this is the discovery of a strain of *Saccharomyces cerevisiae* with improved glucose/ethanol tolerance by culturing a library of mutated SPT15 genes in medium with high glucose content over successive dilutions [92]. Biosensors, with their tunable nature and readable fluorescent protein output, provide an opportunity for modular high-throughput screening protocols to be developed. Single-cell analysis of mutants is made possible by fluorescence-activated cell sorting (FACS) which quarantines cells that display high fluorescence into a collection vessel for further testing [93]. The top-performing mutants can then be analyzed with other low throughput methods like HPLC to confirm their activity. This method has been shown to be successful in the random mutagenesis of *Corynebacterium glutamicum* to increase L-serine yield [94]. Additionally, biosensor-based screening has proven useful for the directed evolution of heterologous proteins to increase their activity in *E.* coli. This includes the optimization of EctB from *Halomonas elongata* for the synthesis of ectoine [95], C2E from *Caldicellulosiruptor saccharolyticus* for the production of lactulose [96], and 2-PS from *Gerbera hybrida* for the synthesis of TAL [97].

1.7 Thesis Objectives

The aim of this work is to develop methods that allow *Escherichia coli* to efficiently utilize a secondary sugar for the synthesis of a value-added product when glucose is also present in the feedstock. We develop a pathway, expression platform, and purification procedure for the production and isolation of D-glyceric acid from galacturonate. We then analyze the efficacy of novel global and sugar-specific methods of CCR relaxation utilizing past studies into central PTS proteins and high-throughput screening of sugar transporters using a carbohydrate-specific biosensor. This thesis demonstrates the development of a novel pathway and expression platform for D-glyceric acid and highlights the importance of investigating CCR relaxation interventions

for the further development of bioprocessing. Additionally, this thesis incorporates a business case analysis for synthetic biology to disrupt the specialty chemicals industry demonstrating the possibilities for the future bioeconomy.

1.8 Thesis Organization

This thesis is organized into five chapters. The first chapter presents an introduction to metabolic engineering strategies for CCR relaxation in *E. coli* and a background on biosensor-based screening of mutant libraries for directed evolution. Chapter 2 presents a method to produce and purify D-glyceric acid from galacturonate and reports a novel pathway expression platform. Chapter 3 describes novel strategies that were developed for the engineering of *E. coli* for both the global and sugar-specific relaxation of CCR. Chapter 4 details an analysis of the opportunity for synthetic biology to disrupt the traditional specialty chemicals industry. Chapter 5 discusses the outlook of bioprocessing for the production of value-added chemicals and the importance of completely utilizing renewable feedstock streams to enable economic feasibility and improve sustainability.

2. Production of D-Glyceric Acid from D-Galacturonate in *Escherichia coli*

Portions of this chapter have been published in:

Fox, K.J., and Prather, K.L.J. Production of D-Glyceric Acid from D-Galacturonate in *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* 2020. **47**: p. 1075-1081.

Ni, C., Fox, K.J., and Prather, K.L.J. Substrate-activated expression of a biosynthetic pathway in *Escherichia coli*. *Biotechnol. J.* 2021. doi.org/10.1002/biot.202000433

Abstract

A microbial production platform has been developed in *Escherichia coli* to synthesize enantiopure D-glyceric acid from D-galacturonate. The expression of uronate dehydrogenase (*udh*) from *Pseudomonas syringae* and galactarolactone isomerase (*gli*) from *Agrobacterium fabrum*, along with the inactivation of *garK*, encoding for glycerate kinase, enables D-glyceric acid accumulation by utilizing the endogenous expression of galactarate dehydratase (*garD*), 5-keto-4-deoxy-D-glucarate aldolase (*garL*), and 2-hydroxy-3-oxopropionate reductase (*garR*). Optimization of carbon flux through the elimination of competing metabolic pathways led to the development of a *ΔgarKΔhyiΔglxKΔuxaC* mutant strain that produced 4.8 g/l of D-glyceric acid from D-galacturonate, with an 83% molar yield. Cultivation in a minimal medium produced similar yield and demonstrated that galactose or glycerol serve as possible carbon co-feeds for industrial production. To demonstrate industrial relevance, a basic purification procedure was developed that enabled a 150-fold increase in D-glyceric acid concentration in the final solution on a mass basis. Finally, a substrate-based induction platform was developed using a transcription factor-based signaling system that enabled expression of *udh* and *gli* upon the addition of D-galacturonate. The removal of the dependence on IPTG on the system decreases materials and purification costs at scale, making it attractive. This novel platform represents an alternative method for the production of D-glyceric acid that demonstrates benefits over reported acetic acid bacteria production methods and introduces an industrially-relevant substrate-based induction system that could be applied to many processes.

2.1 Introduction

Production of commodity chemicals is currently highly-dependent on fossil fuels, a non-renewable and environmentally detrimental feedstock [98]. Finding alternative, renewable starting materials to synthesize these products allows for a more secure production outlook. Renewable resources bring new challenges to chemical production, however, including logistical challenges due to their often de-centralized production as well as the increased complexity of the material itself [99]. Food waste, which is rich in diverse carbon sources and is at a surplus in many parts of the world, is seen as a promising renewable feedstock [100-102]. Food wastes contain a diverse set of sugars naturally structured into pectin, hemicellulose or cellulose that are valuable resources after chemical or enzymatic pretreatment [103, 104]. Bioprocessing, specifically using metabolically engineered microbes such as *Escherichia coli*, is an attractive method to create value-added products from these diverse sugar feeds [105, 106]. Using *E. coli* as a host organism is attractive due to the relatively high degree of understanding of many of its cellular processes as well as the many natural sugar catabolism and transporter proteins that are present in its proteome that enable easier utilization of these feeds. The fermentation of glucose, xylose, and other sugars present in food waste by engineered *E. coli* has enabled the production of many industrially relevant organic acids and alcohols such as succinic acid and 1,4-butanediol [107, 108].

Racemic glyceric acid is an industrially-relevant organic acid that has applications in medicine, polymer synthesis, and in surfactants as a base material [109]. Optically pure D-glyceric acid, however, has a large amount of untapped potential in medicine, due to its bioactivity, as well as in its material properties, as when included in solvents [110]. Currently, racemic glyceric acid and D-glyceric acid are predominately synthesized from glycerol, both biologically and chemically. For biological production, acetic acid bacteria *Gluconobacter fraterurii* and *Acetobacter tropicalis*

naturally produce glyceric acid at high titers from glycerol [111, 112]. These methods, however, utilize the glycerol substrate relatively poorly, resulting in yields around or below 50%. The optimization of culture conditions led to higher yields, near 60%, but these results still fall short of optimal and require an intensified culture strategy [113]. Additionally, a majority of acetic acid bacteria produce both enantiomers of glyceric acid, eliminating the opportunity that producing an enantio-pure product presents. The primary counterexample of this is an isolated strain of *Acetobacter tropicalis* that produces 99% enantiomeric excess (ee) D-glyceric acid [114]. However, this strain has been reported to suffer from stability problems that likely reduce its industrial applicability. Chemically, glyceric acid has also been produced from glycerol at high yields using various selective oxidation reactions [115, 116]. However, these processes also result in a racemic product. Attempts to perform cell-free enzymatic reactions have resulted in the production of enantio-pure D-glyceric acid from glycerol, but these processes are limited by low conversion [117].

The limitations in current D-glyceric acid production methods, including low yield, ee and stability, can be addressed through the metabolic engineering of alternative microbial hosts. We have developed a pathway in *E. coli* for the conversion of D-galacturonate to D-glyceric acid (Figure 2.1). D-Galacturonate is a sugar that is a main component of pectin and is of interest as a pathway substrate due to its high content in many agriculture waste streams [118]. The pathway is composed of two exogenous enzymes, uronate dehydrogenase from *Pseudomonas syringae* and galactarolactone isomerase from *Agrobacterium fabrum*, followed by three endogenous enzymes that are a part of the galactarate degradation pathway. Extensive knowledge of the sugar utilization networks in *E. coli* enables the optimized direction of carbon flux to product through genetic engineering, increasing molar yield. Additionally, the utilization of D-galacturonate as a starting

material dictates that only D-glyceric acid is generated based on previous activity measurements demonstrating that GarR was only active on the D enantiomer [119]. Finally, *E. coli* is known to be a stable microbe that is commonly used in the industrial production of many chemicals.

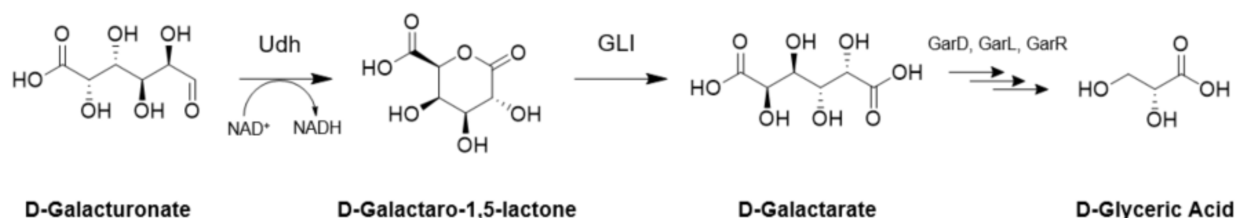

Figure 2-1. Designed pathway for the production of D-glyceric acid from D-galacturonate in *E. coli*. Enzymes uronate dehydrogenase (encoded by *udh*) and galactarolactone isomerase (encoded by *gli*) are exogenous, from *Pseudomonas syringae* and *Agrobacterium fabrum*, respectively. Enzymes galactarate dehydratase (encoded by *garD*), 5-keto-4-deoxy-D-glucarate aldolase (encoded by *garL*), and 2-hydroxy-3-oxopropionate reductase (encoded by *garR*) are endogenous to *E. coli*.

When designing an industrial process utilizing a synthetic pathway, an important factor to consider is the method of heterologous enzyme expression. There are many tuning parameters for expression, including the optimization of host strain, plasmid copy number, and promoter strength [120], however, the induction method specifically becomes important in industrial processes. Currently, those used include chemical induction, such as with IPTG [121], temperature induction [122], and quorum-sensing systems that lead to induction after a period of cell propagation [123]. Chemical induction methods are often useful at a small scale, but become untenable at industrial scale due to hypersensitivity to inducer concentration, high costs, and increased purification burden [124]. Additionally, commonly used induction methods are not tied to substrate presence which necessitates optimizations for new feeding strategies. We have developed a substrate-based induction system to enable expression of *udh* and *gli* in the presence of D-galacturonate. This

strategy allows cellular resources to be directed towards cell proliferation until the substrate is present, optimizing their utilization and improving upon current strategies.

2.2 Materials and Methods

2.2.1 E. coli Strain and Plasmid Construction

E.coli DH5α was used as a cloning strain for all plasmid manipulations. Cultures were propagated in Luria-Bertani (LB) broth (BD, Franklin Lakes, NJ) and all biological manipulations were done in accordance with standard practices [125]. Primers used in this study are included in Table 2.1. The Zyppy Plasmid Miniprep kit was used for all plasmid isolations (Zymo Research, Irvine, CA) and New England Biolabs restriction enzymes were used for all digestions (Ipswich, MA). The *gli* gene, encoding for galactarolactone isomerase from *Agrobacterium fabrum*, was purchased as a gblock from Integrated DNA Technologies (Coralville, IA) (Table 2.2). The pRSFDuet-1 vector with two IPTG-inducible T7 promoters preceding cloning sites was used as a backbone for the expression of these genes. A pRSFDuet vector with the *udh* gene from *Pseudomonas syringae* in the second cloning site was previously constructed in our group [50]. The kanamycin resistance gene was replaced with beta lactamase using circular polymerase extension cloning (CPEC) [126] to enable carbenicillin resistance. A Golden Gate cloning procedure [127] was then followed to insert the *gli* gene into the first cloning site using primers Gli_F and Gli_R (Table 2.1). The biosensor-based expression platform was cloned using a GFP-reporting galacturonate biosensor as the template. The NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs) was used to replace the *gfp* gene with *gli* and *udh* using primers hifi_backbone_F, hifi_backbone_R, hifi_udh_F, hifi_udh_R, hifi_gli_F, and hifi_gli_R (Table 2.1). Plasmids used in this study are presented in Table 2.3.

Mutations were made in host strain MG1655(DE3), carrying the DE3 lysogen encoding the T7 polymerase. Chromosomal inactivation of *garK*, *glxK*, *hyi*, and *uxaC* was completed using the procedure as described by Datsenko and Wanner [128]. Homology region lengths of at least 100

bp, employing two rounds of nested PCR (primers in Table 2.1), were used for the inactivation procedure (e.g., primer pairs dglxK_F_1 + dglxK_R_1 and dglxK_F_2 + dglxK_R_2 to inactivate gene *glxK*). Primers used to verify successful gene deletion are given in Table 2.1. Strains developed through this method and used in this study are presented in Table 2.4.

2.2.2 Culture and Analysis Conditions for D-glyceric acid Production

Cultures were grown in LB or MOPS minimal medium. All chemicals used for medium formulations and analytic standards were purchased from Sigma-Aldrich (St. Louis, Mo). The MOPS minimal medium was prepared from a 10x stock that contained 0.4 M MOPS, 0.04 M tricine, 0.1 mM $FeSO_4$, 95 mM NH_4Cl, 2.76 mM K_2SO_4, 0.005 mM $CaCl_2$, 3 mM $MgCl_2$, 500 mM NaCl, 0.03 µM $(NH_4)_6Mo_7O_{24}$, 4 µM H_3BO_3, 0.55 µM $CoCl_2$, 0.15 µM $CuSO_4$, 1.27 µM $MnCl_2$ and 0.17 µM $ZnSO_4$. This stock was diluted, supplemented with 1.3 mM K_2HPO_4, and titrated with NaOH to achieve a final pH of 7.2. For production experiments, the LB medium was supplemented with 10 g/l of D-galacturonate and the MOPS medium was supplemented with 5 g/l of D-galacturonate as well as 5 g/l of an additional carbon source as indicated in the Results section. LB cultures were inoculated at an OD_{600} of 0.05 and induced with 0.1 mM of IPTG at inoculation. Tubes containing 10 mL of culture were incubated at 30°C with agitation at 250 RPM. MOPS cultures were inoculated at an OD_{600} of 0.05. Tubes containing 10 mL of culture were incubated at 37°C with agitation at 250 RPM until exponential phase. Then, the incubation temperature was decreased to 30°C and the cultures were induced with 0.1 mM of IPTG. For D-galacturonate pathway expression activation fermentations, cultures were grown in LB medium at 37°C in 50 ml baffled flasks. 5 g/l of D-galacturonate was added when the cultures reached the exponential growth phase. Samples were taken at regular intervals for analysis by HPLC, UV-Vis, and qRT-PCR.

For analysis, samples were taken from the cultures at 24 hour intervals. OD measurements were taken, and the supernatant was analyzed by high-performance liquid chromatography (HPLC), using a 1200 Series Agilent Technologies instrument (Santa Clara, CA) with an Aminex HPX-87H Ion Exclusion Column (Bio-Rad Laboratories, Hercules, CA) and refractive index detector. The column temperature was maintained at 65°C and detector temperature was 35°C, with an isocratic mobile phase of 5 mM sulfuric acid run at 0.6 ml/min. This enabled baseline-level separation of elution peaks corresponding to each of the carbon sources tested as well as D-glyceric acid. A 22 minute method was used, with approximate elution times as follows: glucose (9.1 minutes), glycerol (13.6 minutes), galactose (9.75 minutes), arabinose (10.7 minutes), D-galacturonate (8.6 minutes), and D-glyceric acid (11 minutes).

2.2.3 Quantification of mRNA Levels for Expression Platform

Cultures of mutant strain MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi* transformed with pColE1_*exuR*-_*gli_udh*, pColE1_*lacI_gli_udh,* pColE1_*gli_udh*, or pColE1_*exuR* grown in LB medium were sampled before and after the addition of 5 g/l of D-galacturonate. Samples consisted of 10^6 cells that were taken from the fermentation, pelleted, aspirated, and then immediately stored at -20 °C. mRNA extraction was done using the illusta RNAspin Kit (Cytiva, Marlborough, MA). Reverse transcription was completed using the QuantiTect Reverse Transcription Kit (QIAGEN, Hilden, Germany). Quantitative PCR reactions used Brilliant II qPCR High ROX Master Mix (Agilent Technologies) with primers that anneal to *udh* (udh_qpcr_F and udh_qpcr_R in Table 2.1). An ABI 7300 Real Time PCR System (Applied Biosystems, Foster City, CA) was used to perform the reactions and the ABI software was used to determine Ct values. Fold differences were calculated for each sample in comparison to the uninduced biosensor-controlled expression system.

2.2.4 D-glyceric acid Purification

Spent fermentation medium from trials using MOPS minimal medium were used for purification tests. Acidification was done by adding hydrochloric acid dropwise and monitoring the pH of the solution after vigorous stirring. Concentration steps were completed using a hotplate for heating in a fume hood. For the extraction steps, 200 proof ethanol (Sigma-Aldrich) was added in excess to the concentrated solution and then stirred for multiple hours. The extraction step was repeated multiple times until the addition of ethanol did not lead to additional solids crashing out of solution. Final concentration was completed using a rotary evaporator to remove as much solvent as possible.

The final sample was analyzed using a TBDMS derivitization procedure followed by gas chromatography mass spectrometry (GCMS) analysis [129]. The purified sample was resuspended in 20 µL 2% methoxyamine-hydrogen chloride in pyridine (Thermo Scientific, Waltham, MA) and reacted for 90 min at 37 °C. 25 µL N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide with 1% tert-butyldimethylchlorosilane (Sigma-Aldrich) was then added and samples were incubated for 60 min at 56 °C. The supernatant was analyzed on an Agilent 6890N Network GC System coupled to an Agilent 5975B Inert XL MSD. A 3 µL injection volume was used in splitless mode with an inlet temperature of 270 °C. An Agilent J&W DB-35ms column was used with helium as the carrier gas with a flow rate of 1 mL/min. The temperature of the GC oven was initially set at 100 °C for 1 min, increased to 105 °C at 2.5 °C/min, held at 105 °C for 2 min, increased to 250 °C at 3.5 °C/min, and finally increased to 320 °C at 20 °C/min. The MS operated in electron ionization mode. Electron energy was 69.9 eV and the source and quadrupole temperatures were 230 °C and 150 °C, respectively.

Table 2-1. Primer sequences used in vector preparation and gene mutation procedures.

Primer Name	Sequence (5' to 3')
Gli_F	CGACATGGTCTCCTCATGAGCGAACTCGTCAGA
Gli_R	CGACATGGTCTCCTTCTAGGTCGCCTTGACCGG
dglxK_F_1	ACCTTATTCCGTGGGAACGCATCGAATAATTTGTTGAAAAAGGA TTGATAGTGTAGGCTGGAGCTGCTTC
dglxK_R_1	AGTACTGCACTCTTATGCGATATAAATCGCATCCGCTTTAAAGGT AAGGGTTCCGGGGATCCGTCGACC
dglxK_F_2	GCTCAGTAACGGCTTATTGGTCGGGATTTTACTGGCTGTTCTTAT GGATAACCTTATTCCGTGGGAACG
dglxK_R_2	ATTTGCGCGATTTTCAGGAAGGATATGGTACTCATGCCGCATCC GGCATGAGTACTGCACTCTTATGCGATA
dglxK_F_check	CCTTATTCCGTGGGAACG
dglxK_R_check	CATGAGTACTGCACTCTTATGC
dhyi_F_1	ACGCACCGACTGAAACCTGCTTCATGCACTATGAATAAGGGAGA TAAATATGTAGGCTGGAGCTGCTTCG
dhyi_R_1	CCCTGCCAACTTAGCGGCCTAAAAAAGCATTGTCTGAATAGCGT TACGTTTTCCGGGGATCCGTCGACC
dhyi_F_2	CGAACTGGATAACGTCATGGAATTTGAAGATATCGCCGATAACG CAGCGGACGCACCGACTGAAACCTG
dhyi_R_2	TAATTAACCTCTTTTAAATTTCGCTTTTCCTGAATTCAGACAACA CGATCCCCTGCCAACTTAGCGGC
dhyi_check_F	GAAGATATCGCCGATAACGCAG
dhyi_check_R	GGTGTACCCATAATGCCTAAGC
duxaC_F	TTGTAGGGACATTACCTGACGACAGC
duxaC_R	AATCGGCAGGCCATATTTGATGACATTG
dgarK_F	GGCCGAATAAGAGAAACGCCCTCCGGCGTCATGTGGTAACAGGC ATAGGTATGGTGTAGGCTGGAGC
dgarK_R	GCGCAGTATAGAGGGTTTGCGCGCCCTTGTCACCCCGCGTTGCG CATTCCTCCGTCGACCTGCAG
dgarK_check_F	CCCGGATGGCGTATTGCAATC
dgarK_check_R	ATGGAGCCCGGACTTTCCTC
udh_qpcr_F	TGTCCAAGCCCTTAGGAGGTAATACATGGCATCGGCTCATACCA CT
udh_qpcr_R	ACTGAGCCTTTCGTTTTATTTGATGCCTGGTTATTTATCGCCGAA CGGTCCG
hifi_backbone_F	CCAGGCATCAAATAAAACGAAAGGC
hifi_backbone_R	TTGTCCTTAAGCGTGAACGAAAGTTAAACAAAATTATTTGTTCA AAATGTTAACGTTAAC

hifi_udh_F	TGTCCAAGCCCTTAGGAGGTAATACATGGCATCGGCTCATACCA CT
hifi_udh_R	ACTGAGCCTTTCGTTTTATTTGATGCCTGGTTATTTATCGCCGAA CGGTCCG
hifi_gli_F	ACTTTCGTTCACGCTTAAGGACAATTTATAATGAGCGAACTCGTC AGAAAACTG
hifi_gli_R	TATTACCTCCTAAGGGCTTGGACACTAGGTCGCCTTGACCGG

Table 2-2. Sequence of gblock encoding for *gli* from Agrobacterium fabrum that was used in the creation of the pRSFD_*udh_gli* vector.

Sequence Name	Sequence (5' to 3')
GLI	atgagcgaactcgtcagaaaactgagcggcacagcgccgaaccccgcctttccgagaggcgcggtggatacgcaga tgcacatgtatctgcccggctatcctgccttgccgggtggtcctggcctgccgccgggcgctctgccggggccggagg attatcgtcgcctcatgcaatggctcggcatagaccgggtcatcatcacgcagggcaacgcccatcagcgcgataatg gcaacacactggcttgcgttgccgaaatgggcgaggcggcccacgccgttgtcatcatcgatgcgaccacgaccgaa aaggacatggaaaagctcaccgccgccggtacggtcggcgcgcgtatcatggatttgccgggcggcgcggtgaacc tgtccgaattggacgcggtggacgagcgggcacatgcggccgactggatggtggcagtgcagttcgacggcaacgg tcttctcgatcatctgccgcgccttcagaaaatccgctcccgctgggtgtttgatcatcacggcaagtttttcaagggcatc aggacggatgggccggaaatggcggcccttctgaagctcatcgaccgaggcaatctctggttcaaattcgctggcgttt atgaaagctcccgcaaaagctggccctatgccgatgtcgccgccttttcgcgggtgatcgccgcccatgcgccggagc gcatcgtctggggcaccaactggccgcataattcggtgcgcgagacggcggcctatcccgacgatgcccgtcttgcg gaactgacgctcggctggctgccggatgaggcggcgcgtcatcgggcgctggtcgaaaacccggaagcgctgttca agctgtcgccggtcaaggcgacctag

Table 2-3. Plasmids used in this study.

Plasmid	Genotype	Reference
pRSF_*udh_gli*	RSF origin, IPTG-inducible *udh* from *Pseudomonas syringae* and *gli* from *Agrobacterium fabrum*	this study
pColE1_*exuR_gli_udh*	ColE1 origin, galacturonate biosensor that enables a dose-response of *gli* and *udh* expression with galacturonate addition	this study
pColE1_*gli_udh*	ColE1 origin, constitutive expression version of pColE1_*exuR_gli_udh* with *exuR* gene and ExuR operator sites removed	this study
pColE1_*exuR*	Empty vector control version of pColE1_*exuR_gli_udh* with *gli* and *udh* genes removed	this study
pColE1_*lacI_gli_udh*	ColE1 origin, IPTG-induced expression of *gli* and *udh*	this study

Table 2-4. Strains used in this study.

Strain Genotype	Reference
MG1655(DE3)Δ*garK*	this study
MG1655(DE3)Δ*garK*Δ*uxaC*	this study
MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*	this study
MG1655(DE3)Δ*garK*Δ*uxaC*Δ*hyi*	this study
MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*	this study

2.3 Results and Discussion

2.3.1 D-glyceric acid Production using Novel Pathway and Strain Optimization for Titer and Molar Yield from D-galacturonate

Our pathway for the synthesis of D-glyceric acid from D-galacturonate in *E. coli* uses two heterologous proteins, uronate dehydrogenase (Udh), encoded by *udh*, from *Pseudomonas syringae* and galactarolactone isomerase (GLI), encoded by *gli*, from *Agrobacterium fabrum*, to convert D-galacturonate to D-galactarate. The *udh* gene has been utilized previously in our group for the production of glucaric acid and therefore was known to be functionally expressed in *E.coli* [130]. Endogenous enzymes galactarate dehydratase (encoded by *garD)*, 5-keto-4-deoxy-D-glucarate aldolase (encoded by *garL*), and 2-hydroxy-3-oxopropionate reductase (encoded by *garR*) then convert D-galactarate to D-glyceric acid (Figure 2.2). We expressed *udh* and *gli* from a high-copy-number pRSFDuet vector. Induced expression of these genes alone does not allow for significant accumulation of D-glyceric acid in the culture (Figure 2.3). Additionally, the deletion of the most active glycerate kinase, glycerate 2-kinase (encoded by *garK*), does not enable detectable amounts of D-glyceric acid to accumulate without expression of the exogenous genes. However, a Δ*garK* strain with induced expression of *udh* and *gli* was able to synthesize of 3.5 g/l product with a molar yield of 60% after 48 hours.

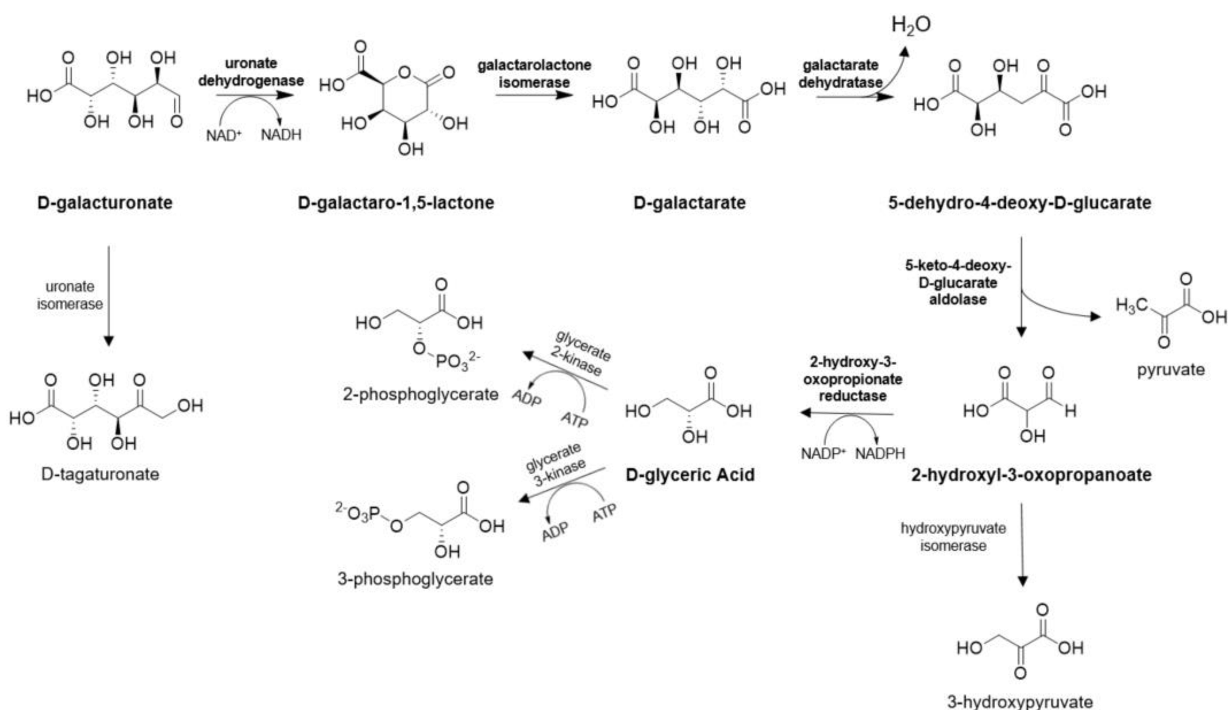

Figure 2-2. Full D-glyceric acid pathway from galacturonate with heterologous enzymes uronate dehydrogenase (encoded by *udh*) and galactarolactone isomerase (encoded by *gli*) and endogenous enzymes galactarate dehydratase (encoded by *garD*), 5-keto-4-deoxy-D-glucarate aldolase (encoded by *garL*), and 2-hydroxy-3-oxopropionate reductase (encoded by *garR*). Also shown are endogenous enzymes uronate isomerase (encoded by *uxaC*), hydroxypyruvate isomerase (encoded by *hyi*), glycerate 2-kinase (encoded by *garK*), and glycerate 3-kinase (encoded by *glxK*) which represent carbon flux diversions from the pathway. Bolded enzyme names and compounds represent the intended direction of flux to increase D-glyceric acid titer and yield.

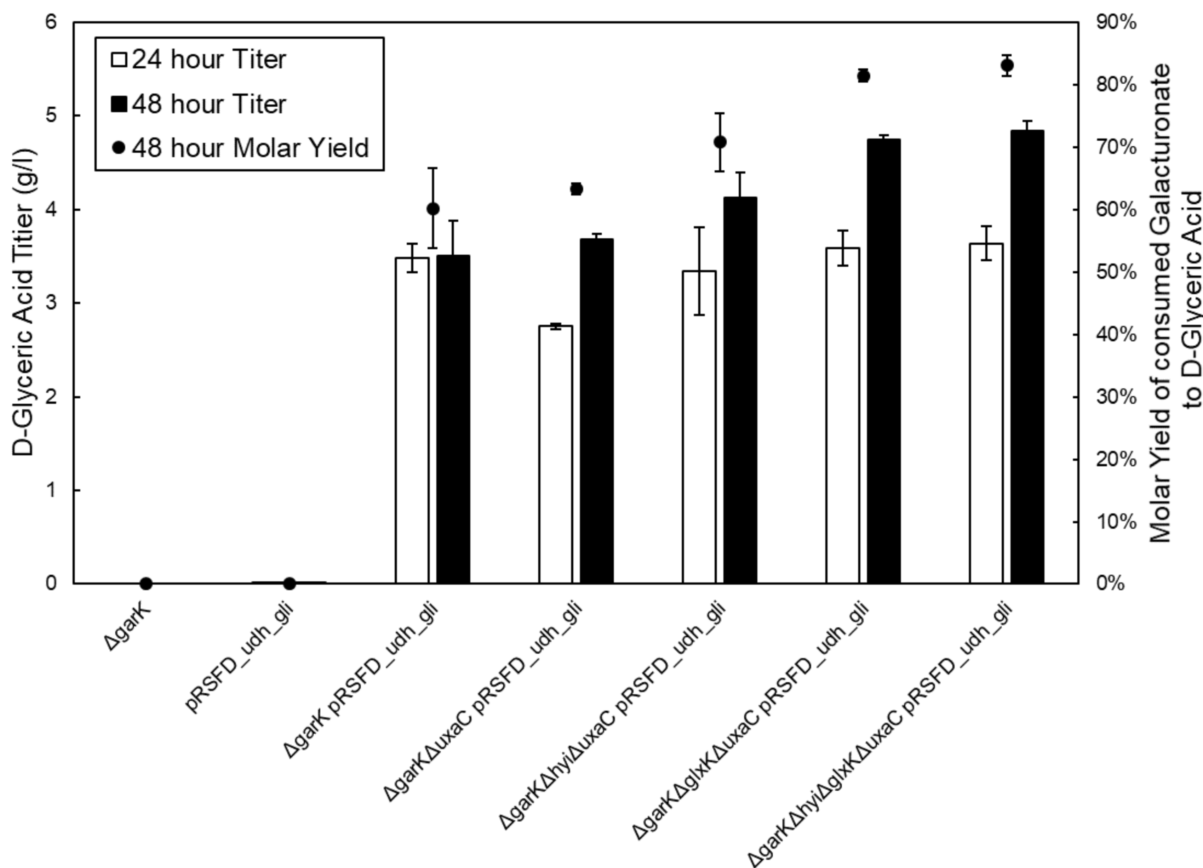

Figure 2-3. Production of D-glyceric acid by different MG1655(DE3) mutant strains. Cultures were grown in LB medium supplemented with 10 g/l of D-galacturonate and 0.1 mM IPTG at 30°C with a shaking rate of 250 RPM. Error bars represent standard deviation of biological triplicates. Molar yield is defined as mol D-glyceric acid produced per mol D-galacturonate consumed in the culture.

To increase the carbon flux from D-galacturonate to D-glyceric acid, we engineered the strain to minimize loss to unwanted byproducts by eliminating competing pathways [107]. Uronate isomerase, encoded by *uxaC*, converts D-galacturonate to tagaturonate and is the first committed step of hexauronate catabolism in *E. coli* (Figure 2.2). It was not anticipated that the deletion of this gene would be necessary to accumulate D-glyceric acid due to the ability of microbes to utilize a pool of substrate to supply multiple fluxes. However, it was surprising that the deletion of *uxaC* led to only negligible increases in both titer and molar yield since it is a major pathway in the

utilization of the feedstock for biomass. In fact, the titer at 24 hours in the Δ*uxaC* strain is actually lower than that produced by the Δ*garK* strain, suggesting that there could be competition for cellular resources at early time points. The similar titer and yield produced by this mutant strain suggests that the recombinant enzyme flux is large enough to outcompete the catabolic pathway for the D-galacturonate substrate but competing reactions in the downstream (endogenous) portion of the pathway direct carbon flux away from the target product (Figure 2.2). The *uxaC* knockout was maintained in strains moving forward, however, since it was anticipated that as the medium complexity changed, there could be larger relative loss from the upstream portion of the pathway. Hydroxypyruvate isomerase, encoded by *hyi*, competes with product formation by converting 2-hydroxyl-3-oxopropaoate to 3-hydroxypyruvate [131]. Deletion of *hyi* increased yield in comparison to the Δ*garK*Δ*uxaC* strain (Figure 2.3, bars 4 and 5). Another glyceric acid kinase, *glxK*, is present in the genome of *E. coli* in addition to *garK.* These two proteins differ based on their thermolability, pH dependence and substrate binding affinity [132]. The mutation of both kinases increased the production of D-glyceric acid, with the more active GarK having a larger effect. The best productivity was observed by the Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC* and Δ*garK*Δ*glxK*Δ*uxaC* mutant strains, with both showing statistically similar titer and molar yield. This suggests that while adding the Δ*hyi* mutation to the Δ*garK*Δ*uxaC* strain increased both titer and yield, its importance is highly diminished when a Δ*glxK* mutation is also present. The MG1655(DE3)Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC* mutant on average produced the most product, 4.8 g/l of D-glyceric acid, with a molar yield of 83%.

The molar yields of D-glyceric acid obtained in this production system are higher than those currently reported from the fermentation of glycerol by acetic acid bacteria [109]. Additionally,

using this production platform allows for only D-glyceric acid to be produced, resulting in complete enantio-purity.

2.3.2 Carbon Source Co-feed Analysis

Production in minimal medium is preferred for commercial purposes as it lessens purification burden and reduces fermentation inconsistencies caused by rich medium. To transition the production into a minimal medium, it was anticipated that the feeding of an additional carbon source, other than D-galacturonate, would be necessary for robust growth due to the modification of the hexuronate utilization pathway in the optimal production strain (Δ*uxaC*). Secondary sugars commonly found in pectin, such as galactose and arabinose, were tested since they would likely be found in natural mixtures with D-galacturonate [118]. In addition, common co-feeds such as glucose and glycerol were tested. The MG1655(DE3)Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC* strain that gave the optimal results in LB medium was used in this experiment. The combination of the high-copy-number vector with induction at inoculation produced a long lag period with delayed production if the same procedure used for the LB medium was used for the MOPS minimal medium. Therefore, we chose to first culture the cells at 37°C until exponential growth was observed, then induce the expression of *udh* and *gli* and reduce the incubation temperature to 30°C.

All combinations of secondary sugars produced some amount of D-glyceric acid, with varying success, and all maintained approximately the same molar yield that was observed in the strain optimization study (Figure 2.4). A pure feed of D-galacturonate was able to produce product since the reaction catalyzed by GarL produces a mol of pyruvate for each mol of D-galacturonate consumed. However, this flux appeared to be too little to enable robust growth and an extended lag period in comparison to the other conditions was observed (Figure 2.5). Galactose and arabinose, both components of pectin along with D-galacturonate, showed differing success when

used as a secondary sugar. Feeding galactose led to high titers of D-glyceric acid while adding arabinose was observed to be detrimental to production. This could be due to the different utilization networks for each of these substrates with galactose entering glycolysis from the Leloir Pathway and arabinose using the non-oxidative section of the pentose phosphate pathway (PPP). By entering in the non-oxidative part of the PPP, arabinose skips crucial NADPH production which is a necessary co-factor for GarR and could explain the lower D-glyceric acid titer with this co-feed. Unsurprisingly, the glucose co-feed produced the poorest results when used as the co-substrate with D-galacturonate. This is likely due to both inducer exclusion and transcriptional regulation of the galacturonate transporter gene *exuT*, effects brought on by carbon catabolite repression that prevent the usage of some sugars when co-fed with glucose [21]. A glycerol co-feed produced yields that were similar to those seen in LB medium. However, secondary carbon source utilization was lowest with glycerol (41% usage) compared to complete utilization in all other co-feeds tested (Table 2.5). In all feeding strategies, the molar yield remained high (often higher than in the LB medium case) showing that the co-feeding strategy did not activate other pathways that would reduce efficiency. This system therefore enables a majority of the D-galacturonate fed to be directed toward product synthesis and all of the additional carbon source to be utilized for growth, similar to the parallel metabolic pathway engineering approach recently described [133]. Further analysis into the effect of arabinose on the transport of D-galacturonate and the usage of it in the designed pathway is necessary to enable this feeding strategy. Additionally, application of carbon catabolite repression relaxation methods would be beneficial for more complete usage of the primary contents of pectin.

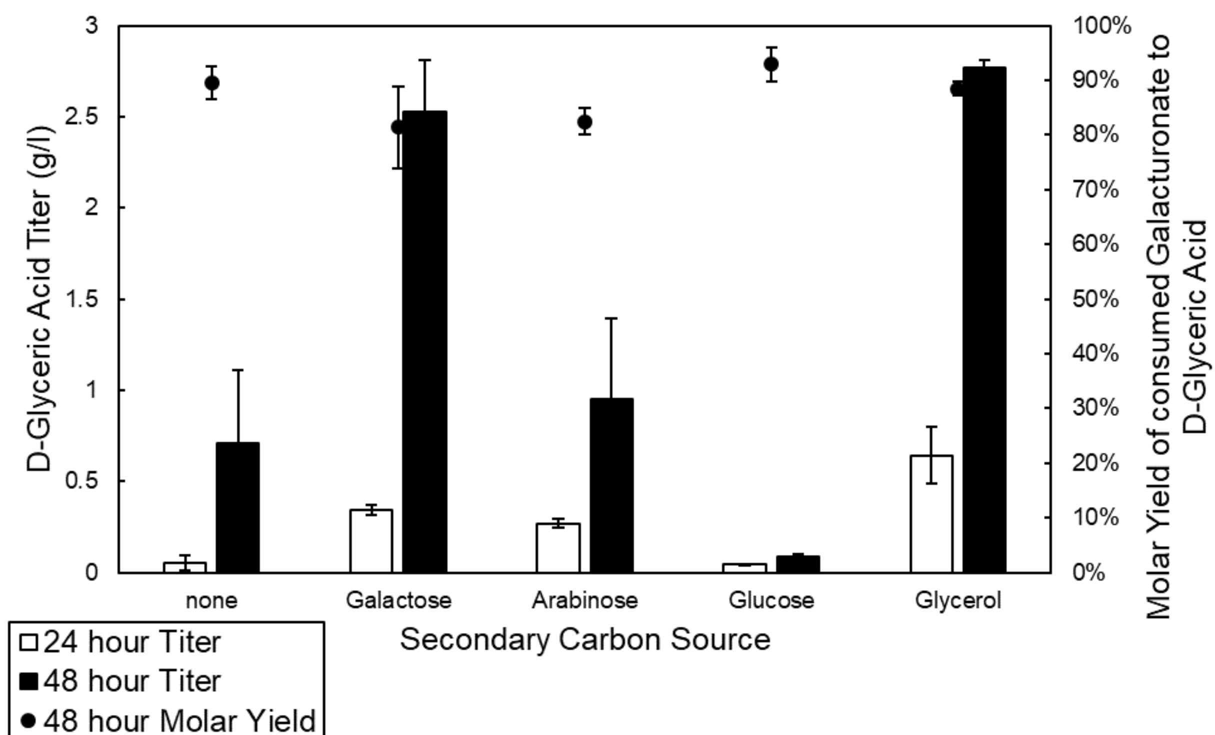

Figure 2-4. Production of D-glyceric acid from MG1655(DE3)Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC* (pRSFD_*udh_gli*) with various secondary sugar co-feeds. Cultures were grown in MOPS minimal medium supplemented with 5 g/l of D-galacturonate and 5 g/l of the listed secondary sugar. Cultures were grown at 37°C until exponential phase and then shifted to 30°C and induced with 0.1 mM IPTG. Error bars represent standard deviation of biological triplicates. Molar yield is defined as mol D-glyceric acid produced per mol D-galacturonate consumed in the culture.

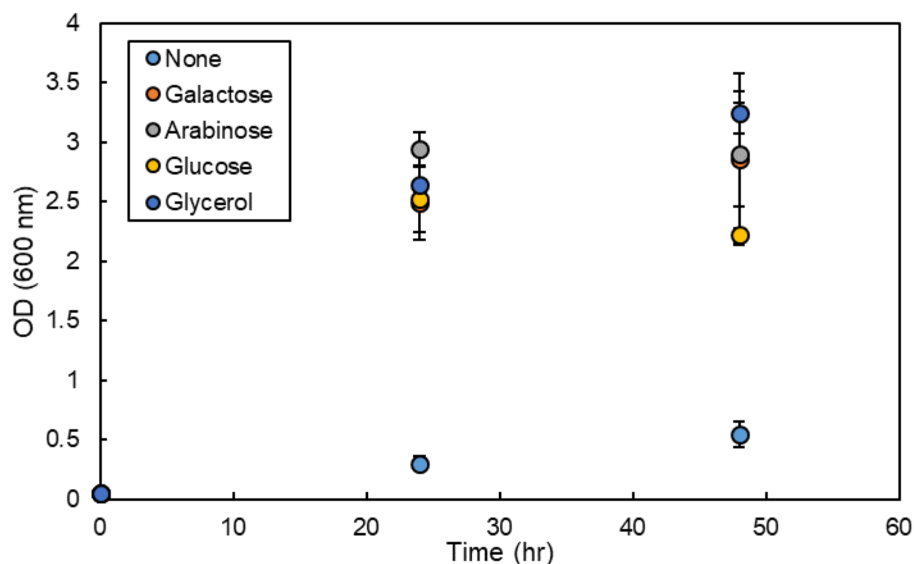

Figure 2-5. Optical density (at 600 nm) measurements for the production of D-glyceric acid in minimal medium from 5 g/l of galacturonate co-fed with 5 g/l of each specified secondary carbon source.

Table 2-5. Fraction of galacturonate and secondary carbon source consumed, titer and molar yield of D-glyceric acid in minimal medium experiment. Results correspond with the data presented in Figure 2.4.

Secondary Carbon Source	Fraction Galacturonate Consumed	Fraction Secondary Carbon Source Consumed	D-glyceric acid titer (g/l)	Molar yield of D-glyceric acid from galacturonate	Final OD_{600}
None	0.25	N/A	0.71	89%	0.55
Galactose	0.99	1.00	2.52	81%	2.86
Arabinose	0.41	1.00	0.95	82%	2.89
Glucose	0.03	1.00	0.09	93%	2.22
Glycerol	1.00	0.41	2.77	88%	3.25

2.3.3 Feedstock-based Expression of D-Glyceric Acid Production Pathway

The IPTG-based expression platform used for the development of the D-glyceric acid pathway is optimal for academic experiments due to its widespread use. However, as processes are scaled up, the necessity of IPTG becomes a hindrance due to added cost and the additional purification

needed. To eliminate this, we developed a novel platform for the expression of *udh* and *gli*, activated by the presence of galacturonate in the medium (Figure 2.6). For this system, a galacturonate-specific transcription factor gene, *exuR* from *Bacillus subtilis*, is constitutively expressed and two operators associated with this transcription factor are placed within the promoter region for the two pathway genes. In the absence of galacturonate, ExuR binds to the operators and prevents the transcription of *udh* and *gli*. When galacturonate is added to the medium, it interacts with ExuR and causes a conformational change that releases this inhibition. Therefore, this expression platform is activated only when the proteins are necessary for product formation, reducing unnecessary energy usage.

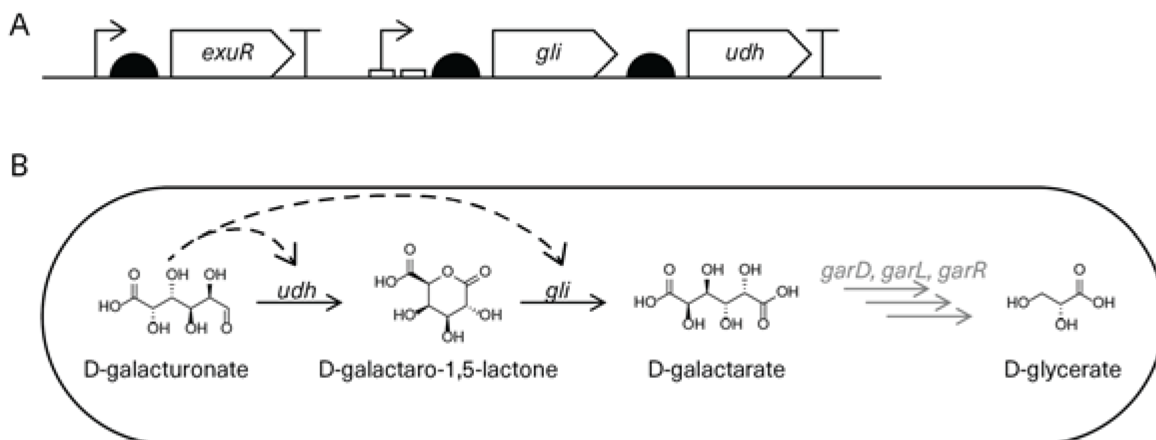

Figure 2-6. Expression platform for the activation of the D-glyceric acid pathway by D-galacturonate. The (A) expression module consists of a constitutively expressed *exuR* gene and an expression cassette that consists of *gli* and *udh* that has its expression regulated by ExuR operators present in the promoter region of the cassette. This construct enables a (B) system where the expression of the pathway to produce D-glycerate is dependent on the presence of D-galacturonate.

A MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi* strain was transformed with the plasmid encoding the expression platform to test the system for its utility in D-glyceric acid production (denoted as P1OO). A constitutive expression version of the expression platform, with no *exuR* gene or

operators (denoted as P1XX), an empty vector version of the vector without *gli* or *gli* and an IPTG-inducible version of the expression platform, with *LacI* replacing *exuR* and a LacO site in place of the ExuR operators, were used as controls. D-galacturonate was added once the culture reached exponential growth. A culture carrying the biosensor was also not fed galacturonate during the experiment to be used as an additional negative control. All conditions grew similarly in LB medium with the exception of the constitutive-expression condition (Figure 2.7(B)). This growth inhibition is likely due to metabolic burden caused by the high-level expression during the early time points. Additionally, it was observed that during the cloning procedure strains carrying the constitutive-expression vector grew poorly which suggests that the culture was likely in a disadvantageous metabolic state even at inoculation.

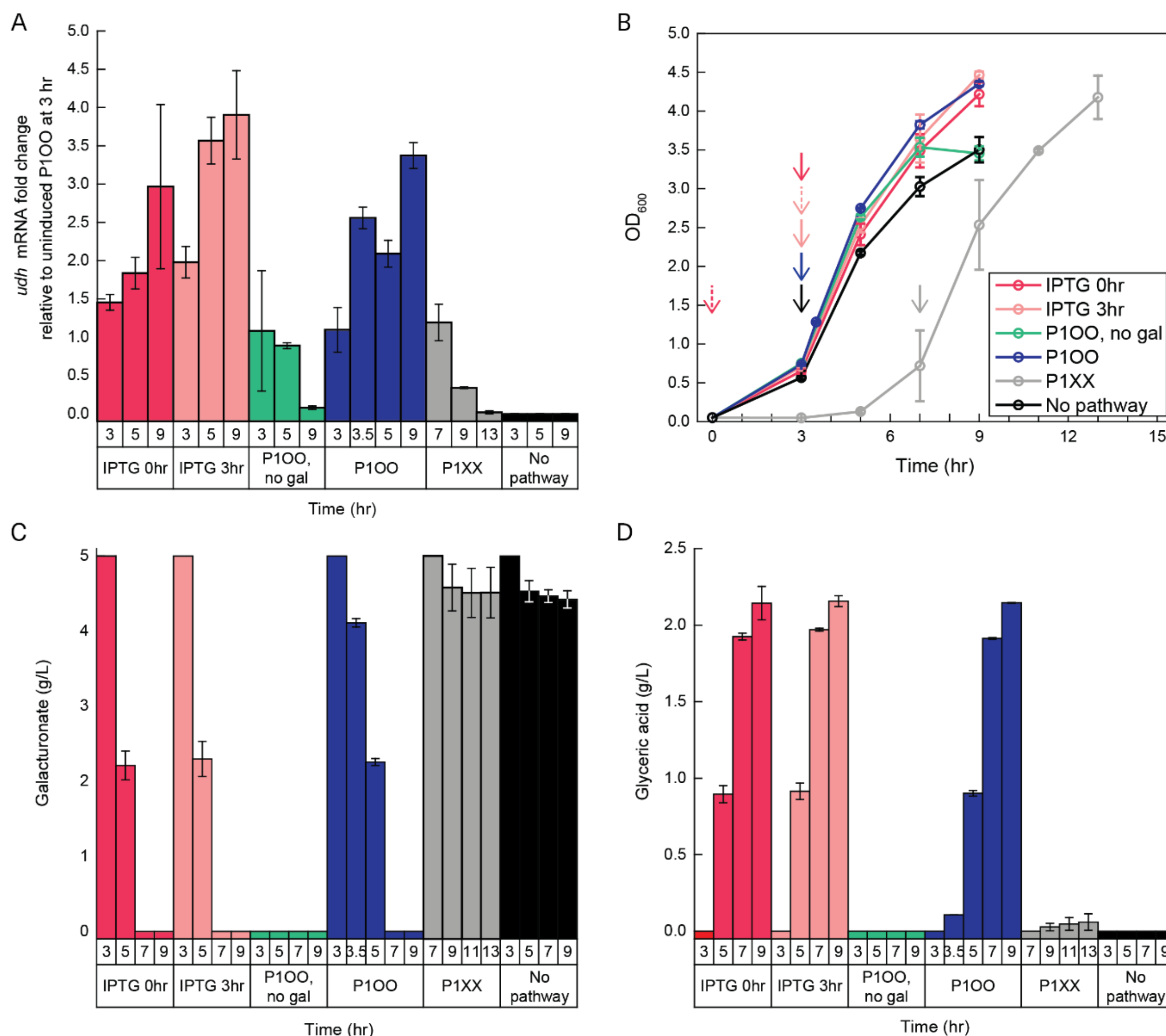

Figure 2-7. Fermentation data from the production of D-glyceric acid from D-galacturonate using different platforms for the expression of *gli* and *udh*. (A) Fold-change of *udh* mRNA compared to uninduced P1OO condition at 3hr for each expression platform over time. (B) OD_{600} of cultures over time, dotted arrows indicate IPTG addition and solid arrows indicate galacturonate addition. (C) Galacturonate and (D) D-glyceric acid concentration over time. Cultures were grown in LB medium at 37°C with a shaking rate of 250 RPM supplemented with 5 g/l of D-galacturonate at 3 hr for empty vector, IPTG-activated: hr=0, IPTG-activated: hr=3, and Biosensor conditions and at 7 hr for constitutive condition. Error bars represent standard deviation of biological duplicates.

Due to the delay in growth, D-galacturonate was not added to the constitutive condition until 7 hours while all other conditions had it added at the 3-hour time point. The D-galacturonate

utilization and D-glyceric acid production rates for the IPTG-based and biosensor expression platforms were similar (Figure 2.7(C)). This suggests that the two methods of induction could be used interchangeably to achieve the same productivity result. Additionally, adding IPTG at either the 0 or 3 hour time points did not affect the results of the fermentation for the IPTG-inducible system (Figure 2.7(D)). However, the constitutive-expression system both did not utilize much D-galacturonate and did not produce appreciable amounts of D-glyceric acid. It is likely again that the metabolic burden that was seen in the OD measurements is causing this difference, preventing adequate resource allocation to produce D-glyceric acid. This demonstrates that having an inducible system that controls the expression of genes is advantageous for this production platform. As expected, the empty vector version of the biosensor both did not utilize D-galacturonate and did not produce D-glyceric acid.

Analysis to determine the strength of induction for each of the expression platform conditions was done using qRT-PCR (Figure 2.7(A)). These results show the activation of expression of *udh* upon the addition of D-galacturonate for the biosensor-based expression platform which is similar to that of the IPTG-induced system. When comparing the biosensor-based expression system with and without D-galacturonate added, there is a clear difference in mRNA content during the later time points demonstrating that it is the D-galacturonate that is activating expression of *udh*. The no pathway condition shows negligible *udh* mRNA which is expected since it is an exogenous gene. The *udh* mRNA content for the constitutive-expression system shows a decreasing behavior, which is expected with that type of expression. However, it is surprising that the level of transcript is lower than the other conditions which have their transcription under regulation. Based on previous experiments with GFP expression (not shown) we expected relatively high levels of expression. It is suspected that this difference is due to the enhanced cell stress of producing *udh*

in an unregulated manner. The low transcript levels are supported by the strain's inability to produce D-glyceric acid and slow growth profile, also indicating cell burden. This underscores the importance of designing systems that do not rely on constitutive-expression platforms, which may be limiting the productivity of the system. These results show that in this system an IPTG-based induction system can be replaced by a substrate-based one. This both did not limit the performance of the strain and eliminated the need to feed an additional chemical (IPTG) into the system, simplifying purification and lowering cost.

2.3.4 Purification of D-Glyceric Acid from Spent Fermentation Medium

The purification of D-glyceric acid from spent fermentation medium is integral for further testing of the enantiopure product for polymer applications. We developed a procedure designed to increase the purity to levels that would allow for chemical manipulation (Figure 2.10). The optimal purification procedure was performed on a defined MOPS-based spent medium which greatly decreased the number of constituents to be separated in comparison to a rich LB medium. First, cell debris was removed by centrifugation and the resulting clear sample was acidified with hydrochloric acid until reaching a pH of approximately 2. This was necessary due to the large amount of salt present in the medium that was likely complexing with the deprotonated D-glyceric acid (pKa=3.42). Upon acidification, a large amount of salt crashed out of the solution. Water was removed through evaporation by heating with solids removed from the suspension by centrifugation, as necessary. Ethanol was chosen as an extraction solvent due to the low solubility of medium constituents and starting materials in it. Solubility information is not available for D-glyceric acid, but it was hypothesized that it would dissolve in ethanol due to the high solubility of glycerol and glycolic acid, which both have similar structures to D-glyceric acid, in this solvent. The concentrated water-based solution was added to a large excess of pure ethanol and the

compounds that crashed out of solution were removed. This solution was evaporated, and pure ethanol was again added in excess. The extraction procedure was repeated until no solids crashed out upon pure ethanol addition. Multiple repetitions of extraction was likely necessary due to the solubility of many sugars in solutions that contain small percentages of water. Therefore, as the solution became closer to pure ethanol, diminishing amounts of sugars and other salts crashed out of solution. Finally, the sample was concentrated using a rotary evaporator to remove the remaining ethanol. The final sample was analyzed by HPLC and contained 1.3 grams of D-glyceric acid, a 65% recovery from the starting concentration. The purity of the final sample was 59.1% by mass, approximately a 150-fold increase of concentration in comparison to the starting spent fermentation medium. The original concentration of D-glyceric acid in the spent fermentation medium was 4 g/l with a total mass of 2 grams.

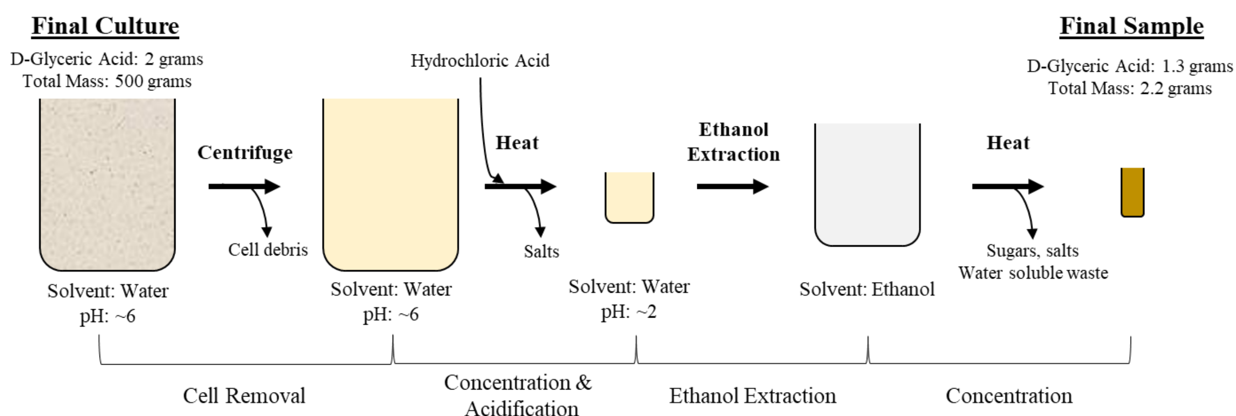

Figure 2-8. Extraction procedure for the enrichment of D-glyceric acid from spent fermentation medium. Cell medium is removed of cell debris by centrifugation, concentrated by heating and then acidified to remove salts. An ethanol extraction is performed to separate non-soluble starting materials and by-products from the desired D-glyceric acid. A final concentration step results in a 150-fold increase in of the concentration of D-glyceric acid in solution compared to the starting spent medium.

Additional confirmation of the identity of D-glyceric acid in the final sample was completed by derivatizing it with tert-Butyldimethylsilyl ether and analyzing the sample with GCMS. The mass

spectra obtained (Figure 2.11(A)) matched both past archived data as well as a control sample of D-glyceric acid that was analyzed with the same method. After purification, the sample was a dark straw color and viscous (Figure 2.11(B)). In comparison, pure (purchased) D-glyceric acid is a colorless syrupy liquid.

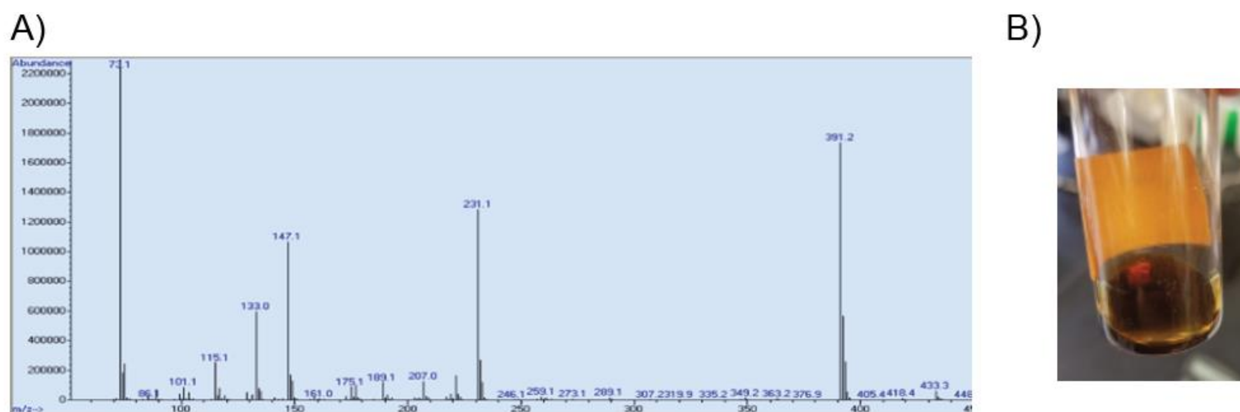

Figure 2-9. Analysis of D-glyceric acid solution after purification procedure. The (A) mass spectrum was obtained using a GCMS after derivitization. This spectrum matches both archived mass spectra for this compound [134] as well as a control sample that was also analyzed (data not shown). Visually (B) the final sample was a dark straw color and viscous.

2.4 Conclusions

The development and optimization of novel utilizations of renewable resources in bioprocessing is imperative for their relevance in the future chemicals industry. We developed a pathway expressible in *E. coli* for the production of D-glyceric acid, a value-added chemical with potential use in the materials industry, from galacturonate, commonly found in pectin. This pathway combines a heterologous upstream module consisting of enzymes Udh and GLI with a native downstream module consisting of GarD, GarL, and GarR. Through strain engineering we developed a MG1655(DE3)Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC*(pRSFD_*udh_gli*) production system that enabled 4.8 g/l of D-glyceric acid titer with a molar yield of 83%. While this molar yield is high, half of the substrate mass is lost as pyruvate in the pathway. Therefore, the yield results in this study compare favorably with those previously reported with the added benefit of an enantiomerically-pure D-glyceric acid product and a novel, renewable substrate. Additionally, production in minimal medium was also achieved with varied success depending on the secondary sugar that was co-fed, with galactose and glycerol enabling the best results. The development of a substrate-based expression platform showed the utility of the system for the reduction of cost in processing and demonstrated performance similar to IPTG-based designs. Finally, the development of a simple purification scheme allowed for the enrichment of D-glyceric acid in solution to enable its usage in chemical processing for research into polymeric applications. This work demonstrates the development and optimization of a complete production system to allow for industrially-relevant bioprocessing and displays many strategies that could be applied to other production platforms to increase their utility.

3. CCR Relaxation Strategies for Global and Sugar-Specific Alleviation

Abstract

The development of bioprocessing to produce value-added chemicals has the potential to improve the sustainability of the chemicals industry. This is largely due to the ability of biological organisms, like *Escherichia coli* (*E. coli*), to utilize renewable resources like food and agricultural wastes as feedstock. Therefore, the optimality of bioprocesses depends partially on the co-utilization of the many sugars present in these wastes. For sugar mixtures containing glucose, this presents an engineering challenge due to the effects of carbon catabolite repression (CCR) which limits the utilization of secondary sugars by *E. coli* when glucose is co-fed. Here we present two strategies for CCR relaxation that were investigated; one employed a global alleviation strategy and the other utilized a sugar-specific strategy. A mutation in $EIIA^{glc}$, an essential part of the phosphoenolpyruvate transferase system (PTS), was investigated as a global CCR relaxation strategy due to its ability to lock the protein in its phosphorylated state, mimicking a lack of glucose. While this engineered strain did co-utilize sugar mixtures, the phenotype was not stable. This result echoes the growth and substrate utilization difficulties that are prevalent in global CCR relaxation strategies presented in the literature. To enable sugar-specific CCR relaxation, the galacturonate-specific permease ExuT was engineered to lessen the effects of inducer exclusion. A galacturonate-specific biosensor was utilized to perform high-throughput screening of an *exuT* mutant library to search for mutants that enabled higher levels of intracellular galacturonate. Utilizing a synthetic pathway to produce D-glyceric acid from galacturonate, a S391R mutant of ExuT increased titer by 20% when a co-feed of galacturonate and glucose was used. These results demonstrate the utility of high-throughput screening for the engineering of sugar transporters for increased production when co-feeding sugars.

3.1 Introduction

The rise of crude oil prices, depletion of non-renewable resources, and detrimental environmental effect of many chemical processes suggest a need for both alternative feedstock and methods for producing goods. Alternative feedstocks, such as complex polysaccharides (cellulose, hemicellulose, pectin), can be obtained from agricultural byproducts such as corn stover or orange peel [16]. A plethora of simple sugars (glucose, lactose, galacturonate) can be extracted from these feedstock and then utilized for the synthesis of a variety of chemicals [135]. Bioprocessing by model organisms, such as *E. coli*, can allow for the creation of these products from a renewable sugar feedstock in a non-intensive and selective manner [136]. Much work has been done to explore the re-programming of the genetics of *E. coli* to optimize both synthetic and native pathways and improve the production of value-added chemicals. However, the applicability of many pathways for use with mixed-carbon feedstock is often hindered by carbon catabolite repression (CCR), an intrinsic carbohydrate consumption regulation network.

CCR-based regulation is observed in many microbes and its effects on cell physiology have been studied extensively [19]. In *E. coli*, CCR is primarily realized through the phosphoenolpyruvate sugar phosphotransferase system (PTS) (Figure 1.1) which phosphorylates glucose upon its transport into the cell. The PTS interacts with secondary-sugar catabolism networks within *E. coli* through both transcriptional regulation and inducer exclusion. Transcriptional regulation involves the activation of secondary-sugar catabolic genes through the cyclic AMP receptor protein (CRP)-cyclic AMP (cAMP) complex [20]. Phosphorylated $EIIA^{glc}$, a key protein in CCR regulation, activates adenylyl cyclase, which cyclizes ATP into cAMP and leads to an increased concentration of cAMP-CRP complex. The inducer exclusion phenomena involves the interaction of non-phosphorylated $EIIA^{glc}$ with secondary-sugar permeases, inhibiting them [21]. This prevents the

transport of secondary sugars and therefore both inhibits the activation of catabolic genes, which is often induced by the secondary sugar, and prevents the catabolism of the sugar due to its absence from the cytosol. The phosphorylation state of $EIIA^{glc}$ is determined by the glucose transport rate of the PTS, with more transport correlating to more $EIIA^{glc}$ in a non-phosphorylated state. Therefore, the phosphorylation state of $EIIA^{glc}$ is the primary switch in *E. coli* that modulates the cell between catabolite repression and secondary-sugar utilization.

The potential upside of carbon source co-utilization in bioprocessing has led scientists to investigate strategies for the alleviation of CCR in *E. coli*. These strategies generally fall into two categories: global remediation, which modifies central signaling pathways, and sugar-specific remediation, which engineers orthogonal sugar-specific systems. Global relaxation strategies often involve the mutation of the PTS, which has been shown to increase the titer of xylonate from a mixed sugar feed [30]. However, modifications to central signaling pathways, like the PTS, generally decrease both the growth rate of the mutant strain as well as its ability to utilize glucose [31]. Therefore, additional interventions are often necessary to restore the mutants ability to consume glucose, such as employing adaptive evolution [33]. Sugar-specific relaxation strategies are often more difficult to develop due to the necessity of orthogonality of the regulation network being engineered. A mutant XylR, with amino acid substitutions R121C and P363S, was found to increase the activation of the xylose operon in the presence of glucose [58]. Additionally, it has been suggested that the engineering of transport systems could be another way to accomplish sugar-specific alleviation [61, 65]. This chapter investigates novel methods to relax CCR effects through both the global and sugar-specific approaches. Depending on the medium formulation and application, it is possible that one of these approaches would be preferred over the other, so the development of both cases is important for the further advancement of bioprocessing technology.

3.1.1 Global Alleviation through the Mutation of EIIA^{glc}

The two primary cell-wide signaling cascades that are a part of the PTS are the cAMP and EIIA^{glc} systems. Modification of cAMP production is often detrimental to cell growth, reducing its utility for engineering solutions [137]. Since the phosphorylation state of EIIA^{glc} is a primary determinant for the sugar preference of *E. coli,* its modification provides an opportunity to globally relax CCR. Presper et al. discovered that mutating the active sites of EIIA^{glc} altered its ability as both a phosphate donor and acceptor [138]. Two active histidine residues act to both accept a phosphate from phosphoenolpyruvate and then donate that phosphate to glucose upon its transport into the cytosol. The mutation H90Q completely prevented the protein from being phosphorylated, while the mutation H75Q allowed the protein to be phosphorylated, but prevented its ability to donate a phosphate [138]. This chapter explores the effect that introducing the H75Q mutation to EIIA^{glc} has on the co-consumption characteristics of *E. coli*. Analysis of mutants of EIIA^{glc} have previously shown that their modification affects cAMP levels [139]; however, the specific mutants presented in the Presper et al. study have not been analyzed for their effect on sugar co-catabolism.

3.1.2 Sugar-Specific Alleviation through the Directed Evolution of a Galacturonate Permease

Inducer exclusion regulates secondary sugar catabolism on a sugar-specific basis by the interaction of non-phosphorylated EIIA^{glc} with sugar-specific permeases, such as LacY [21]. ExuT, a galacturonate permease present in *E. coli*, is likely a proton symporter like LacY and therefore is anticipated have a similar inducer exclusion effect. Mutations in the LacY [140], MalK [140], and RafB [141] permeases have been shown to reduce the effect of inducer exclusion by weakening the strength of the interaction of non-phosphorylated EIIA^{glc} with the transporter. In the example of LacY, one of the most studied permeases in the *E. coli* proteome, it is expected that the transporter interacts with non-phosphorylated EIIA^{glc} at a cytoplasmic loop between two

transmembranal spanners [142]. Therefore, it is hypothesized that the inducer exclusion effect that is likely occurring in the galacturonate transport system could be relaxed with the addition of mutations that weaken the interaction of non-phosphorylated $EIIA^{glc}$ with ExuT.

Our group developed a galacturonate-specific biosensor that responds to intracellular concentrations of galacturonate by inducing transcription of green fluorescent protein (GFP) [143]. Figure 3.1 displays the effect that the introduction of glucose to the growth medium has on the action of the biosensor. Galacturonate, which interacts with ExuR to enable transcription of *gfp*, has its transport hindered by the interaction of non-phosphorylated $EIIA^{glc}$ with ExuT. Therefore, the fluorescence that is observed without glucose is higher than that with glucose. This system establishes the opportunity to use the galacturonate biosensor as a screening mechanism for mutant strains with alleviated inducer exclusion effects.

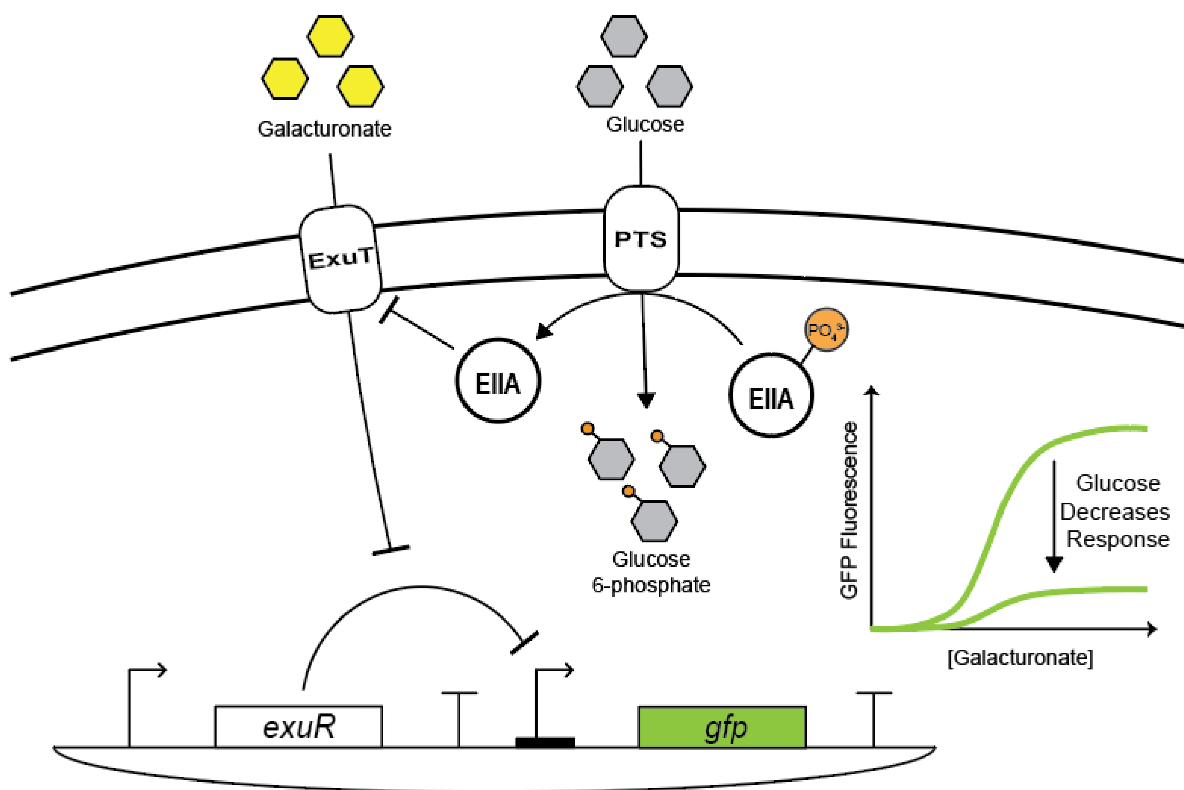

Figure 3-1. Effect of glucose on the action of a galacturonate-specific biosensor. The transcription factor-based biosensor was tuned to express *gfp* as a function of galacturonate concentration [143]. In the presence of glucose, inducer exclusion hinders the uptake of galacturonate into the cell and reduces the *gfp* expression of the biosensor.

Directed evolution is a framework that mimics nature by screening a diverse population of mutants for a desired phenotype and then selecting for top respondents to evolve the population towards better performance [87]. The value of directed evolution is improved when biosensor-based screening and FACS are used to find optimal mutants, increasing the throughput of the process [93]. This process has been extensively used to modify heterologous proteins to increase their activity in *E. coli* [95-97]. Figure 3.2 displays a flowchart for the application of biosensor-based directed evolution for the mutation of ExuT to increase galacturonate transport in the presence of glucose. Genetic diversity was obtained within the *exuT* gene by applying an error-prone PCR procedure to generate a mutant gene library. This gene library was ligated with a constitutive expression vector backbone and then used to transform the sensor strain introduced in Figure 3.1. The resulting mutant strain library was incubated in both glucose and galacturonate and then sorted using FACS to find the top GFP performing mutants. The top performing mutants were isolated and tested in an application for galacturonate usage in the presence of glucose to confirm their activity. For this study, the D-glyceric acid pathway was used as the application (Chapter 2). The pathway utilizes two heterologous enzymes, uronate dehydrogenase (*udh*) and galactarolactone isomerase (*gli*) to convert galacturonate to galactarate. Galactarate is then converted to D-glyceric acid by three endogenous proteins, GarD, GarL, and GarR. D-Glyceric acid production in mutant strains was determined by HPLC and top performers were then subjected to additional rounds of directed evolution.

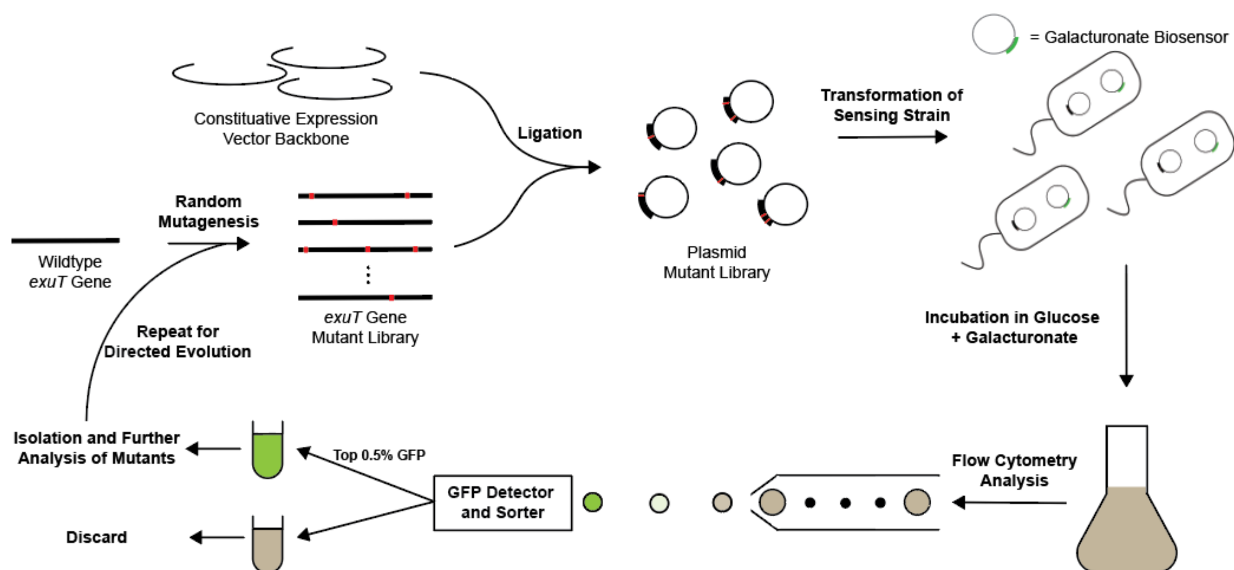

Figure 3-2. Schematic of the procedure for the directed evolution of *exuT* to increase galacturonate transport in the presence of glucose. Random mutagenesis is performed on the wildtype *exuT* gene to generate a mutant library of *exuT* genes. A plasmid library is generated through the ligation of the mutant gene library with a constitutive expression vector backbone. A Δ*exuT E. coli* strain is then transformed with both the galacturonate biosensor (Figure 3.1) and the mutant plasmid library. The resultant mutant strains are then incubated in a mixture of both glucose and galacturonate overnight. FACS is then used to sort for top performing mutants, and then collected strains are isolated and tested separately for their efficacy. A top performing mutant is then used as the template for additional rounds of random mutagenesis and testing.

3.2 Materials and Methods

3.2.1 E. coli Strain and Plasmid Construction

E. coli strain DH5α was used for all DNA manipulation procedures. These procedures were performed in accordance with standard practices [125]. Cultures were propagated in Luria-Bertani (LB) broth (BD, Franklin Lakes, NJ) and the Zyppy Plasmid Miniprep kit was used for all plasmid isolations (Zymo Research, Irvine, CA). New England Biolabs (Ipswitch, MA) restriction enzymes were used for all digestion procedures. All genes overexpressed on plasmids in this study, including *crr, exuT, garR, garL,* and *garD,* were amplified from the MG1655 *E. coli* genome with primers found in Table 3.1 (e.g., garL_NcoI_F and garL_HindIII_R for *garL*). These were then genetically manipulated to be expressed on a plasmid by using either the digest-ligate or Golden Gate cloning method [127]. The *crr* gene was inserted into a pCOLADuet backbone using the BamHI and SacI restriction sites (primers crr_BamHI_F and crr_SacI_R). The *exuT* gene was inserted into a vector with a p15a origin and constitutive promoter using the Golden Gate ligation method (primers exuT_GG_F_1, exuT_GG_R_1, exuT_GG_F_2, and exuT_GG_R_2). The vector expressing *garR, garL,* and *garD* was generated by a combination of the digest-ligate and Golden Gate cloning methods. First, the *garL* and *garD* genes were inserted into a pCloDF13Duet vector using restriction sites NcoI and HindIII for *garL* and NdeI and AvrII for *garD*. The *garR* gene was inserted into a separate pCloDF13Duet vector using the NdeI and AvrII restriction sites. A Golden Gate procedure was then used to insert the entire *garR* expression cassette into the vector containing *garL* and *garD*. Directed point mutations in genes were introduced using the Site Directed Mutagenesis kit from New England Biolabs and primers in Table 3.1 (e.g., crr_H75Q_F and crr_H75Q_R for mutation H75Q in *crr*). Error-prone PCR was performed using the GeneMorph II Random Mutagenesis kit from Agilent Technologies (Santa Clara, CA) using

primers epPCR_exuT_F and epPCR_exuT_R. The *exuT* library generated from this procedure was then ligated into the constitutive-expression vector using the EcoRI and HindIII restriction sites.

Chromosomal inactivation of the genes *exuT* and *crr* was completed using the procedure as described by Datsenko and Wanner [128]. Homology region lengths of at least 100 bp, employing two rounds of nested PCR (primers in Table 3.1), were used for the inactivation procedure (e.g., primer pairs dexuT_F_1 + dexuT_R_1 and dexuT_F_2 + dexuT_R_2 to inactivate gene *exuT*). Insertion of mutant *exuT* genes into the genome was completed using the clonetegration procedure with primers exuT_POSIP_F, exuT_POSIP_R, Clonetegration_P1, Clonetegration_P2, Clonetegration_P3, and Clonetegration_P4 for insertion into the 186 site of the genome [144]. When necessary, these mutations were made in the strain MG1655(DE3), carrying the DE3 lysogen encoding the T7 polymerase. Table 3.1 displays the primers used for verification of strain mutations and plasmid sequences. Tables 3.2 and 3.3 contain the plasmids and strains, respectively, used in this study.

3.2.2 Culture Conditions for Mutant EIIA^{glc} Strain Fermentation of Sugar Mixtures

Cultures were propagated in LB medium for later inoculation into fermentation medium. Fermentations were done in MOPS minimal medium supplemented with the necessary sugars and 0.1 mM IPTG. All chemicals for the minimal medium were purchased from Sigma-Aldrich (St. Louis, Mo). The MOPS minimal medium was prepared from a 10x stock that contained 0.4 M MOPS, 0.04 M tricine, 0.1 mM $FeSO_4$, 95 mM NH_4Cl, 2.76 mM K_2SO_4, 0.005 mM $CaCl_2$, 3 mM $MgCl_2$, 500 mM NaCl, 0.03 µM $(NH_4)_6Mo_7O_{24}$, 4 µM H_3BO_3, 0.55 µM $CoCl_2$, 0.15 µM $CuSO_4$, 1.27 µM $MnCl_2$ and 0.17 µM $ZnSO_4$. This stock was diluted, supplemented with 1.3 mM K_2HPO_4, and titrated with NaOH to achieve a final pH of 7.2. At this point, filtered concentrated sugar mixtures were added to achieve the desired concentration as dictated in the Results and Discussion

section. To prepare for fermentation, a media train of strain inoculation into LB, then to MOPS medium without IPTG, and finally to the fermentation medium (MOPS) was followed to limit the media-induced stress on the strain. Fermentation cultures were inoculated at an OD_{600} of 0.05. Cultures were grown in tubes with a 10 ml working volume and incubated at 37°C with agitation at 250 RPM.

Table 3-1. Primer sequences used in genetic manipulation procedures.

Primer Name	Sequence (5' to 3')
dexuT_F_1	GACCCAGCGTAGTCATAACCCGTACTGGAAAAACTGCACGAA CACATTGGTGTAGGCTGGAGCTGCTTC
dexuT_R_1	GCGTAGACTCTTGCTGGCGAACCGCGGACCAGGCCAGCAACAG CGTGCCGTTCCGGGGATCCGTCGACC
dexuT_F_2	GCTCTGGCGACGCAGAACTCCGCCCTGGCGGCTATCGTGGAAA AAATGTGGACCCAGCGTAGTCATAACC
dexuT_R_2	TGAAAAACCGTCTGGCATCGTTGTGCCTTGATGAACGGCACGA ATCGCCAGCGTAGACTCTTGCTGGC
dexuT_F_check	CAGAACTCCGCCCTGGC
dexuT_R_check	CATCGTTGTGCCTTGATGAACGG
dcrr_F_1	GCTGGTTAACAAGTTCATTGAAGAAAAACAATCTGCTAATCC ACGAGATGCGGCCCAATTTACTGCTTAGGAGAAGATCTGTAGG CTGGAGCTGCTTCG
dcrr_R_1	GCGATGAATTGATTTTGCCGCCGCTGGCGGAAGCATAAAAAAA TGGCGCCGATGGGCGCCATTTTTCACTGCGGCAAGAATTCCGG GGATCCGTCGACC
dcrr_F_2	ATCCGTAACACGAACTTCGAAGATGCGAAGGTGTTAGCAGAGC AGGCTCTTGCTCAACCGACAACGGACGAGTTAATGACGCTGGTT AACAAGTTCATTGAAGAA
dcrr_R_2	AGGGAAGGCGTTAACCGATGCAGTGCACCGAGCGGGGTTGCGC GTACTGGAAGTGATGCGCTACACCCAGCAGCATGAGAGCGATG AATTGATTTTGCCG
dcrr_F_check	GTCGGAACTGGTATTTAACCAGACT
dcrr_R_check	ACCTTCCCGAAGCGTATCGA
exuT_POSIP_F	AAAAAAGGTACCTTGACAATTAATCATCCGGCTCGT
exuT_POSIP_R	AAAAAACTGCAGTTAATGTTGCGGTGCGGG
Clonetegration_P1	CTCATTCGAAACCACCCACCG
Clonetegration_P2	ACTTAACGGCTGACATGG
Clonetegration_P3	ACGAGTATCGAGATGGCA
Clonetegration_P4	GATCATCATGTTTATTGCGTGG
crr_BamHI_F	AAAAAAGGATCCATGGGTTTGTTCGATAAACTGAAATCTCTG
crr_SacI_R	AAAAAAGAGCTCTTACTTCTTGATGCGGATAACCGG
crr_H75Q_F	CAGGCATTCTCTATCGAATCTGATAGCG
crr_H75Q_R	GTTGGTTTCAAAGATTTTACCAATGGTGCC
pCOLA_seq_F	TTGTGAGCGGATAACAATTCCCCTG
pCOLA_seq_R	TAAGATGGGGAATTGTTATCCGCTCA
exuT_GG_F_1	GCCAAAGGTCTCCCACTGTCAGACCAAGTTTACTCA
exuT_GG_R_1	GCCAAAGGTCTCCTGTATTTAGAAAAATAAACAAAAAGAGTTTG
exuT_GG_F_2	GCCAAAGGTCTCGTACATTACGCCCCGCCCT
exuT_GG_R_2	GCCAAAGGTCTCGAGTGATCGGCACGTAAGAGG
exuT_F_check	TTGCGCCGACATCATAACGG
exuT_R_check	ATCAGACCGCTTCTGCGTTC
epPCR_exuT_F	CCCAAGGAGCGGAATTCATG

epPCR_exuT_R	CGCCAAAACAGCCAAGCTTTTA
exuT_I314N_F	ATGGCCCGGGTATGATCGG
exuT_I314N_R	TCATCAGCACTGCACCCAG
exuT_S391R_F	CGCCCGCTGTTCGCAGTTC
exuT_S391R_R	GAAGCCGATGGTGTCAGCC
exuT_S391R + F322Y_F	ACACCAACCCGTATGTCGCA
exuT_S391R + F322Y_R	ACAGGCCGATCATACCCGG
exuT_S391R + T245I_F	TTTTTAACGCGTGGATCCCG
exuT_S391R + T245I _R	TACCCCAGGCCGGTTC
exuT_S391R + T375S_F	TCGCTGTTTGCGCTGGTT
exuT_S391R + T375S _R	ACTTGCCAGCCATGCG
exuT_S391R + E131K_F	AAGCGTTCCATCGCAGTA
exuT_S391R + E131K _R	TTTCGCCGGGAACCATT
exuT_S391R + E131K+F253Y_F	ACATGTTTAAAGTTTACGGCTTTAACCTG
exuT_S391R + E131K+F253Y_R	ACAGCGGGATCCACG
garL_NcoI_F	AAAAAACCATGGATGAATAACGATGTTTTCCCGAATAAATTC
garL_HindIII_R	AAAAAAAAGCTTTTATTTTTTAAAGGTATCAGCCAGTTTCTG
garD_NdeI_F	AAAAAACATATGGCCAACATCGAAATCAG
garD_AvrII_R	AAAAAACCTAGGTCAGGTCACCGGTGC
garR_NdeI_F	AAAAAACATATGAAAGTTGGTTTTATTGGCCTGG
garR_AvrII_R	AAAAAACCTAGGTTAACGAGTAACTTCGACTTTCGCC
garL_garD_GG_F	TGCATGGGTCTCGATTGTACACGGCCGCATA
garL_garD_GG_R	TGCATGGGTCTCGTAACGATTACTTTCTGTTCGACT
garR_GG_F	TGCATGGGTCTCGGTTAATACGACTCACTATAGGGGA
garR_GG_R	TGCATGGGTCTCGCAATTTAACGAGTAACTTCGACTTTC
garL_F_check	CGCTCTCCCTTATGCGAC
garL_R_check	CTGTTCGACTTAAGCATTATGCG
garD_F_check	GGCCGCATAATCGAAATTAATACGA
garD_R_check	GTGGCAGCAGCCTAG
garR_F_check	AGTCGAACAGAAAGTAATCGTTAATACGAC
garR_R_check	GATTATGCGGCCGTGTACAA

Table 3-2. Plasmids used in this study.

Plasmid	Genotype	Reference
pCOLADuet_*crr*	ColA origin, IPTG-inducible expression of *E. coli crr*	this study
pCOLADuet_*crr*(H75Q)	ColA origin, IPTG-inducible expression of *E. coli crr*(H75Q)	this study
p15a_*exuT*	P15a origin, constitutive expression of *E. coli exuT*	this study
p15a_*exuT*(I314N)	P15a origin, constitutive expression of *E. coli exuT*(I314N)	this study
p15a_*exuT*(S391R)	P15a origin, constitutive expression of *E. coli exuT*(S391R)	this study
p15a_*exuT*(S391R+F322Y)	P15a origin, constitutive expression of *E. coli exuT*(S391R+F322Y)	this study
p15a_*exuT*(S391R+T245I)	P15a origin, constitutive expression of *E. coli exuT*(S391R+T245I)	this study
p15a_*exuT*(S391R+T375S)	P15a origin, constitutive expression of *E. coli exuT*(S391R+T375S)	this study
p15a_*exuT*(S391R+E131K)	P15a origin, constitutive expression of *E. coli exuT*(S391R+E131K)	this study
p15a_*exuT*(S391R+E131K+F253Y)	P15a origin, constitutive expression of *E. coli exuT*(S391R+E131K+F253Y)	this study
pCloDF13_*garL_garD*	CloDF13 origin, IPTG-inducible expression of *E. coli garL* and *garD*	this study
pCloDF13_*garL_garR*	CloDF13 origin, IPTG-inducible expression of *E. coli garL* and *garR*	this study
pCloDF13_*garL_garD_garR*	CloDF13 origin, IPTG-inducible expression of *E. coli garL, garD* and *garR*	this study
pColE1_*exuR_gfp*	ColE1 origin, galacturonate biosensor that enables a dose-response of gfp expression with galacturonate addition	[143]

Table 3-3. Strains used in this study.

Strain Genotype	Reference
MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*	this study
MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*Δ*exuT*	this study
MG1655Δ*exuT*	this study
MG1655(DE3)Δ*crr*	this study
MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*Δ*exuT* 186::*exuT*	this study
MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*Δ*exuT* 186::*exuT* (S391R)	this study

3.2.3 Directed Evolution Protocol for ExuT

To generate a large diversity in the plasmid library, electro-competent DH5α cells from New England Biolabs were used for DNA isolation from the ligation procedure described in Section 3.2.1. The transformed cells were incubated for one hour after electroporation and plated onto dishes with a diameter of 13.5 cm. This was performed for a total of 8 plates to increase the number of colonies isolated. After overnight growth at an incubation temperature of 30°C, the colonies on the plates were removed by applying SOB medium and scraping the surface of the agar. The cells were collected in a Falcon tube and diluted to an OD_{600} of approximately 4. A mini-prep plasmid extraction procedure was performed on the diluted sample to extract the plasmid library. The sensing strain, MG1655Δ*exuT* transformed with the galacturonate biosensor, was then transformed with the plasmid library. Again, after a one hour incubation period, the culture was plated on eight dishes with a diameter of 13.5 cm. After overnight growth at an incubation temperature of 30°C, 8-10 colonies from these plates were chosen and sequenced to determine the mutation rate of the library. The colonies were then scraped and collected off of these plates. The collected strain library was stored at -80°C after adding 15% glycerol.

To test for beneficial mutations, the frozen culture was added to LB medium supplemented with 10 g/l glucose and galacturonate and allowed to incubate at 37°C overnight. After incubation, the cell medium was centrifuged and the cell pellet was re-suspended with PBS + 1% w/v sucrose. After this wash step, the cells were immediately sorted with FACSAria III (BD Biosciences, Franklin Lakes, NJ) with a 488 nm laser and a 530/30 band-pass filter. Populations were gated on forward and side scatter to remove cell debris. The top 0.5% GFP-positive population was collected in 1 ml of LB and was incubated at 37°C for 30 minutes to recover. This culture was

expanded into a 5 ml culture with antibiotic selection and incubated at 37°C overnight. After outgrowth, the culture was plated on LB agar to yield single colonies for further testing.

Isolated colonies were inoculated into individual wells of a 48-well flower plate that each contained LB medium supplemented with glucose and galacturonate. This plate was incubated at 37°C and analyzed with a BioLector (m2p-labs, Baesweiler, Germany) to determine trends in both biomass and fluorescence. Cultures that displayed fluorescence confirming their high activity were sequenced to determine the mutations present in the *exuT* gene. These mutations were introduced into a newly cloned vector to remove any side effects from the sorting procedure. Finally, these vectors were introduced into a D-glyceric acid production system to determine the applicability of the mutations for the production of a value-added product from a sugar mixture.

3.2.4 Culture Conditions for D-Glyceric Acid Production

Cultures were grown in LB medium supplemented with 0.1 mM IPTG, galacturonate, and glucose. Sugar concentrations varied depending on the experiment and are listed in the Results and Discussion section. Fermentation cultures were inoculated at an OD_{600} of 0.05 and IPTG was added at hour 0. Cultures were grown in tubes with a working volume of 10 ml and incubated at 30°C with agitation set to 250 RPM.

3.2.5 Analysis Conditions for Sugar, D-Glyceric Acid, and Fluorescence Detection

Samples from sugar consumption experiments were taken at regular intervals of less than 24 hours. Samples for D-glyceric acid production were taken at 24 hour intervals unless otherwise stated. Collected samples were used for OD_{600} measurements and the supernatant was analyzed by high-performance liquid chromatography (HPLC). A 1200 Series Agilent Technologies instrument with an Aminex HPX-87H Ion Exclusion Column (Bio-Rad Laboratories, Hercules, CA) and refractive index detector were used in the chromatography experiments. The column temperature was

maintained at 65°C and the detector was at 35°C, with an isocratic mobile phase of 5 mM sulfuric acid run at 0.6 ml/min. A 22 minute method was used; approximate elution times are as follows: glucose (9.1 minutes), glycerol (13.6 minutes), lactose (7.9 minutes), D-galacturonate (8.6 minutes), and D-glyceric acid (11 minutes).

Fluorescence measurements were taken either using an Infinite F200pro plate reader (Tecan, Mannedorf, Switzerland) or with a BioLector. For analysis with the plate reader, cells were first washed twice with PBS to reduce auto-fluorescence effects introduced by LB medium. For the BioLector, 1 ml cultures were grown for 24 hours while the device recorded biomass and GFP.

3.3 Results and Discussion

3.3.1 Global Alleviation through the Mutation of EIIA^{glc}

To ensure that the sugar mixtures of interest elucidate CCR, fermentations in minimal medium were performed. These experiments displayed that glucose is preferentially consumed over both galacturonate and lactose (Figure 3.3). Additionally, when glucose is replaced by glycerol, a non-PTS sugar, the secondary sugars are consumed immediately. It was observed that the consumption of both galacturonate and lactose is faster than that of glycerol, suggesting additional sugar preferences other than just the glucose effect. This result is supported by previous reports that additional intrinsic sugar hierarchies are present in *E. coli* other than those caused by the PTS [56].

The EIIA^{glc} mutant H75Q was chosen to be investigated for its effect on the co-consumption of sugar mixtures due to its reported ability to remain phosphorylated independent of the glucose concentration of the medium [138]. A pCOLA vector with IPTG-inducible expression of the mutant EIIA^{glc} was constructed and a MG1655(DE3) Δ*crr* strain was transformed with this vector. This strain was cultured in MOPS minimal medium supplemented with both glucose and galacturonate and induced with 0.1 mM IPTG upon inoculation. Compared to controls, the mutant strain displayed a 3-fold longer lag period, indicating cell stress due to the expression of the mutant (Figure 3.4(A)). However, both glucose and galacturonate were consumed by the H75Q mutant strain at approximately the same rate once the culture began growing (Figure 3.4 (B) and (C)). The galacturonate consumption rate of the mutant strain was greater than that of the Δ*crr* control, indicating that the expression of the mutant protein was impacting catabolism. Interestingly, a strain transformed with an empty pCOLA vector also experienced a growth lag followed by sugar co-utilization, suggesting that it is possibly simply a stress response that led to this phenotype rather than the mutant protein itself (data not shown).

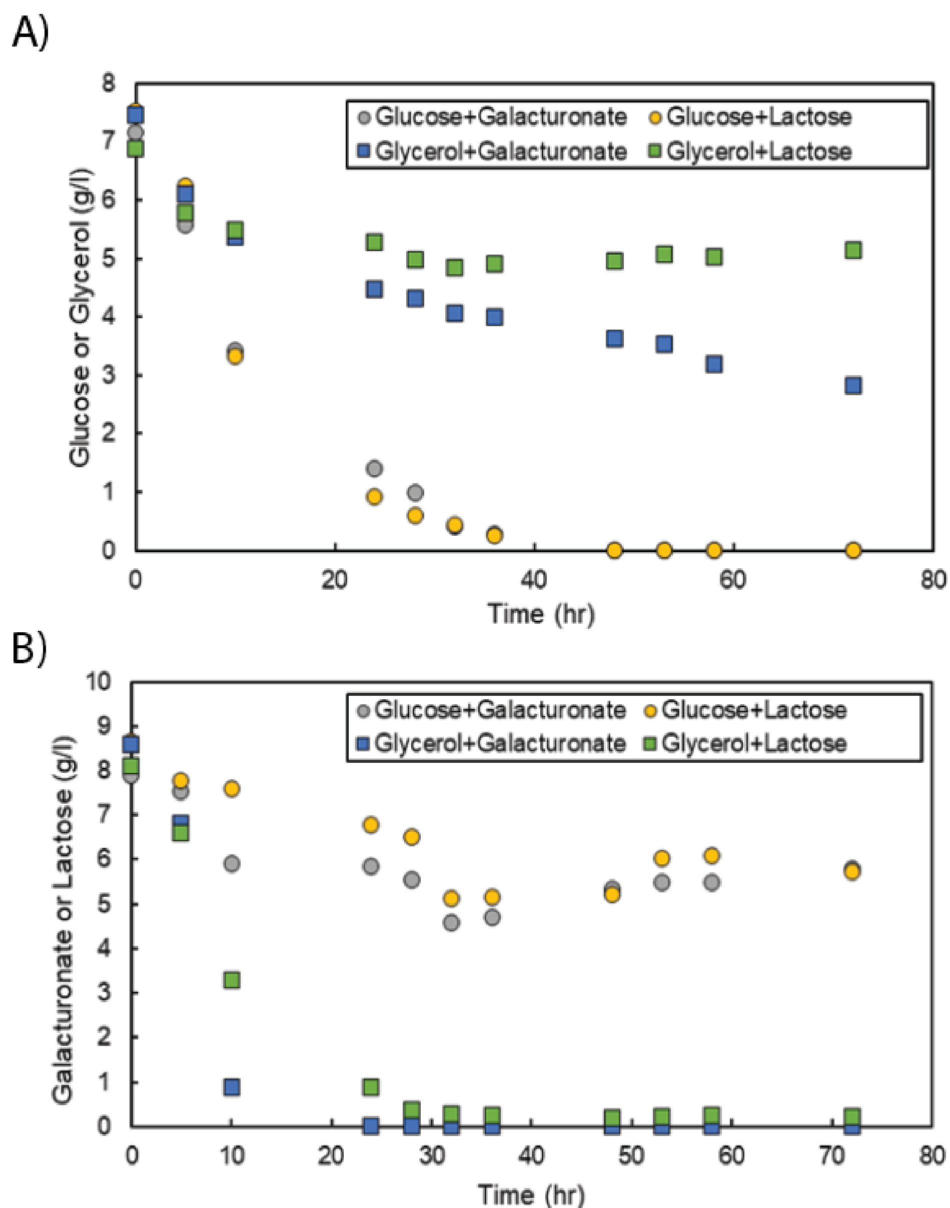

Figure 3-3. Sugar consumption trends of co-feeds of glucose and glycerol with galacturonate and lactose by wildtype *E. coli*. Consumption trends of (A) glucose and glycerol and (B) galacturonate and lactose are displayed. Fermentations were performed in MOPS medium with a working volume of 10 mL at 37 °C and a shaking rate of 250 RPM.

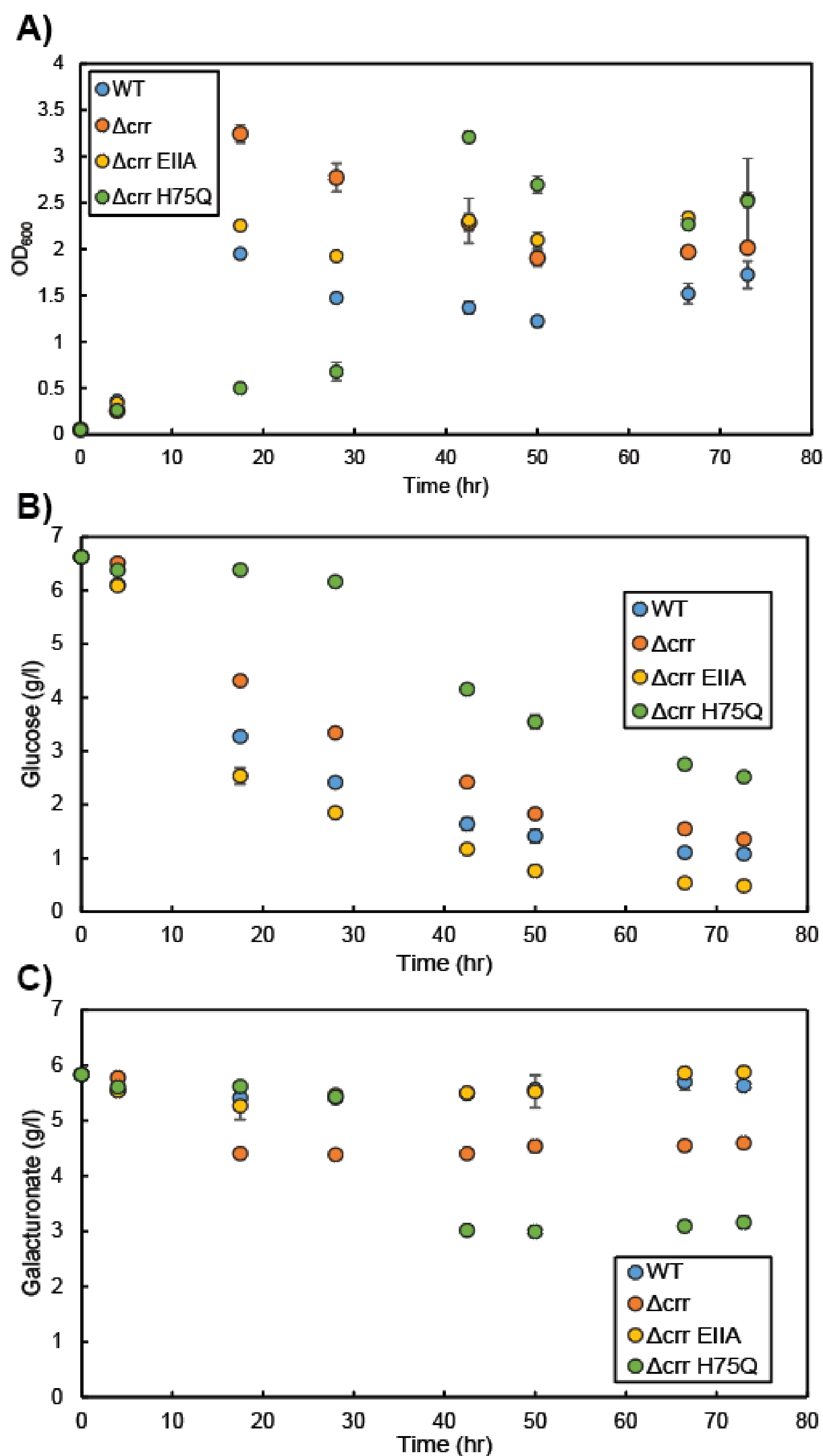

Figure 3-4. Fermentation of glucose-galacturonate co-feed by strain expressing H75Q mutant EIIA^{glc}. (A) Optical density, (B) glucose concentration, and (C) galacturonate concentration of cultures. All fermentations were performed in MOPS minimal medium supplemented with 0.1 mM IPTG at 37 °C and a shaking rate of 250 RPM. The working volume of the cultures was 10 mL.

The culture of the mutant strain from the end of the fermentation was re-inoculated into the same MOPS medium with glucose and galacturonate. This time, the growth delay was decreased significantly, and the co-consumption of the sugars occurred immediately (Figure 3.5). Additionally, upon storing the culture at -70 °C and then re-testing, the results in Figure 3.5 were reproduced. A similar test of the empty vector control displayed that the re-inoculation strategy did not lessen the lag experienced. Therefore, the expression of the mutant EIIA^{glc} is likely the cause of the phenotypic change observed.

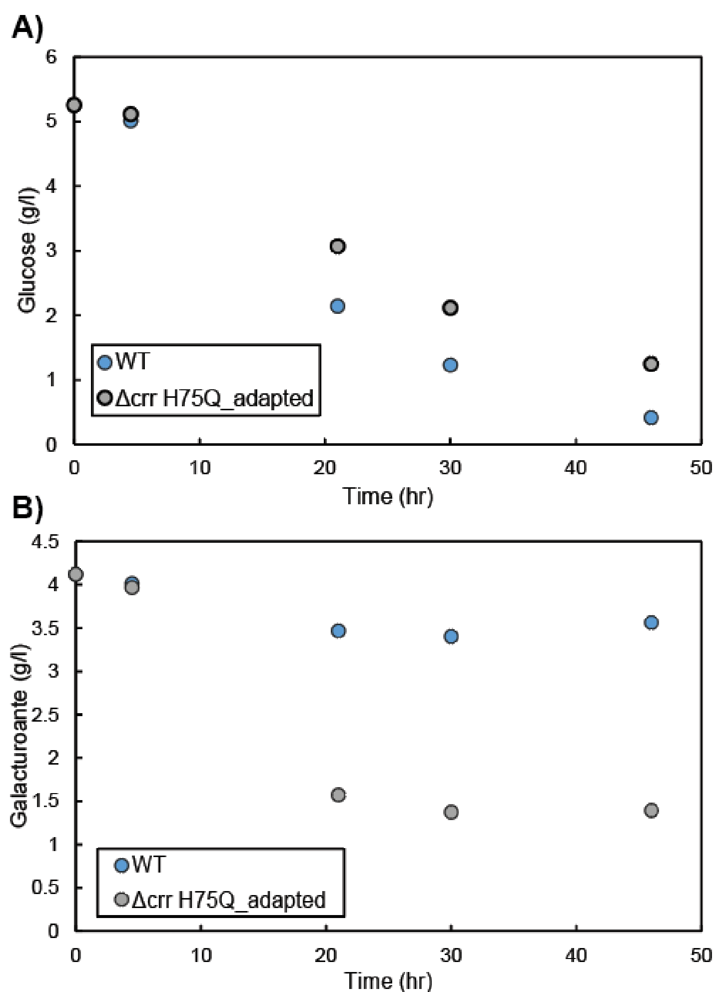

Figure 3-5. Fermentation of glucose and galacturonate co-feed by adapted strain expressing H75Q mutant EIIA^{glc}. (A) Glucose and (B) galacturonate concentrations of the adapted mutant strain compared to wildtype consumption. Fermentations were done in MOPS minimal medium supplemented with 0.1 mM IPTG at 37°C and a shaking rate of 250 RPM.

Since this approach is intended as a global CCR relaxation strategy, the mutant strain's ability to co-consume a mixture of three sugars was also investigated. The adapted mutant strain, from a freezer stock, was inoculated into MOPS minimal medium supplemented with glucose, galacturonate, and lactose. The wildtype strain, as expected, only consumed glucose and the addition of IPTG to the wildtype culture enabled a small amount of lactose consumption as well (Figure 3.6). This is likely due to the induction of the endogenous lactose operon caused by the lactose analogue, IPTG. It was anticipated that the Δ*crr* strain would have difficulties consuming glucose [31], however it was unexpected that this strain would be able to consume lactose, but not galacturonate. Further analysis revealed that the presence of galactose in place of lactose caused the same effect of galacturonate catabolism repression in this strain, even when glucose was not present. This may be explained through the intrinsic hierarchy of secondary sugar preference that *E. coli* possesses. The adapted mutant strain demonstrated an ability to consume all three sugars at approximately the same rate and outperformed the Δ*crr* strain in both glucose and galacturonate consumption rates. This suggests that there is something phenotypically different between the adapted mutant and the control strains. Full-genome sequencing did not reveal any discernable mutations in the adapted sample compared to the sample before the growth lag. This points to other modifications in the phenotype, which are more difficult to analyze. Therefore, it was decided to instead further characterize the sugar catabolism characteristics of the adapted mutant that was isolated.

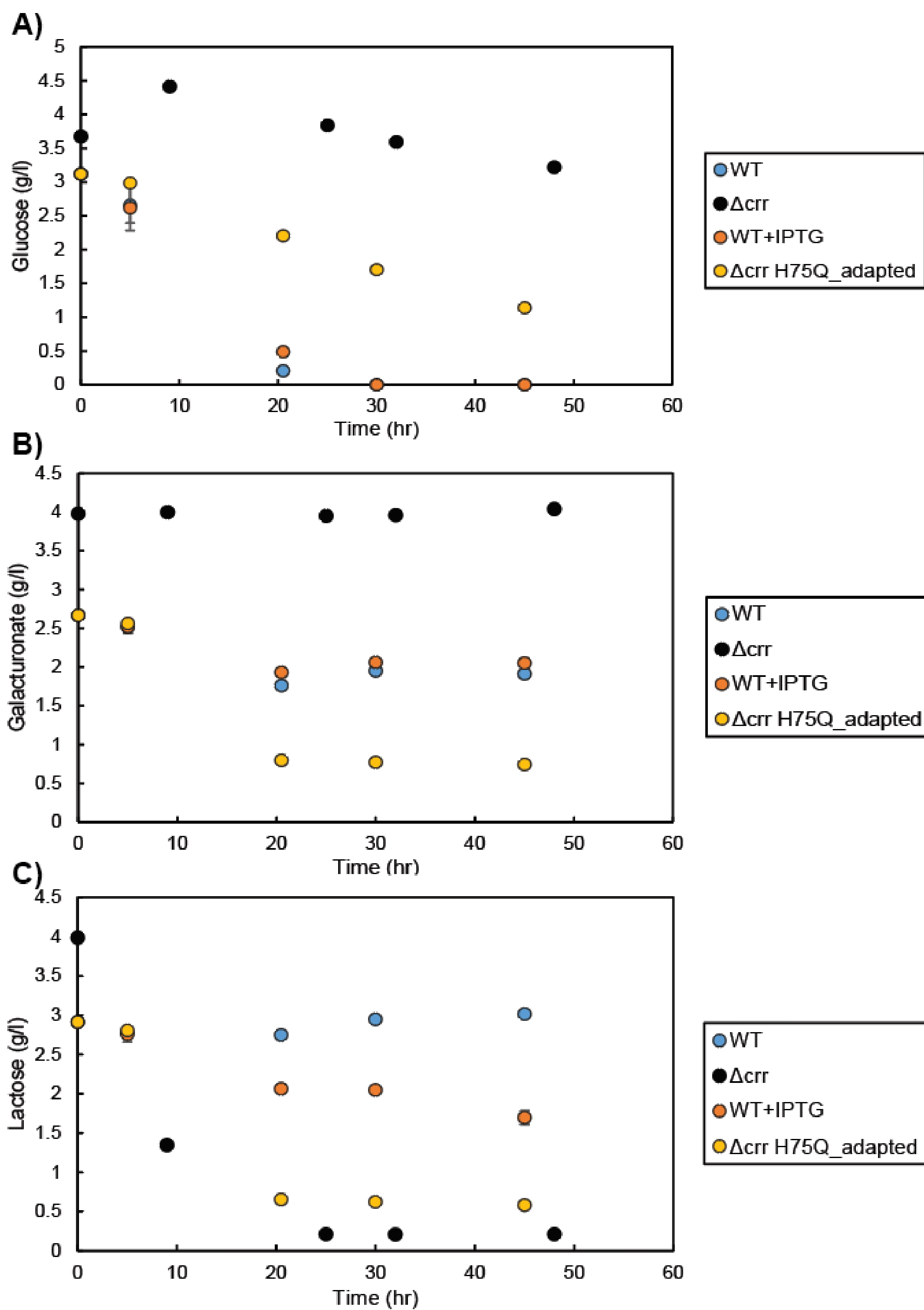

Figure 3-6. Fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing H75Q mutant $EIIA^{glc}$. (A) Glucose, (B) galacturonate, and (C) lactose concentration in fermentations with a strain expressing a H75Q mutant $EIIA^{glc}$ compared to a wildtype with and without IPTG and a Δ*crr* mutant. Fermentations were done in MOPS minimal medium supplemented with 0.1 mM IPTG when relevant at 37°C and a shaking rate of 250 RPM.

To characterize the homogeneity of the adapted mutant strain, clonal isolation was performed on agar plates made of both LB and MOPS media. Unexpectedly, the sugar consumption phenotypes of mutants isolated using different media differed from one another (Figure 3.7). The adapted strain isolated on LB medium plates displayed a sugar consumption phenotype resembling the Δ*crr* mutant, consuming glucose and lactose but not galacturonate. Alternatively, the adapted strain isolated with MOPS medium plates consumed glucose and galacturonate, but not lactose. This phenotype had not been observed previously in this study. Additionally, the medium used to expand the culture after plate-based isolation also influenced the sugar consumption phenotype (Figure 3.8). Transferring a colony isolated on a MOPS plate to MOPS minimal medium for fermentation led to a glucose and lactose consumption phenotype while transferring the clone to LB medium led to a glucose and galacturonate consumption phenotype. This suggests that the media train is important for the elucidation of a strain isolate that consumes glucose and galacturonate. Any deviation from growth on a MOPS plate for strain isolation and then propagation in LB medium leads to behavior similar to a strain that does not express $EIIA^{glc}$ (Δ*crr*). Inoculating a fermentation experiment with both the strain propagated in LB and MOPS media enables the experimental results from Figure 3.6 to be observed, with the consumption of glucose, galacturonate, and lactose. It is hypothesized that the initial growth lag experienced by the adapted strain created a co-culture of strains with the two sugar consumption phenotypes discussed. Further testing of strains isolated on MOPS plates by propagation in MOPS minimal medium demonstrated this co-culture diversity. Of three clones isolated from MOPS plates and propagated in MOPS minimal medium, two consumed glucose and lactose while the other consumed glucose and galacturonate (Figure 3.9).

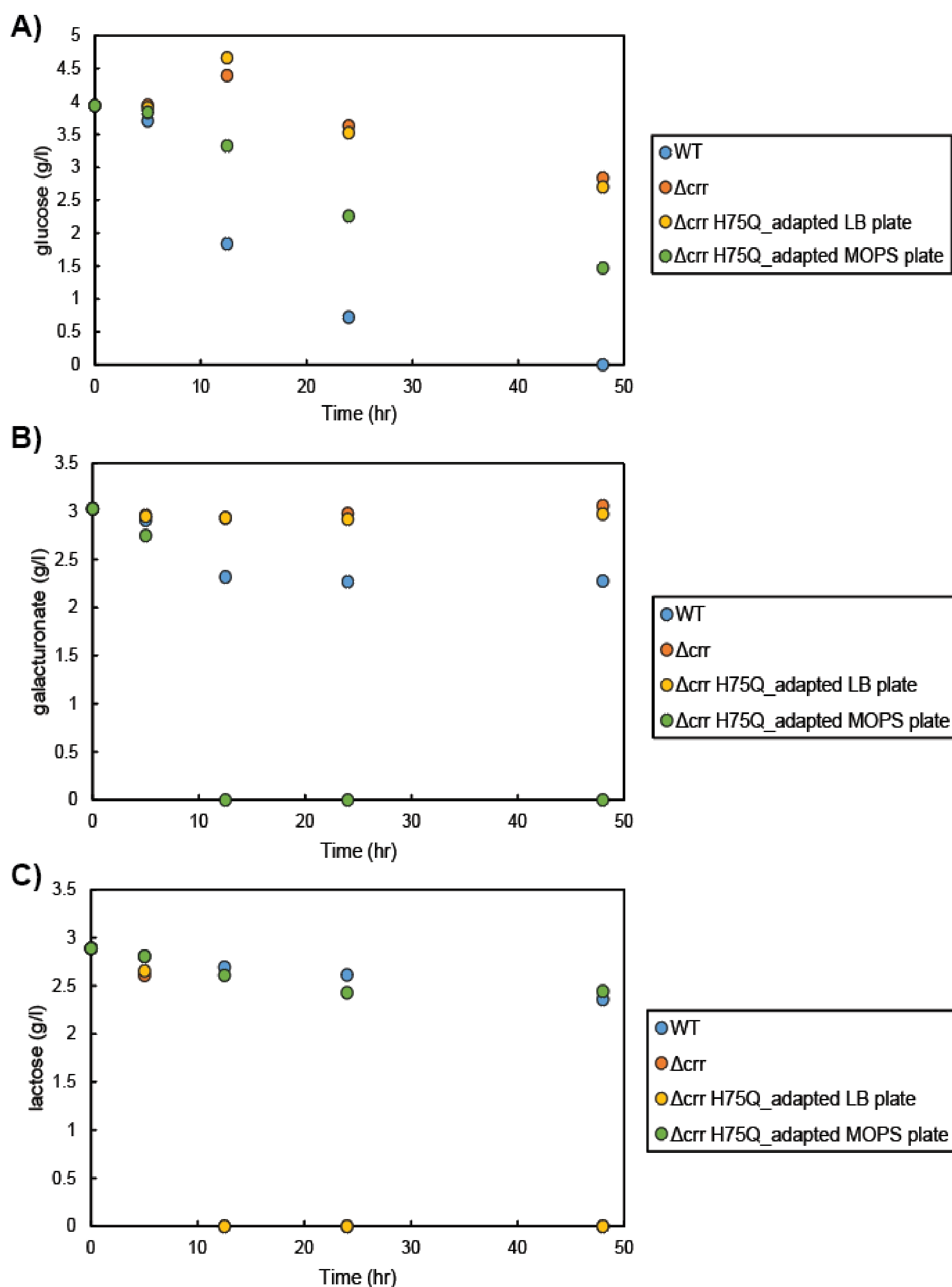

Figure 3-7. Fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing H75Q mutant EIIA^{glc} that was plated either on LB or MOPS agar. (A) Glucose, (B) galacturonate, and (C) lactose concentration in fermentations with a strain expressing a H75Q mutant EIIA^{glc} plated on either LB or MOPS agar compared to wildtype and a Δ*crr* mutant. Fermentations were done in MOPS minimal medium supplemented with 0.1 mM IPTG when relevant at 37°C and a shaking rate of 250 RPM.

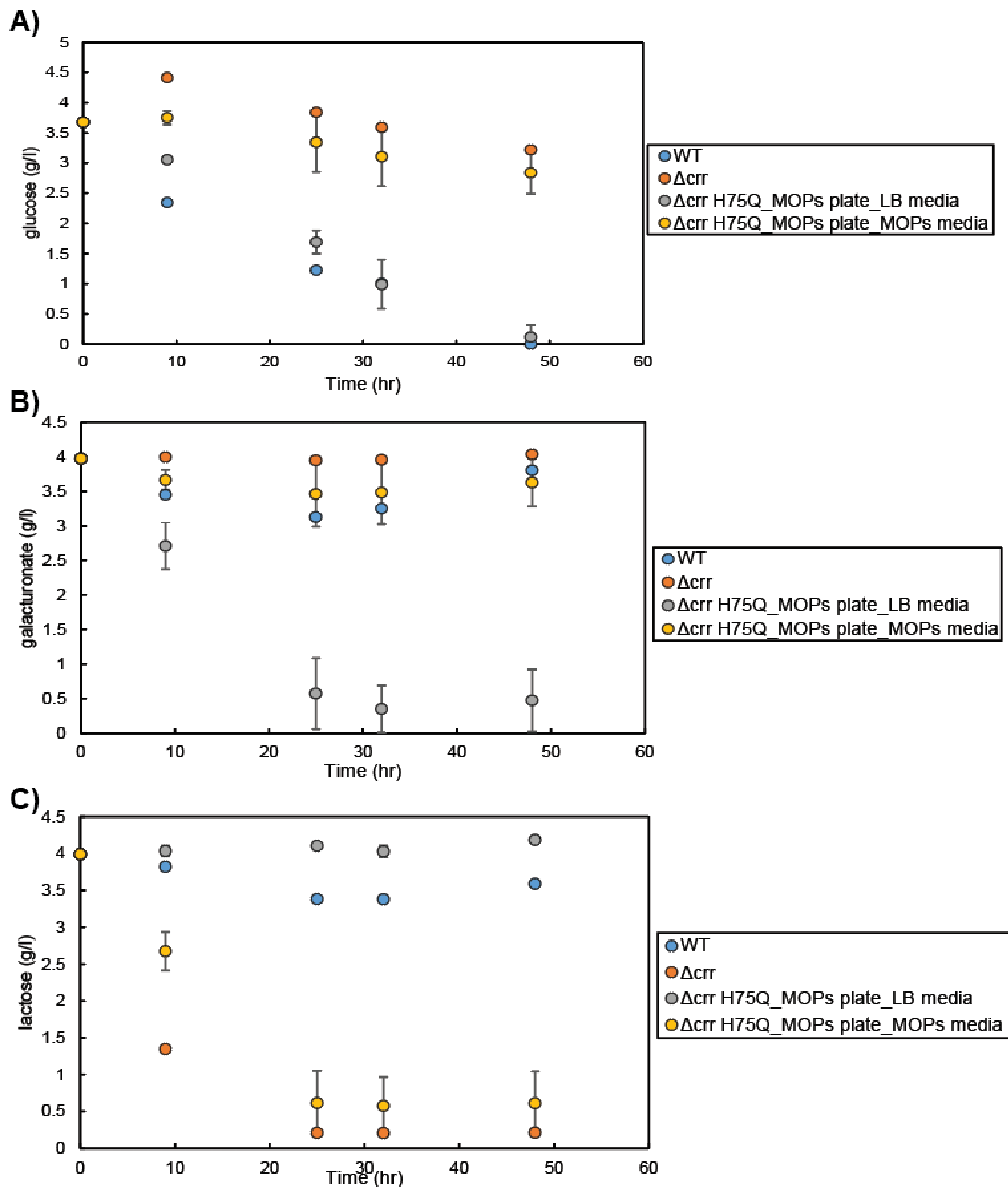

Figure 3-8. Fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing H75Q mutant EIIA^{glc} plated on MOPS agar and expanded in LB or MOPS medium. (A) Glucose, (B) galacturonate, and (C) lactose concentration in fermentations with a strain expressing H75Q mutant EIIA^{glc} expanded in either LB or MOPS medium compared to wildtype and a Δ*crr* mutant. Fermentations were done in MOPS minimal medium supplemented with 0.1 mM IPTG when relevant at 37°C and a shaking rate of 250 RPM.

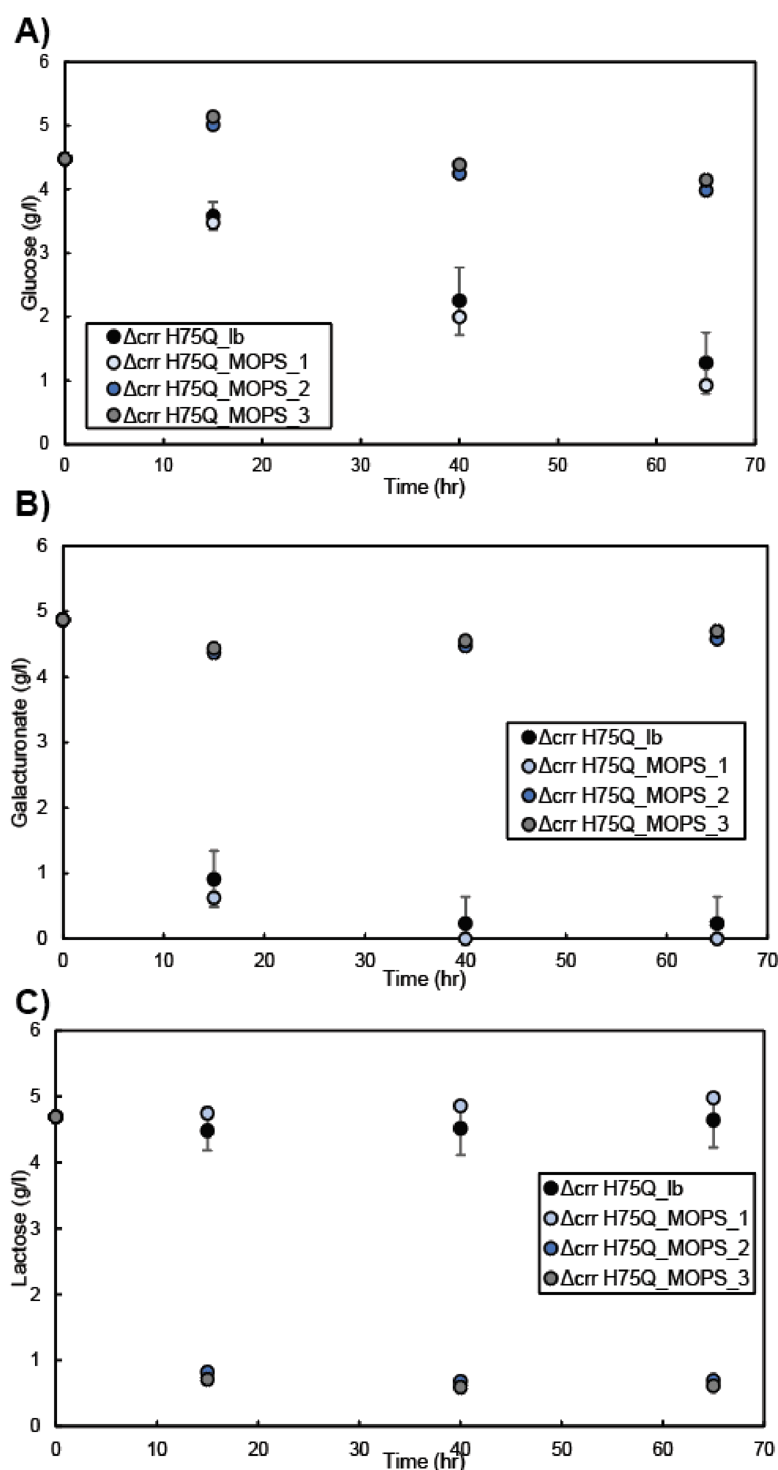

Figure 3-9. Clonal variation shown in fermentation of co-feed containing glucose, galacturonate, and lactose with adapted strain expressing mutant $EIIA^{glc}$ plated on MOPS agar and then expanded in MOPS medium. (A) Glucose, (B) galacturonate, and (C) lactose concentration in fermentations with a strain expressing a H75Q mutant $EIIA^{glc}$ expanded in either LB or MOPS medium. Fermentations were done in MOPS minimal medium supplemented with 0.1 mM IPTG when relevant at 37°C and a shaking rate of 250 RPM.

While the sugar consumption phenotypes that were observed during this study were interesting, the timescale of my research did not allow for a full investigation into the phenomenon. Therefore, the results observed showing the differing sugar consumption phenotypes are included to enable a future group or scientist to expand upon these ideas.

3.3.2 Sugar-Specific Alleviation through the Directed Evolution of a Galacturonate Permease

3.3.2.1 Screening System Validation

The efficacy of the galacturonate sensing system for the determination of intracellular galacturonate was validated prior to use in the screen. MG1655 Δ*exuT* was transformed with the galacturonate biosensor and co-fed galacturonate and glucose for comparison to a feed of pure galacturonate (Figure 3.10). As expected, the fluorescent response is effectively eliminated with the knock-out of the *exuT* gene, showing that the ExuT permease is the primary means for galacturonate transport. The signal response is recovered with the constitutive expression of *exuT* from a plasmid. The addition of glucose decreases the response upon the addition of galacturonate. This suggests that the signaling system demonstrates an ability to differentiate between galacturonate transport activity with and without glucose in the medium.

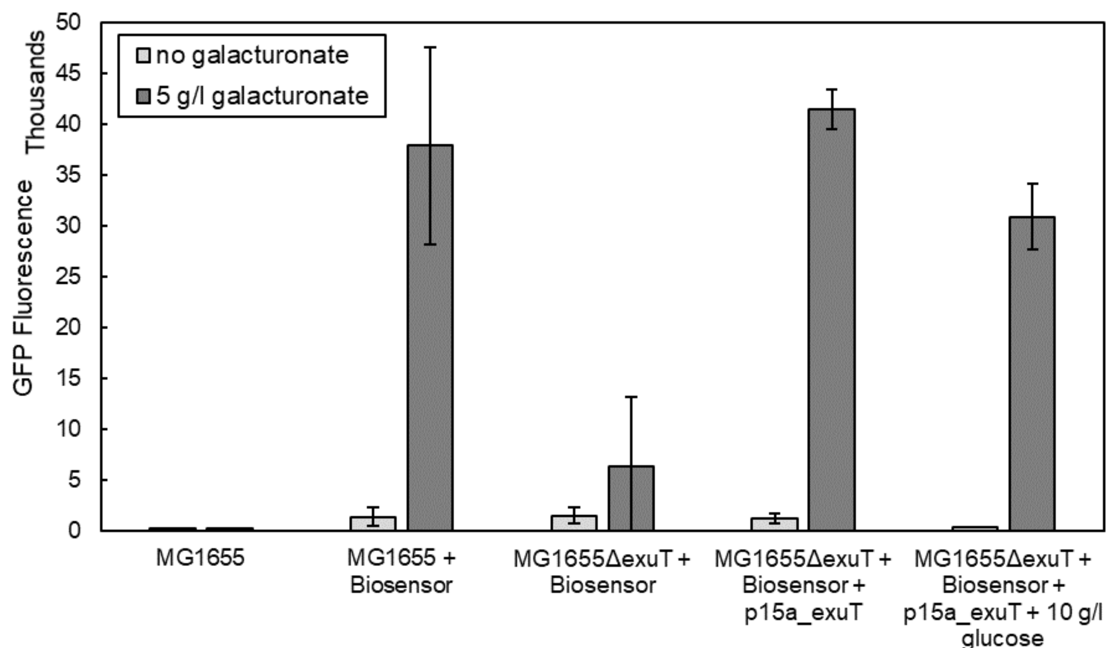

Figure 3-10. GFP response upon the addition of galacturonate to different strains transformed with the galacturonate biosensor. Cultures were grown in LB medium overnight with or without galacturonate, as labeled, washed with PBS and then analyzed using a plate reader to determine the GFP output.

3.3.2.2 First ExuT Mutant Library Screening and Testing

The first *exuT* mutant gene library had an error rate of approximately 3 base pairs/kb. This was ligated with a constitutive expression backbone with a p15a origin compatible with the galacturonate biosensor. Upon the transformation of the sensing strain with the mutant plasmid library, the strain library was incubated in LB medium supplemented with glucose and galacturonate overnight. Single-cell fluorescence data was collected using flow cytometry analysis (Figure 3.11). The negative control strain carried a plasmid that constitutively expressed the wildtype *exuT* gene and was fed both glucose and galacturonate. As expected, the largest population seen in the negative control is near the low end of GFP signal, indicating that galacturonate entry into the cell is being limited (Figure 3.11(A)). The positive control was a strain carrying the galacturonate biosensor cultured without any glucose added to the medium. This

sample showed a high GFP signal, indicating that galacturonate is entering the cell and able to activate the biosensor (Figure 3.11(B)). Interestingly, in both the negative and positive controls there were populations present on the opposite side of the GFP spectrum than what was expected. This suggests that there is a level of leakiness or inconsistency present in the signaling system on a single cell level. Analysis of the mutant strain library displayed a wide distribution of GFP signals, with three major populations observed (Figure 3.11(C)). A wide distribution was expected as a library provides variance both due to the mutagenesis of the *exuT* gene, as well as other genetic discrepancies that may be present from colony-level diversity. The top 0.5% GFP signaling population was isolated and collected from this experiment for further analysis. This population displayed a GFP signal close to that of the positive control, suggesting that beneficial mutations may have occurred. Upon propagation in rich medium, a portion of this population was plated to enable isolation of single mutant colonies for clonal analysis.

Isolated colonies were cultured in medium supplemented with glucose and galacturonate using a BioLector apparatus which measures GFP signal and biomass continuously (Figure 3.12). A number of mutants performed similarly to the positive control. Alternatively, many of the isolated clones from the top GFP signaling population performed similar to the negative control, which expressed wildtype *exuT*. False positives are a common problem when using flow cytometry for high-throughput screening [145], so this observation was not unexpected, and methods have been developed to try to lessen the effect of this problem [146]. Clones that displayed a ratio of GFP to biomass higher than 3, determined to be a positive signal, were sequenced to determine the mutation(s) present in the *exuT* gene being constitutively expressed from the plasmid.

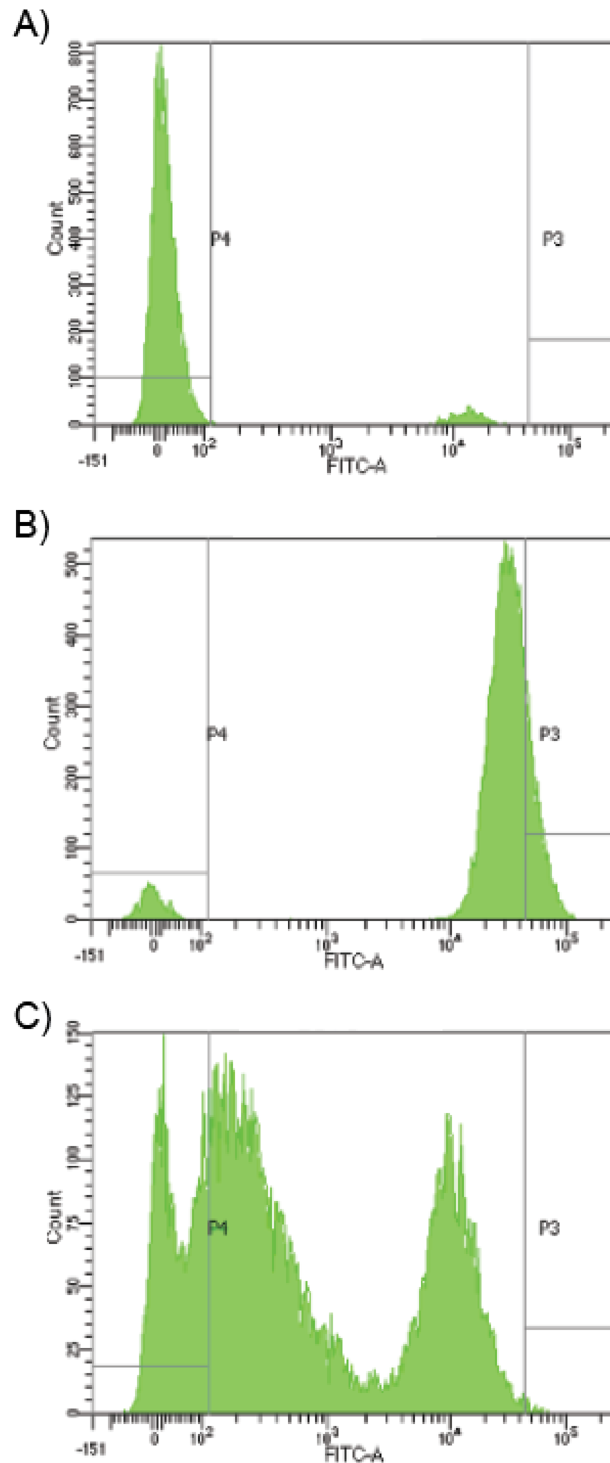

Figure 3-11. GFP distribution profiles for first ExuT mutant library sorting procedure, using flow cytometry. The (A) negative control was wildtype *exuT* expressed on a plasmid and fed both glucose and galacturonate. The (B) positive control was the wildtype *exuT* expressed on the genome fed only galacturonate. The (C) library was constructed as described and fed both glucose and galacturonate. The top 0.5% GFP positive population was collected for further analysis.

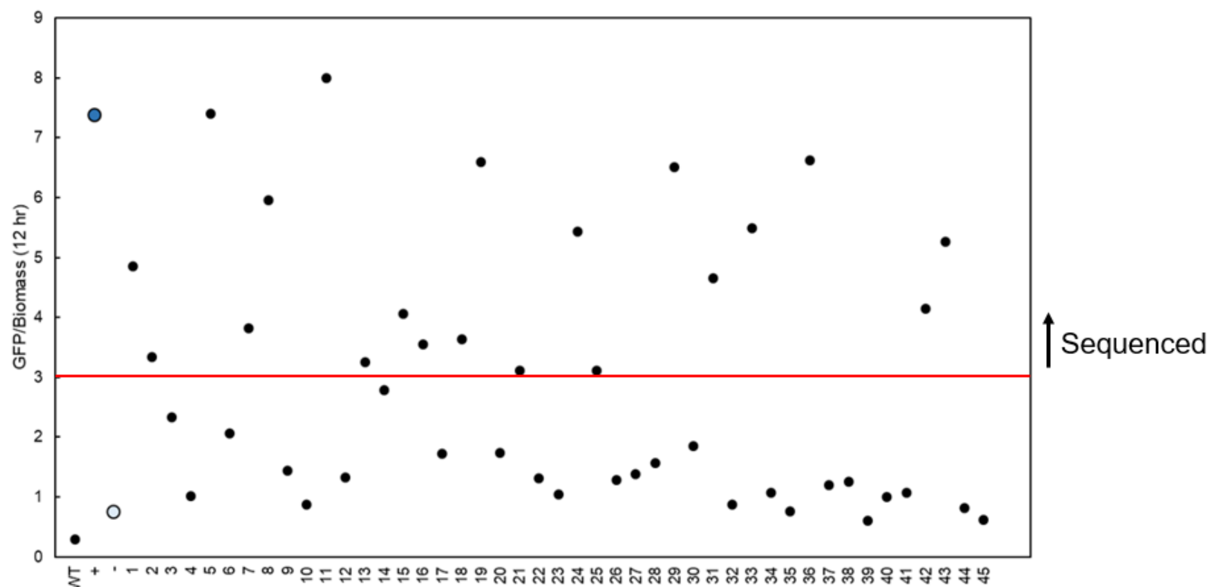

Figure 3-12. GFP analysis of isolated clones done in a BioLector from the sorting experiment of the first mutant ExuT library. Samples were cultured in LB medium supplemented with 5 g/l of glucose and galacturonate. The positive control is a sample without glucose fed and the negative control is expressing wildtype *exuT*. Any isolated colony that displayed activity above the red line shown was sequenced to determine the mutations present in the *exuT* gene.

Upon sequencing, it was discovered that almost all of the top performers contained one of two mutations: I314N or S391R. Additionally, one of the top performers was wildtype *exuT* and the I314N clones all also contained a silent mutation. It is anticipated that the large prevalence of the same two mutants was likely due to the method that was used for cell propagation after the sorting procedure. Growth advantages that may have been present in these mutants may have led to their dominance of the culture and therefore a large number of them being chosen during the colony isolation step. These mutations were then introduced into plasmids for utilization in the production of D-glyceric acid from a co-feed of glucose and galacturonate using the pathway introduced in Chapter 2.

MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*Δ*exuT* was transformed with both a plasmid constitutively expressing different variants of the *exuT* gene and a plasmid enabling IPTG-inducible expression

of the *udh* and *gli* genes necessary for D-glyceric acid production. When glucose and galacturonate were co-fed to this production system, the molar yields of D-glyceric acid from consumed galacturonate were depressed in comparison to previous results using this pathway (Figure 3.13). It is hypothesized that this was due to the lack of transcriptional activation of the endogenous genes that are a part of the D-glyceric acid pathway: *garD*, *garL*, and *garR*. It is possible that this effect had not been observed in the sugar co-feeding study presented in Chapter 2 due to the equal effect of catabolite repression on the expression of *exuT*. In this study, since *exuT* is being constitutively expressed, the effects of catabolite repression on the expression of these endogenous proteins is more apparent. The overexpression of *garD*, *garL*, and *garR* from a plasmid elevated the molar yield back to levels that were consistent with initial testing of the D-glyceric acid production platform. This displays the importance of eliminating the effect of transcriptional regulation when looking at the effect that the ExuT mutants have on D-glyceric acid production so that only inducer exclusion inhibition is being considered.

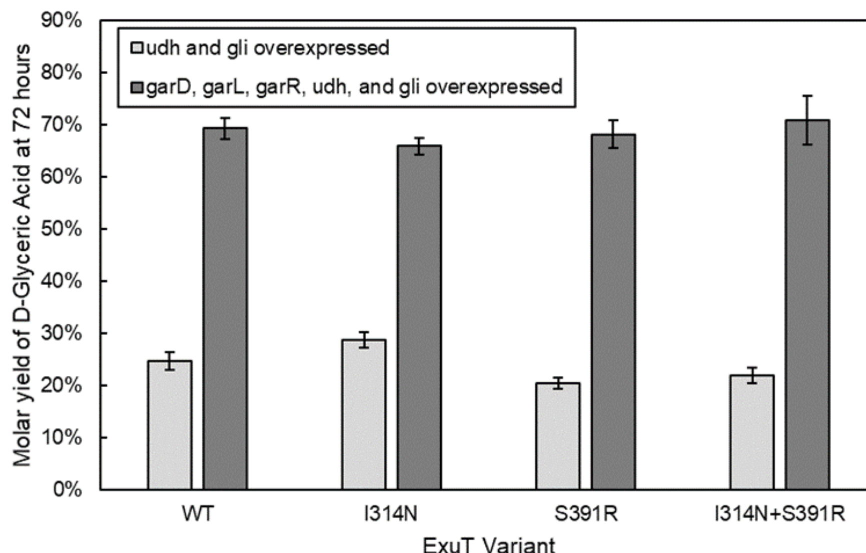

Figure 3-13. Relief of effect transcriptional inhibition on the molar yield of D-glyceric acid when utilizing plasmid-based expression of *exuT* and co-feeding glucose and galacturonate by overexpressing endogenous pathway genes. Cultures were grown in LB medium supplemented with 10 g/l of galacturonate, 5 g/l of glucose and 0.1 mM IPTG. Fermentations were done in tubes with 10 ml working volume and incubated at 30 °C with a shaking rate of 250 RPM.

Using this modified D-glyceric acid production platform, studies were performed to determine the efficacy of the isolated mutants for increased production in the presence of glucose. When fed pure galacturonate the S391R mutant performed similar to the wildtype, while the strain expressing the I314N mutant produced less D-glyceric acid (Figure 3.14). Additionally, when both mutants were introduced to the ExuT protein the system produced less D-glyceric acid compared to the wildtype. In this experiment, a successful mutant would perform the same as the wildtype as it would show that there is no transport limitation introduced by the mutation. When supplemented a co-feed of glucose and galacturonate, a similar trend is seen where strains that express an ExuT protein with the I314N mutation perform similarly, or worse, than the wildtype (Figure 3.15). This indicates that the I314N mutation is likely detrimental to transport action and is therefore not interesting for engineering applications. The S391R mutation, however, allowed the strain to produce 20% more D-glyceric acid in comparison the wildtype. A t-test comparing the titer when expressing a WT

exuT and a S391R mutant *exuT* gave a p-value of 0.006. This suggests that the mutant transporter is improving the ability for the system to utilize the galacturonate present in the medium for the production of D-glyceric acid. The titer of D-glyceric acid reported in this experiment by the wildtype condition is much higher than that reported in Chapter 2. However, this is expected as the entire production pathway, as well as the transporter itself, are being overexpressed in this experiment. This removes the transcriptional activation part of catabolite repression, which itself increases the ability of the system to utilize galacturonate. The larger increase in titer from the removal of transcriptional regulation in comparison to the attempt to alleviate inducer exclusion suggests that the former may dominate in this system.

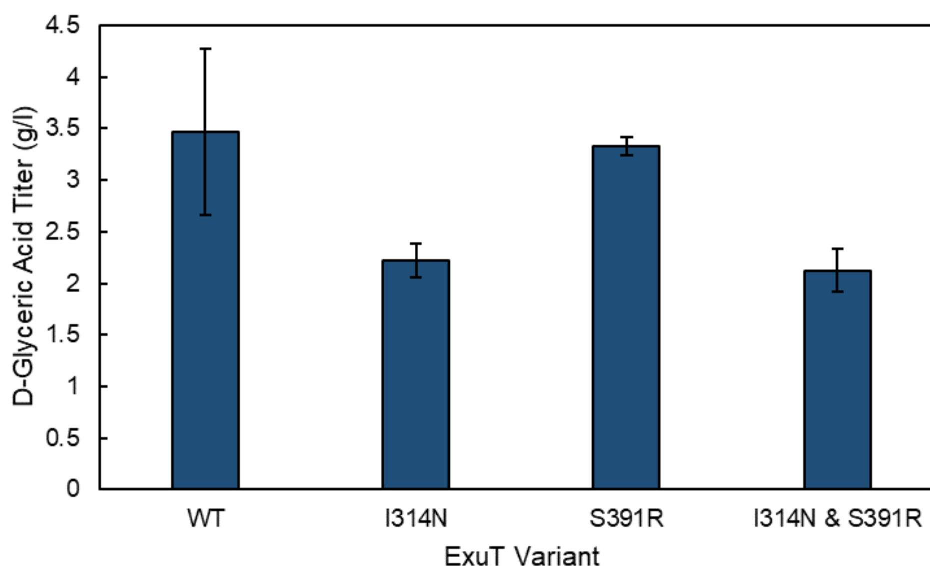

Figure 3-14. D-glyceric acid titer from the plasmid-based expression of different ExuT mutants with the feed of only galacturonate. Fermentation was performed in LB medium supplemented with 10 g/l of galacturonate and 0.1 mM of IPTG. Cultures were done in tubes with a 10 ml working volume, incubated at 30 °C and agitated at a shaking rate of 250 RPM. Samples were taken after 62 hours of fermentation.

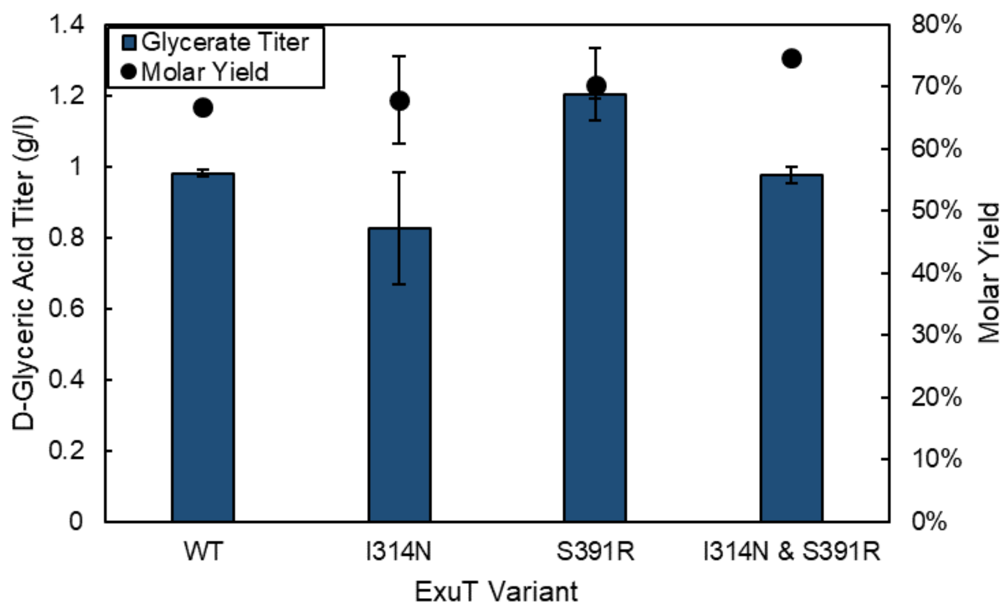

Figure 3-15. D-glyceric acid titer and molar yield from the plasmid-based expression of different ExuT mutants isolated from the first mutant library with a feed of glucose and galacturonate. Fermentation was performed in LB medium supplemented with 5 g/l of glucose, 10 g/l of galacturonate and 0.1 mM IPTG. Cultures were done in tubes with a 10 ml working volume, incubated at 30 °C and agitated at a shaking rate of 250 RPM. Samples were taken after 62 hours of fermentation.

To increase the stability of the system, constitutive expression cassettes for both the wildtype and S391R mutant *exuT* genes were inserted into the 186 phage insertion site of the genome of the MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*Δ*exuT* strain. These strains were transformed with plasmids enabling the overexpression of the entire D-glyceric acid production pathway and co-fed glucose and galacturonate. The results of this experiment were similar to when the genes were expressed via plasmid, with the mutant ExuT enabling an approximately 72% increase in D-glyceric acid titer (Figure 3.16). The decrease in titer is likely due to the lower copy number of the *exuT* gene intrinsic to genome-based expression. When no *exuT* gene was introduced into the genome or expressed via plasmid, no D-glyceric acid was observed (data not shown). These results suggest that the S391R mutation in ExuT is beneficial for the production of D-glyceric acid when

glucose and galacturonate are co-fed and therefore make it a good starting point for the next round of random mutagenesis and screening for continued directed evolution. It should also be noted that while the crystal structure of this protein is not known, it is anticipated by models that the serine residue 391 in the wildtype transporter resides in the cytosolic space [147]. This adds some validity to the beneficial S391R mutation as it would make sense that a cytosolic residue changing could alter the interaction between it and $EIIA^{glc}$, which is also present in the cytosol.

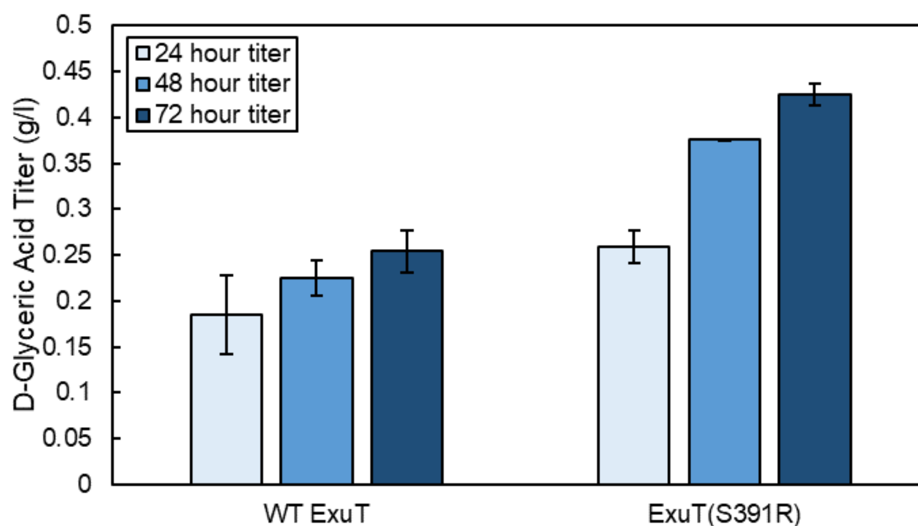

Figure 3-16. D-glyceric acid titer from the genome-based expression of wildtype and S391R ExuT with a feed of glucose and galacturonate. Fermentation was performed in LB medium supplemented with 10 g/l of glucose and galacturonate and 0.1 mM IPTG. Cultures were performed in tubes with a 10 ml working volume, incubated at 30 °C and agitated at a shaking rate of 250 RPM.

3.3.2.3 Second ExuT Mutant Library Screening and Testing

For the second round of directed evolution, MOPS minimal medium was used for cell propagation and sorting. This was to reduce auto-fluorescence effects caused by LB medium and also prevent other CCR mechanisms from being activated due to the diverse array of carbon sources present in LB. Random mutagenesis was performed on the S391R mutant of *exuT* and an error rate of

approximately 2.2 bp/kb was achieved. The corresponding strain library was incubated in MOPS minimal medium supplemented with 10 g/l of glucose and galacturonate overnight. Analysis by flow cytometry was performed on a freezer stock of the first library (Figure 3.17(A)) and the second library (Figure 3.17(B)). The GFP distribution profile for the first library has a larger peak in the negative regime in comparison to the second library. However, the first library also appears to have a larger peak towards the positive GFP regime. Generally, the population of the second library has shifted towards higher GFP signal in comparison to the first library, which is expected as a top performing mutant was chosen as the initial seed for this round of mutations. The top 0.5% GFP positive population of Figure 3.17(B) was collected for strain isolation and further analysis to determine if beneficial mutations were generated.

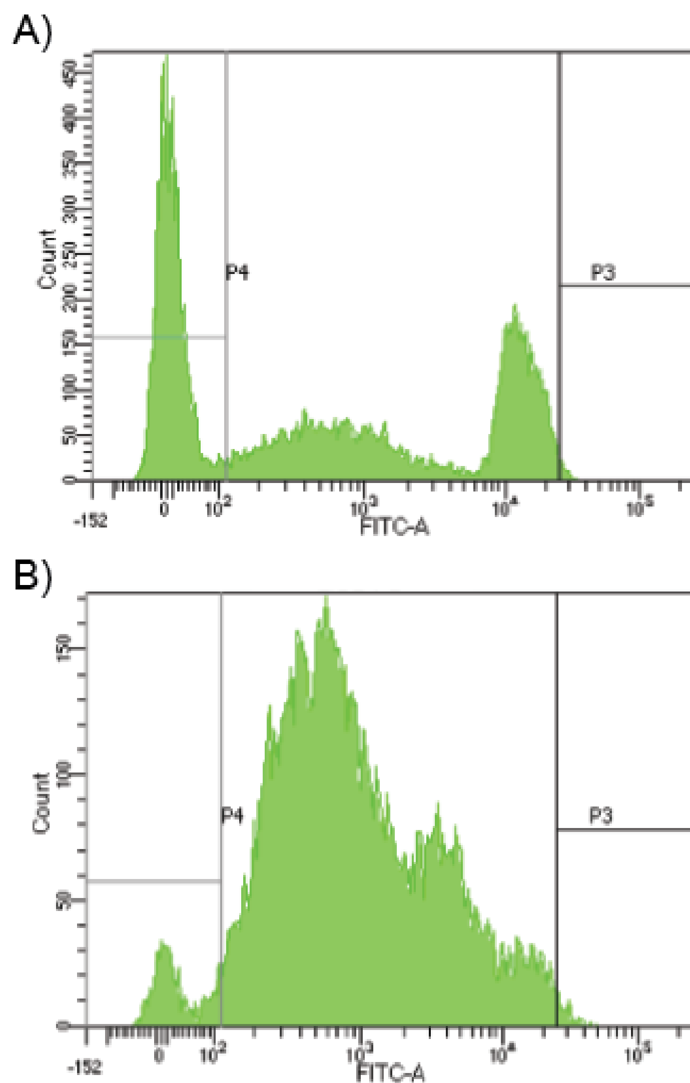

Figure 3-17. GFP distribution profiles for second ExuT mutant library sorting procedure, detected using flow cytometry. For this round of sorting, cultures were incubated in MOPS medium supplemented with 10 g/l of glucose and galacturonate to eliminate any fluorescence that could be caused by a rich medium. The (A) first library was re-analyzed to compare to the (B) second library. The top 0.5% GFP positive population of the second library was collected for further analysis.

As was done with the first library, isolated clones were assessed for their efficacy using a BioLector to measure GFP signal (Figure 3.18). Multiple clones were found to exhibit a GFP/Biomass ratio higher than that of the positive control, a phenomenon that was much rarer in the analysis of the first library. This again makes sense as this second round of mutations is an

attempt to enrich the performance seen by the first library. The mutants that performed better than the positive control were collected and sequenced to determine the mutations present in the constitutively expressed *exuT* gene.

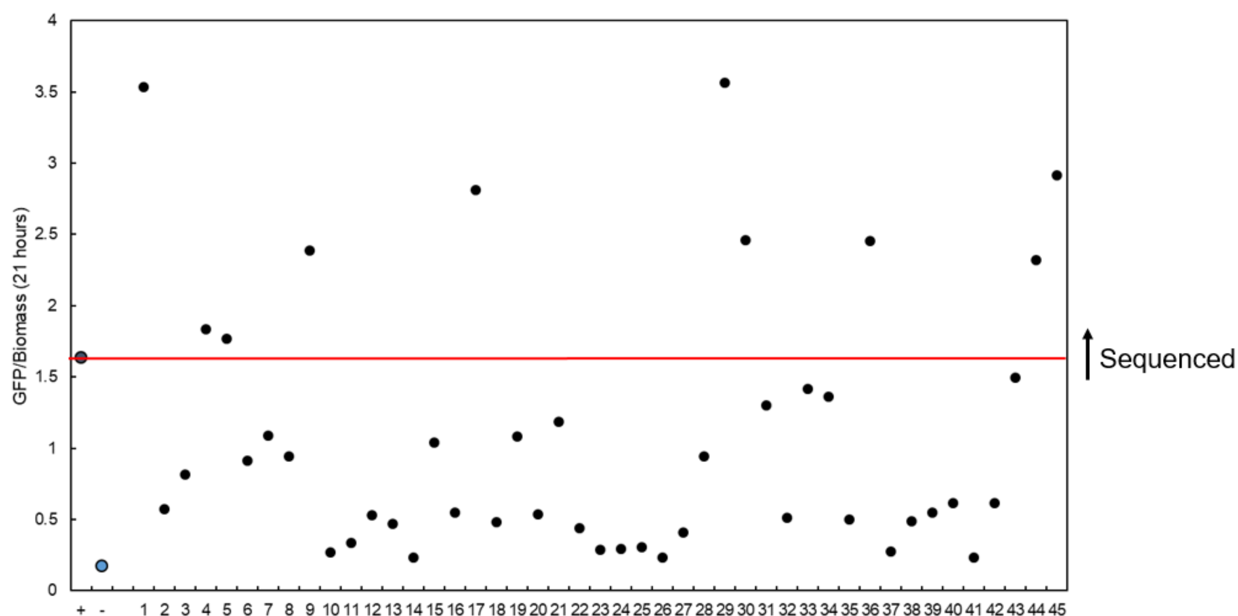

Figure 3-18. GFP analysis of isolated clones done in a BioLector from the sorting experiment of the second mutant ExuT library. Samples were cultured in MOPS medium supplemented with 10 g/l of glucose and galacturonate. The positive control is a sample without glucose fed and the negative control is expressing wildtype *exuT*. Any isolated colony that displayed activity above the red line shown was sequenced to determine the mutations present in the *exuT* gene.

A wider diversity of mutations was discovered in this round of screening in comparison to the first library. The mutants discovered were: S391R+F322Y, S391R+T245I, S391R+T375S, and S391R+E131K+F253Y. Additionally, one of the mutants isolated was the original S391R *exuT* gene. MG1655(DE3)Δ*garK*Δ*glxK*Δ*uxaC*Δ*hyi*Δ*exuT* was transformed with a plasmid constitutively expressing each of these mutants as well as with two plasmids that allowed for the overexpression of the entire D-glyceric acid pathway. The titer of D-glyceric acid observed from a co-feed of glucose and galacturonate was measured over a 72 hour period for each of these strains (Figure 3.19). Surprisingly, 3 of the newly identified ExuT mutants greatly reduced the D-glyceric

acid titer. Additionally, the S391R+T245I mutant performed similarly to the S391R mutant, suggesting that the addition of the T245I mutation was not beneficial. Therefore, the second round of mutagenesis and screening in the directed evolution procedure did not produce a mutant protein that performed better than the optimum from the first library. This could be due to the S391R mutation being a dead end in terms of evolutionary development, or the change in medium from rich to minimal could have affected these results and prevented the elucidation of better mutants.

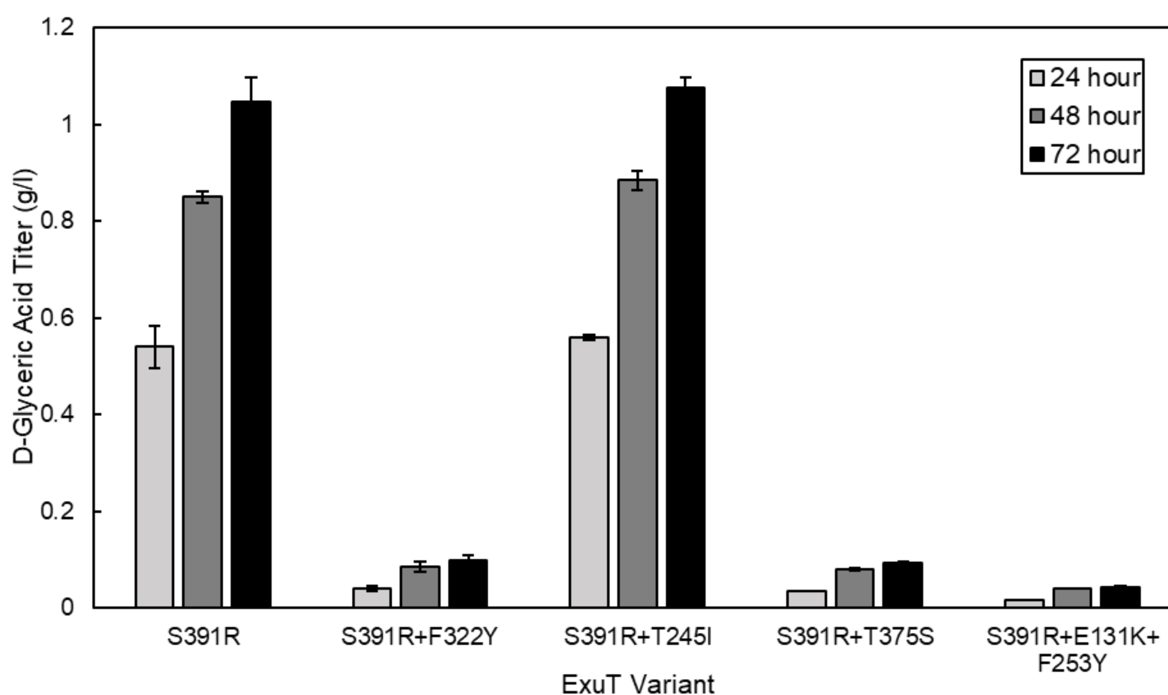

Figure 3-19. D-glyceric acid titer from the plasmid-based expression of different ExuT mutants isolated from the second mutant library with a feed of glucose and galacturonate. Fermentation was performed in LB medium supplemented with 10 g/l of glucose, 10 g/l of galacturonate and 0.1 mM IPTG. Cultures were done in tubes with a 10 ml working volume, incubated at 30 °C and agitated at a shaking rate of 250 RPM.

3.4 Conclusions

The co-utilization of carbon sources commonly present in renewable agricultural wastes is imperative for the optimization of bioprocessing for the production of value-added chemicals. Central to this overarching goal is the ability to co-utilize sugar mixtures containing glucose, the primary trigger for CCR effects in *E. coli*. Two approaches for CCR relaxation were analyzed, one being a global alleviation strategy and the other targeting a specific sugar for co-utilization. Global alleviation through the mutation of $EIIA^{glc}$, while allowing co-utilization of sugar mixtures, led to phenotypic inconsistencies that are commonly observed in the literature when parts of the PTS are mutated. Mutation of the galacturonate permease ExuT led to improved production of D-glyceric acid when glucose and galacturonate were co-fed. This mutant was obtained by combining directed evolution with a novel sugar transport screening system that utilized a galacturonate specific biosensor to detect intracellular galacturonate level. However, the study also displayed that removing transcriptional inhibition of D-glyceric acid producing genes led to a larger increase in titer than the mutation of the permease did. This work displays both the complexity and potential present in solutions for the relaxation of CCR effects for the co-utilization of sugar mixtures through the engineering of intrinsic sugar catabolism pathways.

4. Market Opportunity of Synthetic Biology in the Specialty Chemicals Industry

Abstract

Innovation in the chemicals industry is necessitated by increasing calls for more sustainable and higher-performing compounds and processes. While the market for specialty chemicals remains large and attractive, market forces will likely decrease profitability in the future. Biological production, enhanced with synthetic biology, presents an opportunity to disrupt this industry with a differentiated offering, combining sustainable feedstock usage with novel product portfolios. In this analysis the traditional specialty chemicals industry was investigated in terms of market attractiveness and the effect of Porter's five forces on the market future. The capabilities of synthetic biology to alter these effects on new entrants to the market were determined and displayed a limiting of negative future effects on these firms. Therefore, it is anticipated that the specialty chemicals industry will be more attractive for synthetic biology companies than traditional incumbents in the future. Current challenges for synthetic biology companies, including ethical and regulatory uncertainty, difficulty in scale-up, and under-developed logistic networks, are also presented. Finally, a competitive landscape of current industrial players in synthetic biology and an assessment of firm strategy is presented.

4.1 Introduction

The specialty chemical industry has blossomed through continued innovation in the development of novel chemical applications. This has led to more sophisticated chemical processes to produce these specialty chemicals as well as a further global reliance on new products. Generally, the chemical industry can be split into two classifications: commodity and specialty chemicals. Commodity chemicals are differentiated by their wide-scale use, for example fuels and organic building blocks for other chemicals, while specialty chemicals are produced in smaller batches and are often optimized for a specific use, for example fragrances (Figure 4-1) [148]. While advances have been made in the commodity chemical industry, especially in the biological production of biofuels [149, 150] and ethanol [151, 152], this analysis will focus on the specialty chemical industry and analyze how synthetic biology could further revolutionize and disrupt the field.

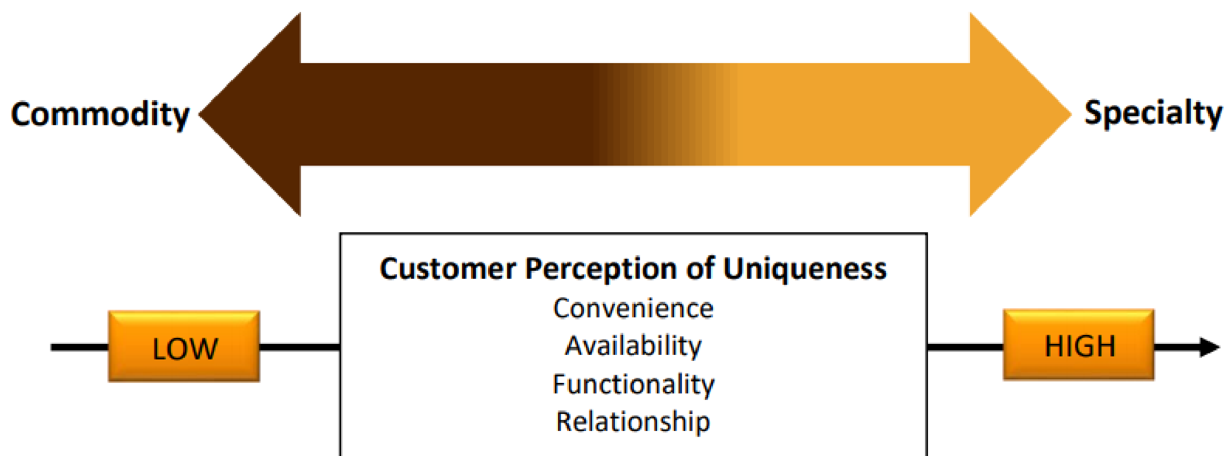

Figure 4-1. Graphical representation of the differentiating factors between commodity and specialty chemicals. Figure taken from Hatzikyriakou, M (2019) [148].

Global spending for specialty chemicals was estimated as $873 billion in 2019 with a 3% compound annual growth rate (CAGR) expected through 2024 [153]. Of the total chemical

industry, specialty chemicals represented 26% of the market in 2019 which was segmented among multiple product classes including textile chemicals, coatings, food products, and more (Figure 4-2) [153]. The highly diversified nature of the specialty chemical industry enables opportunity for new players to enter the industry due to the possibility of identifying a distinctive niche which would enable stable growth and security in market share. This can be seen in the increased prevalence of production methods using both electrochemical [154] and biological [155] means, among others. The market is eager for alternatives in production as issues involving environmental degradation, sustainability, and future economics become concerns with our increasing use of these products [156]. Biotechnology has garnered interest due to the capability of synthetic biology to modify organisms and generate novel chemicals in a platform-based manner [157] and its utilization of renewable feedstock [158] which improves the sustainability of the process. The massive influx of innovation in the area of synthetic biology has created a large technology push effect in the market [159]. This coalesces in technology being introduced into the market with the goal of trying to find buyers that did not know they needed/wanted the product. Another effect of this is that many small-scale companies enter the industry with their innovation, hoping to generate a pull effect in the future through recognition by bigger industry players. An example of this would be a large-scale specialty chemicals company, like Exxon Mobil, acquiring or investing in a start-up's technology due to their knowledge of the market for it. While effects like this are occurring in the industry [160], the slow speed of adoption is surprising to many and is likely due to multiple challenges in using synthetic biology for the biological production of chemicals including limited scalability [161] and ethical concerns [162].

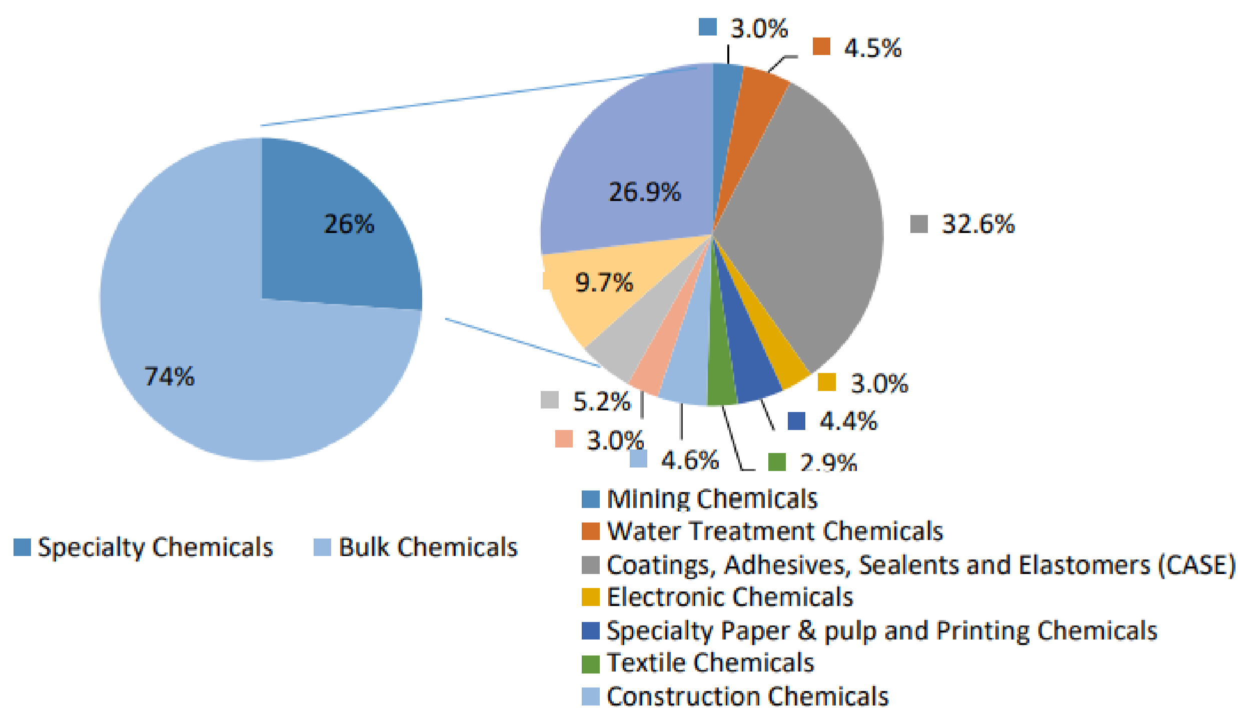

Figure 4-2. Make-up of the specialty chemical industry in 2019 and its positioning in relation to the global chemical industry. Figure taken from Bala, S (2020) [153].

This analysis will first overview the specialty chemical industry from the lens of global market attractiveness and dynamics. While giving a sense of the opportunities and stagnation in the market, this analysis will also illuminate gaps that are present that can be filled through synthetic biology applications. Opportunities for synthetic biology to disrupt the market are presented along with a competitive landscape of current organizations that are beginning to accomplish this.

4.2 Market Attractiveness of the Specialty Chemicals Industry

The attractiveness of an industry is a measurement of how enticing a specific market is to entry by new competitors. Many frameworks have been developed to better understand and define an industry's attractiveness to inform market entry and strategy [163-165], however in this analysis we will use three characteristics to define it: profitability and volatility, industry size and growth, and total addressable market. The specialty chemicals industry is reasonably attractive based on average profitability, low volatility, large market size, low market growth, and a large total addressable market (Table 4-1). Primary risks associated with the industry include the risk of commoditization, which would lower profitability, and high exposure to trade disputes that could limit access to developing economies that will be integral to the growth of the industry.

Table 4-1. Summary of analysis of the attractiveness of the specialty chemicals industry.

Factor	Attractiveness	Rationale
Profitability Volatility	Medium High	• Profitability is in line with US average compared across industries • Volatility is low across 2001-2020, with general upward trend • There is a risk of commoditization which would lead to decreased profitability and increased volatility
Industry Size Industry Growth	High Low/Medium	• Industry size is large (near $1T USD) • Growth is lower than global total market • Slower growth could be inconsequential if new entrant has novel product
Total Addressable Market	High	• Total addressable market is large due to high innovation • New developing nations enable future markets • There is a risk of loss of access to developing areas if trade disputes arise

4.2.1 Profitability and Volatility

The profitability of an industry is an important metric for attractiveness as it suggests what the return of a venture into a specific market may be. Return on invested capital (ROIC) is often used as a proxy for this evaluation. In his landmark paper, Porter curated ROIC measurements in multiple US industries from 1992-2006, finding that the average value was 14.9% [166]. This data set illuminated many trends in industries that are very competitive due to the commoditization of their product, such as the airline industry with an average ROIC of 5.9%, or less price competitive due to the uniqueness of each offering, such as the pharmaceutical industry with an ROIC of 31.7% [166].

Historically, the top specialty chemical conglomerate companies have achieved an ROIC between 10-20% (2001-2020) [167, 168]. This is similar to the US average ROIC, suggesting that the industry could be attractive. Commodity chemicals in a similar time frame (2001-2016) displayed an industry ROIC of 6-14%, always lagging behind specialty chemicals by approximately 4% [167]. Specialty chemical organizations enjoy a higher profitability than commodity chemicals due to the differences displayed in Figure 4-1, leading to a higher willingness to pay for their product. Stratification of the specialty chemical industry into pure players, only producing a single product, and conglomerates, producing a portfolio of products, shows an 11% boost in ROIC for pure players [168]. This differentiation is likely due to the highly specialized nature of the pure players, either requiring highly complex manufacturing or possessing a technical advantage that enables them to act as a monopoly in the space.

The volatility of an industry's profitability gives an additional sense of the risk involved with entering a market since a highly volatile industry could lead to negligible growth opportunities, depending on the time of entry. An industry with no volatility is preferred due to its predictability,

enabling smart and well-informed decision making. The specialty chemicals industry has relatively low volatility, showing a gradual increase in ROIC over the period of 2001-2020 [168]. In comparison the commodity chemical industry is highly volatile, displaying an oscillating pattern of ROIC with an amplitude of approximately 3% [167]. This difference is likely due to the increased competition and dependence on raw materials pricing in the commodity chemical industry in comparison to specialty chemicals.

However, there is always a threat in specialty industries of commoditization which could be initiated through increased demand, technological innovation, or the expiration of patent protection. Upon commoditization, firms must shift their strategy to become more lean and must now compete on price rather than product [169]. This risk of commoditization has a high effect on the attractiveness of the specialty chemical industry since the commodity chemical industry is highly volatile and has a lower-than-average profitability. Pure players are especially wary of commoditization as they may be deposed if their single product is made unprofitable without a portfolio of other unique offerings to support the firm.

The specialty chemicals industry has a moderate attractiveness in terms of profitability and volatility. While historically its ROIC has been increasing and relatively high, there is a constant risk that products would become commoditized. As innovation increases in the chemicals industry, this risk increases and forces a new player to generate a moat to prevent commoditization or prepare for its inevitability.

4.2.2 Industry Size and Growth

The global market for the specialty chemicals industry was estimated to be $873.4 billion USD in 2019 with a CAGR of 3% from 2019-2024 [153]. The massive size of the industry is very attractive for entrance, as taking even a small market share would be impactful. The estimated growth rate

is below the estimated growth rate of the world economy, estimated to be 3.5% in 2019 and 3.6% in 2020 by IMF [170]. This suggests that the market growth in the specialty industry lags the total global market which is likely due to the commoditization of parts of the industry and new products making up a small amount of the market initially. Additionally, specialty chemical users could be slow to adapt to new products, even with the fast innovation in the industry currently.

The size of the specialty chemical industry is attractive. However, the industry's growth lags the global market. Therefore, based on industry size and growth the specialty chemical industry is moderately attractive. The slower growth rate could be inconsequential, however, if a player enters the market with a new method of production or product that makes another obsolete. This would remove market share from another player while quickly capturing the same piece with the superior technology. Therefore, if a player is entering the market with a market-disrupting technology, the specialty chemical industry would be very attractive due to its large market size.

4.2.3 Total Addressable Market

The total addressable market (TAM) is a measurement of the total market size that is possible for an industry. A larger TAM means that the growth of the market can be sustained for a longer amount of time, enabling capture of part of the market by a new entrant, making the industry more attractive. Alternatively, if the TAM is close to the current market size, the industry is not attractive as it would be necessary to push a current player out of the space to capture any part of the market.

For the specialty chemicals industry, it is difficult to quantify the TAM since there is constant innovation in the space, enabling new product opportunities. This bodes well for a new entrant with a novel product, as there will likely be a market share available to take or the market will expand to encompass the product. Specifically of interest is the expected rapid growth of the market size for specialty chemicals in the Asia-Pacific region (3.3% CAGR) which already

represented 47.9% of the market share in 2019 [153]. Through the rapid expansion of countries in this region, the demand for specialty chemicals will continue to grow into the future. It is also expected that the same effect will eventually be felt in other low GDP countries as they improve infrastructure, trade for new goods, and participate in other specialty chemical-fueled areas.

The global total addressable market for the specialty chemicals industry is attractive due to the further development of many parts of the world. Additionally, innovation constantly opens new market share for novel products. Risks associated with this market growth include trade disputes which could hamper ability to reach developing markets and cannibalism of market size through the introduction of superior products.

4.3 Market Dynamics of the Specialty Chemicals Industry

While it is essential to understand the current state of attractiveness of the market, it is even more important to estimate how that attractiveness will evolve in the future. A common framework used for this type of analysis are Porter's five forces: buyer power, supplier power, industry rivalry, substitutes, and new entrants [166]. Based on this analysis, it is expected that the specialty chemicals industry will become less attractive in the future due to increased pressure both from buyers and suppliers as well as heightened competition through industry rivalries (Table 4-2). Increasing commoditization lowers profitability and increases pressure on firms to differentiate which necessitates investment in new products and services.

Table 4-2. Summary of analysis of Porter's five forces on the specialty chemicals industry.

Force	Effect on Industry	Rationale
Buyer Power	High	• Rapid speed of commoditization increasing supplier options for buyers • Increasing buyer demand for sustainable products and flexible relationships
Supplier Power	High	• Limited optimization of supply chain management leaving profit on the table • Volatility in raw materials price and buyer demands lead to increased supplier power
Industry Rivalry	High	• Rivalry based on price for many products due to commoditization • Necessitates firms invest in differentiating technology
Substitutes	Medium	• Substitutes being required by buyers to limit current hazards • Many current products are optimized, difficult to find worthwhile replacements
New Entrants	Low	• Large investment in R&D and capital needed for entry into market • Niche markets may allow new entrants, but would not affect market share of incumbents

4.3.1 Buyer Power

In a capitalist economy, the market determines which products and companies stay in business and which do not. Therefore, the power of the buyer and how it is expected to evolve over time is incredibly important to understand the dynamics of an industry. An industry where the buyer has negotiation power, such as airlines, is much riskier to enter in comparison to industries where customers are price takers and therefore have low price elasticity, such as gasoline stations.

Buyer power in the specialty chemicals industry is high. This is especially driven by the rapidly increasing speed of commoditization of specialty chemical products. Upon commoditization, buyers enjoy multiple possible supplier options and can play them off one another to gain benefit. Additionally, the sustainability and flexibility of production has been increasingly prioritized by buyers [171, 172]. This forces specialty chemicals companies to invest more capital into these initiatives to improve their marketability, lowering their profitability. Finally, due to the high innovation in the industry buyers can force organizations to compete with one another on both price and function. It is anticipated that these demands will increase in the future and would necessitate large investment into R&D to develop novel production methods and chemicals that both perform at a high level and are sustainable.

4.3.2 Supplier Power

To make a product an organization needs to source raw materials for them to process. Therefore, the relation to suppliers and the power that the suppliers have over the industry is an important factor in the dynamics of the attractiveness of the industry. If there are, for example, more suppliers to choose from, then a bidding war could be initiated which would lead to higher value generation by the industry of interest. Supplier power becomes problematic when suppliers take a large or volatile share of the value from the industry.

The specialty chemicals industry has been unable to perfect its sourcing platform, leading to moderately high supplier power. The value of purchasing expenditures can range from 20-60% which leads to uncertainty in the industry as well as giving suppliers the upper hand in negotiations [173]. These expenditures vary due to both the volatility in raw materials pricing as well as the additional demands by consumers that the industry be flexible in terms of product output. Pressure in the area of procurement has led many in the specialty chemicals industry to rethink its purchasing model which has been accompanied by further research into optimum raw materials management [174]. Further optimization of these techniques should lead to a lower supplier power in the future of the industry; however, suppliers will still be able to pressure margins.

4.3.3 Industry Rivalry

Competition within an industry is detrimental to profitability due to the prevalence of price cutting to acquire market share. While competition is healthy for industry, an excess of it will lead to a less attractive industry in the long run and could prevent new entrants from taking hold. Rivalries are especially problematic in commodity industries where all firms are attempting to sell the customer a similar product.

In specialty chemicals, industry rivalries are high due to the increased speed of commoditization which leads to competition on price. To limit competition, firms are forced to focus a large amount of their budget on R&D to continue to develop new products and maintain differentiation from their competitors. Additionally, business models incorporating flexibility in their offerings are utilized by some firms to gain market share at the expense of the additional cost of this service [175]. Companies that can leverage differentiating value statements, such as a more sustainable product/process or IP that prevents commoditization, are also able to lessen the effect of industry rivalries. Therefore, a new company entering this market should have a plan to display

differentiation to the market and compete with the capabilities of current players [176]. It is anticipated that this competition pressure will remain in the future as new developing markets open and multiple organizations move to gain maximum market share.

4.3.4 Substitutes

A substitute product presents the possibility of a firm's offering becoming irrelevant in a market due to superior performance or lower pricing. Additionally, a substitute could split a market leading to a lower share held by the incumbent firm. Both situations present great risk which can be mitigated through constant improvement of a firm's offerings, IP protection, and an understanding of the competitive landscape to understand substitutes before they hit the market.

Substitutes present a moderate risk in the specialty chemicals industry. Often substitutes are requested by buyers as some currently used specialty chemicals are dangerous or pose an environmental hazard [177]. Additionally, novel chemicals and processes are constantly being developed and present the possibility of outperforming a current product on the marketplace. However, it is resource-intensive to create these new products, and there is no guarantee that they will perform better than current offerings. Buyers also are not incentivized to break current supply relationships unless the cost or performance is significantly improved, so not all substitutes will pull a large market share when introduced.

4.3.5 New Entrants

The introduction of new competition to a market is never welcomed by incumbents since it likely means that they must lose some of their market share. If it is easy for new companies to join the market and gain a foothold, this leads to greater competition and a less attractive industry.

The specialty chemicals industry has a low risk to new entrants taking market share. This is primarily due to the large amount of capital investment needed to both advance R&D far enough to have a marketable product and to purchase equipment to produce the product. There is a slight risk to new entrants in niche markets that necessitate highly specialized products, however this is not problematic to incumbents because these represent small portions of the market, and they are not likely to lose any of their market share to one of these new entrants.

4.4 Opportunity for Synthetic Biology to Disrupt the Specialty Chemicals Industry

Biotechnology and its optimization through synthetic biology presents an opportunity to capitalize on the shortcomings of traditional specialty chemicals synthesis. Entrants would be able to take advantage of an attractive industry and challenge incumbents with differentiation that would enable them to capture market share. These differentiations include but are not limited to innovative feedstock usage and new product capabilities. Due to these capabilities, the market for non-energetic biorefined products is expected to increase at a CAGR of 9.9% from 2020-2025 (Figure 4.3) [178]. Different classes of bio-derived specialty chemicals are anticipated to increase at a CAGR of between 8.6-17.7% during this period, which is much higher than the estimation for the current specialty chemicals industry [153, 178]. It is also of note that the primary growth is predicted to be in non-energetic bioproducts, which mirrors current results seen in industry where pivots away from biofuels and towards specialty products are occurring for many synthetic biology companies [179]. This is also shown through an analysis of the patent landscape which displays an increase in both patents for production of specialty chemicals over time as well as an increase in patents for the optimization of currently reported chemical production patents [180].

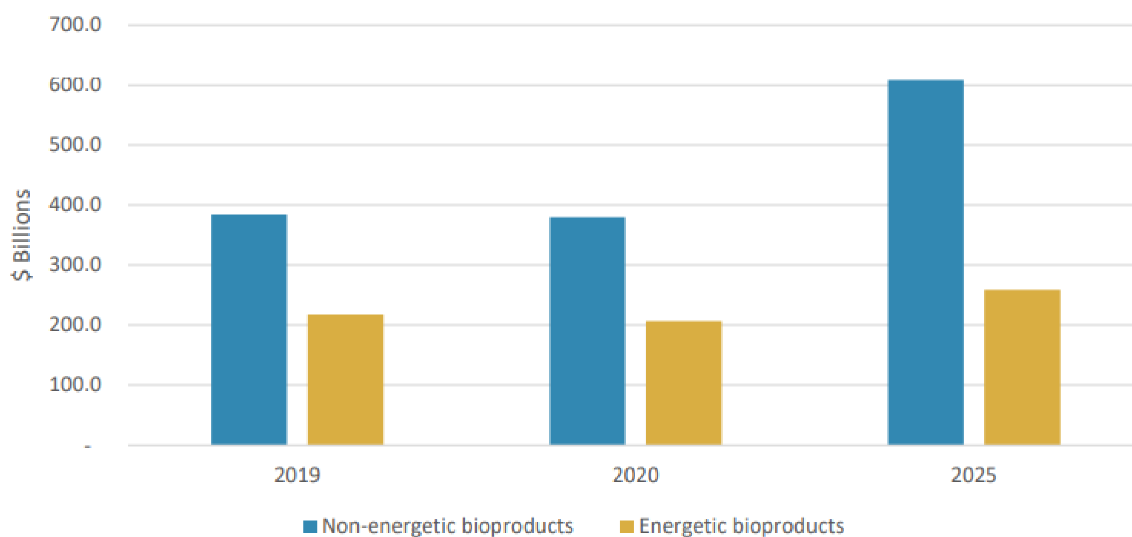

Figure 4-3. Global market for biorefinery products by bioproduct category, 2019-2025. Figure taken from "Biorefinery Products: Global Markets" by BCC Research 2021 [178].

The key capabilities that synthetic biology enables leads to an altering of the attractiveness of the industry and the forces that will dictate attractiveness in the future. A decrease of the pressure from buyer and supplier power as well as a lessening of industry rivalries due to differentiation of products enables a more attractive industry as well as a more positive industry outlook in the future for new entrants to the specialty chemicals industry utilizing synthetic biology.

However, as the bioproduction industry takes more of the market share from traditional specialty chemicals producers, it must continue to innovate and overcome visible challenges in front of it. Ethical and regulatory issues involving the public perception of genetic engineering, manufacturing and scale-up difficulties associated with a lack of infrastructure and low yields, and a fragmented supply chain for these renewable feedstocks all present hurdles that must be optimized before accelerated market takeover can occur.

4.4.1 Benefits of Synthetic Biology Application

4.4.1.1 Novel Feedstock Usage for Increased Sustainability

The ability of microbial organisms to utilize renewable feedstock to produce chemicals of interest has been discussed in Chapters 2 and 3 of this thesis. An effective combination of genetic engineering and biomass procurement can lead to a sustainable circular process chain in which an effectively unlimited amount of feedstock for necessary chemicals is available for use [181]. Through the power of metabolic engineering, many chemicals of industrial significance have been produced from a variety of feedstocks, including corn, sugar cane, lignocellulosic biomass and many more [182]. Further development is still being done on optimizing the pre-processing of these feedstocks to enable efficient usage of the material as well as utilizing C1 carbon sources which are abundant, cheap, and often damaging if not correctly processed [183].

The ability to process a separate feedstock than competitors in an industry increases the attractiveness of the specialty chemicals industry for synthetic biology firms. The utilization of these starting materials decreases the dependence of the firm on fossil fuels, which raise concerns both due to their limited nature and the political climate that surrounds their supply. Therefore, the expected volatility of the price of feedstock is expected to decrease upon the development of a dependable and sustainable supply chain, leading to a higher and more predictable profitability for synthetic biology firms.

The buyer and supplier powers that will dictate the attractiveness of the industry in the future will also affect synthetic biology firms differently than traditional producers. As discussed in Section 4.3.1, buyers have been applying pressure to specialty chemicals producers to synthesize products in a more sustainable fashion. Utilizing biomass-derived feedstocks will enable synthetic biology firms to hold a competitive advantage over much of the industry which likely will lead to an

increased number of contracts. Supplier power that is fueled by consistent usage of common starting feedstocks would also be lowered for synthetic biology firms. Many of the renewable feedstocks being used for bioproduction are waste materials from agriculture activity. This allows synthetic biology firms to be price-makers instead of price-takers. Additionally, until bioproduction becomes more prevalent, the competition for acquisition of these materials will be low, further decreasing the supplier power. The primary difficulty in this area for synthetic biology firms is the development of an effective supply chain for these feedstocks which will be discussed further in Section 4.4.2.

4.4.1.2 Novel Product Capability

Chapter 2 of this thesis presented the biological production of a high-value chemical from a low-value feedstock. This demonstrates another benefit of utilizing synthetic biology for the biological production of specialty chemicals: the ability to produce a diverse number of products in a predictable and selective manner. In fact, the application of synthetic biology allows for not just the production of known high-value chemicals [184], but also for “new-to-Nature” products, such as fluorine-containing monomers [185] or innovative bioplastics [186] to name a few. Production is enabled by the ability of biological systems to express exogenous genes, be predictable for engineering of metabolic pathways, and produce stereo-selectively. In particular, the fragrance and flavor industry has seen the impact of these technologies [187] leading to the development of grapefruit flavoring (nootkatone[188]), vanilla (vanillin[189]), and other products that enable continuity in supply indeterminate of harvest yield complications. These novel products enable new functionality in a multitude of areas, and the potential appears limitless as the ability to simulate, engineer, and screen for optimized mutants is advanced and continuously developing [190, 191]. The promise of new high-quality chemicals has led many groups, including in the

United States [192] and Asia [193], to develop investment plans for future development and industrialization of synthetic biology.

The attractiveness of the specialty chemicals industry is increased by the ability to produce novel chemicals that are inaccessible by industry incumbents. Novel products that offer new functionality product new markets, which increase both the size of the industry and the TAM. Additionally, a firm can demand a higher profitability for a product that replaces a current offering but performs better.

The buyer power, industry rivalry, and substitutes forces that highly impact the specialty chemicals industry affect synthetic biology firms differently due to the ability to synthesize novel products. Buyer power is decreased due to the increased range and specificity of products that are available through bioproduction. While the number of products that are available through bioproduction is still not equal to that of chemical synthesis, the ability to produce multiple high-value targets that are not practical through chemical synthesis will lead to an increase in industry attractiveness for synthetic biology firms over time. For specialty products that are difficult to synthesize chemically, industry rivalry and substitute forces will also decrease for bio producers. The technology of these firms will generate a moat for a market share that will not be accessible for incumbents. This is a primary reason why the initial opportunity for the industrialization of synthetic biology is in specialty and not commodity chemicals. Industry rivalries would be fierce for synthetic biology entrants competing in commodity chemicals, likely causing them to exit unless they possess a large advantage in production methodology or raw materials sourcing.

4.4.2 Challenges Still Needing to be Addressed for Synthetic Biology

4.4.2.1 Ethical and Regulatory Issues

Many new innovations that intend to disrupt current markets are subject to increased scrutiny, and rightly so, for their ethics and the regulatory aspects that should control it. Synthetic biology and genetic engineering carries with it a large amount of scrutiny from the public who both do not understand it completely and are afraid of the possibilities that may come from it [194]. Therefore, there is a push to increase public and institutional understanding of the ramifications of synthetic biology more completely as well as plan for ways to mitigate possible disasters that could occur due to this innovation [195]. Additionally, ethical considerations of access, land usage, and the limits of creating life also necessitate complete thought before effective regulation should be introduced [162]. While addressing these risks is incredibly important for the safe usage of this powerful technology, there is a risk that the regulation that is introduced could be unnecessarily restrictive. Although it is not currently apparent that this will happen [196, 197], mistakes in the community or new-found technology may lead to increased regulation that would strain the development of the industry [198].

Increased regulatory action would lead to a decrease in the ability of synthetic biology firms to compete with traditional specialty chemicals manufacturers. It could decrease profitability depending on the action that firms would have to take to abide by regulations. Buyer power would be increased as there would be less of a reason to choose synthetic biology firms over traditional incumbents. Additionally, the threat of substitutes would be increased for synthetic biology firms due to competition from traditional firms that would be able to operate without regulations specific to biological production.

4.4.2.2 Manufacturing and Scale-up Difficulties

A consistent challenge for the synthetic biology industry is the ability to produce compounds at scale to compete with incumbents. This is a much more prevalent problem in commodity chemicals, where in 2007 it was estimated that capital and materials costs would need to be halved in order to compete with the industry at large [199]. For specialty chemicals, the capital cost per capacity is still higher than that of incumbents [200], however higher profit margins enable more competition by synthetic biology firms. Impressive improvements in cost optimization have been achieved both in commodity and specialty chemicals over the past few decades [200], but competition with incumbent producers with large infrastructure networks will also add risk to market entry. Additionally, low yields have often led to difficulties in scaling processes up to commercial production levels. Improvements in the design-build-test cycle have led to increases in throughput of testing [201], however the difficulty of scaling biological processes remains a source of risk where it may be difficult to obtain useful yields once production is moved from the benchtop.

High capital costs and inefficient scale-up would decrease the attractiveness of the specialty chemicals industry for synthetic biology firms due to a decrease in profitability. Additionally, industry rivalry would be increased as a synthetic biology firm may find itself unable to compete in a specific market due to being unable to compete on price with incumbents. Pressure on increasing yields of processes may lead to additional investment into R&D, decreasing profit margins further.

4.4.2.3 Supply Chain Challenges

An essential capability that has been identified to be holding back the proliferation of bioproduction is the development of an efficient and sustainable supply chain for raw material

feedstock [202]. The renewable materials that enable a key capability of synthetic biology are often located in diverse parts of the world and are non-centralized. This presents difficulties in supplying a consistent flow of feedstock, a challenge that is magnified when growing seasons are also considered [203]. Solutions such as centralized treatment plants and storage facilities have been suggested [202] and analysis has even been done to supply these feedstocks to remote areas as far away as space [204]. However, until an effective logistics platform is put into action for the supply of feedstock to synthetic biology firms there will be additional risk for these new entrants into the specialty chemicals industry.

Supply chain inefficiencies would decrease the attractiveness of the industry due to a lowering of profitability from an increase in raw materials cost. Additionally, investment into the development of a supply chain would initially decrease the profitability of new entrants. Supplier power would be increased from the scarcity of starting materials and industry rivalry would be increased as incumbents could compete on price and likely force synthetic biology firms out of the market.

4.5 Competitive Landscape of Synthetic Biology Firms

The development and initiation of production of goods using biological engineering has been touted as one of the major contributors to the next industrial revolution [205]. Synthetic biology plays an important role in this revolution and development of synthetic biology-leveraging organizations has begun in earnest towards this transition. The pioneering companies that are a part of the current competitive landscape of specialty chemicals producers using synthetic biology are helping to define the future bioeconomy [206]. Additionally, there has been increasing support from government-sponsored research organizations across the world for the development of these capabilities [207]. The understanding of the strategies of current players in this field will enable better planning for future organizations as well as hopeful avoidance of past missteps.

An analysis of current players utilizing synthetic biology as their primary method of production in the specialty chemicals industry was performed to understand the maturity of current companies and their respective strategies. A list of selected synthetic biology companies was developed (Table 4-3) using Pitchbook [208] and their primary product offerings were identified. Not included in this collection are large companies such as Dupont that are investing in the technology as this was intended as an analysis of novel players. Additionally, companies that are investigating food replacement technology were not included in this list as it is considered its own industry apart from specialty chemicals.

Table 4-3. Selection of synthetic biology companies in the specialty chemicals industry.

Name	Year Founded	HQ	Products
Genomatica	1998	San Diego, CA	1,4-butanediol, 1,3-butylene glycol, bio-Nylon, Long-chain chemicals
Cathay Industrial Biotech	2000	Shanghai, China	Long chain dibasic acids, pentylenediamine, polyamide
Amicogen	2000	Jinju-si, South Korea	Multiple products
Amyris	2003	Emeryville, CA	Multiple high value chemicals
Evolva	2004	Reinach, Switzerland	Nutrition, healthcare, and wellness products
Ginkgo Bioworks	2008	Boston, MA	Platform generated organisms for application in many areas
Bioworks	2009	Edmonton, Canada	Bioplastics
Conagen	2010	Bedford, MA	Amino acids, lipids, flavonoids, phenolics, terpenoids, and cartenoids
Lygos	2010	Berkeley, CA	Aspartic, Glyceric, Isobutyric, Lactic, and Molonic acids
Kalion	2011	Milton, MA	Glucaric Acid
Manus Bio	2011	Cambridge, MA	Sweeteners, flavors, and others
Biosyntia	2012	Copenhagen, Denmark	Biotin
Zymergen	2013	Emeryville, CA	Products in electronics, consumer care, and agriculture
Biotechnologies	2014	Boulder, CO	Amino acids, organic acids, terpenoids
Hyasynth Bio	2014	Montreal, Canada	Cannabinoids
Ardra	2014	Toronto, Canada	Leaf-aldehyde, Butylene-glycol, Keto-esters
Magenta BioLabs	2016	Cork, Ireland	Hyaluronic acid
Joyn Bio	2017	Boston, MA	Microbial products for agriculture
Demetrix	2017	Berkeley, CA	Cannabinoids
Huue	2019	Oakland, CA	Indigo
Cysbio	2019	Kongens Lyngby, Denmark	Amino acids, sulfated biochemicals

This list gives important insight into the current strategy of many synthetic biology firms. Smaller companies such as Huue, Demetrix, and Kalion have chosen single, high value, compound offerings. This enables quicker optimization of production and therefore a faster realization of revenue and value. However, this strategy also brings with itself the risk of putting the entire fate of the organization into advancements in and the market of a single product. Therefore, product choice is integral and selections of indigo and cannabinoids, both highly demanded and valuable, de-escalates the risk.

In comparison, larger companies such as Amyris have transitioned into a suite of product offerings as they have de-risked their platform through previous product success. This success was not always apparent for Amyris, however, as they initially sought to compete in the biofuels market, but then opted to focus on specialty chemicals instead [209]. As discussed in previous sections, the ability for synthetic biology to compete in the commodity chemicals industry is difficult. Multiple biofuel and commodity chemicals companies that were initially highly valued based on their intellectual assets went bust due to an inability to generate competitive products [210]. Amyris has displayed both sides of this equation and shows the possible success that comes with a focus on a differentiating platform.

There is a large amount of momentum for synthetic biology companies with two notable companies, Zymergen and Ginkgo Bioworks, performing or planning an initial public offering [211, 212]. These companies' primary businesses are the production of novel materials biologically and un-locking the potential of nature. Additionally, in the United States investment in synthetic biology by the government has become more organized through the creation of BioMADE, whose goal is to build up the bio industrial manufacturing ecosystem in the United States [213]. The efficient organization of producers will enable more substantial market capture

for synthetic biology firms in the specialty chemicals market and a collective effort to tackle some of the challenges described in Section 4.4.2.

While a large push for synthetic biology has been occurring in the United States, globally there has also been development as shown in Table 4-3. Globalization of this technology is important to understand its true value due to the opportunity present in developing countries for many specialty chemicals. Therefore, collaboration across global organizations, like has been done in academia, will help to enable further emergence of synthetic biology and the full realization of its potential.

For the large potential market, there are still very few players in this industry and very few large revenue generators. As synthetic biology R&D enables better and more competitive production methods and address the challenges presented in Section 4.4.2 it is anticipated that more specialty chemical products will be optimized for biological production. The competitive landscape of synthetic biology-leveraging firms is continuing to evolve towards a future more efficient bioeconomy.

4.6 Conclusions

Firms leveraging synthetic biology for the optimization of bioproduction have an opportunity to disrupt the specialty chemicals industry through differentiation in terms of feedstock usage and novel product offerings. The specialty chemicals industry is attractive to new entrants due to its large market and reasonable profitability, however, market forces such as buyer and supplier power, industry rivalry, and substitutes present risks that the industry will grow less attractive in the future. Synthetic biology firms can leverage their technical moats and clear differentiation to take advantage of an attractive market with additional protections against many of the market forces that incumbents face. The current competitive landscape of firms taking advantage of this technology highlights its many strategic benefits but also reveals multiple challenges that must be addressed. A more efficient and sustainable bioeconomy is possible with continued investment and innovation surrounding the synthetic biology field.

5. Conclusions and Outlook

5.1 Summary and Future Directions

In this work, we developed a production platform for the biosynthesis of D-glyceric acid from D-galacturonate and expand upon its application to mixed-sugar feeds. The utilization of mixed feedstock is important to the industrial relevance and sustainability of the process as it allows for the use of renewable agricultural wastes. However, carbon catabolite repression (CCR), an intrinsic signaling network in *Escherichia coli*, complicates the usage of mixed feedstocks as it dictates a utilization preference for glucose over other sugars. Therefore, this work also focused on examining reported strategies for the relaxation of the effects of CCR as well as introducing two novel strategies. All these strategies were classified as being either global (engineering global systems) or sugar-specific (engineering orthogonal sugar systems) relaxation methods. Additionally, this work presented a business case analysis for the disruption of the specialty chemicals industry by synthetic biology leveraging firms. The capabilities of synthetic biology to utilize sustainable feedstock and synthesize novel products were highlighted.

The partially synthetic D-glyceric acid pathway included two heterologous enzymes, *udh* and *gli*, and three native enzymes, *garD, garL,* and *garR*. Through optimization, strain MG1655(DE3) Δ*garK*Δ*hyi*Δ*glxK*Δ*uxaC* was engineered and enabled production of 4.8 g/l of D-glyceric acid with a molar yield of 83%. The application of this production system to a minimal medium displayed a hierarchy in secondary sugar feeds, with galactose and glycerol performing the best, and enabled the development of an extraction-based purification procedure that enriched the concentration of D-glyceric acid 150-fold in solution. We also developed a novel substrate-based induction platform that enabled the expression of *udh* and *gli* upon the addition of D-galacturonate to the fermentation medium. Further work could be done on this system by analyzing substrate feeding strategies to determine whether a bolus or continuous addition gives better production.

Additionally, the production platform could be scaled up to larger volumes using a bioreactor to determine how the system performs at scale.

The introduction of glucose for the minimal medium experiment led to a low titer due to CCR. To enable global CCR relaxation, we developed a system that utilized a previously characterized EIIA (H75Q) protein that enabled the co-consumption of sugar mixtures containing glucose, galacturonate, and lactose. As is common with global interventions, however, we found that this system was phenotypically unstable and was not useful for an application to our D-glyceric acid production platform. To enable sugar-specific CCR relaxation we developed a screening platform to enable the high-throughput analysis of a library of mutant galacturonate permeases (ExuT). This strategy implemented a galacturonate-specific biosensor that enabled FACS-based screening. Expressing an ExuT (S391R) mutant enabled a 20% increase in D-glyceric acid titer in comparison to expressing the wildtype ExuT when glucose and galacturonate were co-fed. Further analysis into the structural consequences of this specific mutation could be done to understand its effects on galacturonate transport. It also would be interesting to repeat the directed evolution studies to discover if this mutant is consistently isolated or if another evolutionary path is revealed. Analysis into methods to further optimize the screening system would be beneficial as in its current state, it often has a large amount of noise and gives contradicting results at times. Finally, a project that utilizes this same framework but with another sugar transport system would be interesting to pursue to determine the ubiquity of this analysis and engineering method.

An analysis of the potential for synthetic biology to disrupt the specialty chemicals industry showed both the attractiveness of the market as well as the capabilities that biotechnology firms should utilize to differentiate themselves. The specialty chemicals industry is attractive based on its high profitability with low volatility, large size, and large total addressable market. Synthetic

biology companies should leverage their ability to utilize sustainable feedstocks and synthesize novel products to differentiate from incumbents, generating a technology-based moat that will limit the threat of commoditization. The current competitive landscape of synthetic biology companies in the specialty chemicals industry shows that there are multiple challenges to still be addressed, including regulatory and supply chain issues. However, continued innovation in this space will enable further market capture by these firms, leading to a more prevalent bio-based economy. Continued analysis of the temporal evolution of this industry will enable a better understanding of both successful and unsuccessful strategies that will be useful for new entrants.

5.2 Outlook

5.2.1 Metabolic Engineering for the Production of Sustainable Monomers

In the past couple of decades, the development of both sustainable processes and products has become an important issue in the chemical industry. This has led companies such as DuPont and BASF to rework their business plans to center around 'greener' applications and market many of their products based on the environmental impact of their production platform or biodegradability of their bioplastics [214].

Utilizing bioprocessing for the synthesis of monomers has been shown to be effective at both making processes more sustainable as well as opening the door to many novel plastics. The ever-expanding toolbox of a metabolic engineer has enabled the production of many monomer classes by engineered microorganisms including: diols, dicarboxylic acids, aromatics, diamines, and hydroxyl acids [215]. Polymerization of these monomers have led to many novel plastics with attractive properties, such as the development of a sustainable polyurethane from β-methyl-δ-valerolactone [216]. Metabolic engineering also enables the tuning of plastic properties by modifying side chains of monomers through additional enzymatic reactions, further increasing its utility for novel plastic development. Incredibly, methods have even been developed that construct microbial 'plastic factories' that produce polymers, such as PHAs [217]. Further development of this field lends itself to designer polymers that are optimized for their specific use while being synthesized using sustainable practices [218].

Another benefit of biopolymers is their increased biodegradability due to being produced from natural sources. In comparison to most fossil-based polymers, bioplastics have a much shorter anaerobic degradation time [219]. Additionally, interesting work has been done for the development of engineered microorganisms that are able to degrade synthetic plastics, decreasing

the burden on the development of biodegradable bioplastics [220]. Continued work to develop novel bioplastics with improved degradation characteristics, through approaches such as the modeling of degradation rates [221], will further increase the prevalence of bioprocessing in the plastics industry.

5.2.2 Utilization of Sustainable Feedstock for Chemical Production

Another major draw to the use of bioprocessing for the production of chemicals is the ability of microorganisms to be easily engineered to utilize sustainable feedstock. Agricultural wastes, including lignocellulosic biomass [222] and food waste [102], present opportunities for optimized usage due to their high sugar content. This has the potential to decrease our dependence on petroleum, which causes diplomatic, economic, and environmental uncertainty, and also enable the increased valorization of these materials, which are often under-utilized. However, the use of agricultural waste also presents new challenges that must be addressed before its use can be ubiquitous.

From a logistical perspective, these materials present a complication due to their often de-centralized production [99]. In addition, the complexity of the material itself also presents challenges. Harsh chemical pre-treatment is often necessary to break down the complex carbohydrates present in the agricultural waste, which can hinder the sustainability of the process [223]. Additionally, the complex mixture of sugars present in agriculture waste can also lend itself to necessitate 'biological funneling', designs for selective valorization of substrate to a specific compound, to enable pure chemical production and limit byproduct formation [224]. However, strategies for the effective utilization of each constituent in these sugar mixtures have been developed and show promise compared to single sugar feeds [18]. Harnessing the full capability

of these mixed feedstocks will enable bioprocessing to become more optimized for both chemical production and overall sustainability.

5.2.3 Novel Engineering Approaches to Carbon Catabolite Repression Alleviation

The complexity of agricultural waste feedstock introduces the challenge of effectively utilizing all the feed constituents. Carbon catabolite repression is a common cause of preferential carbon source usage due to intrinsic signaling mechanisms that prioritize the catabolism of 'valuable' sugars, namely glucose. Many strategies for the co-utilization of sugar mixtures containing glucose have been reported in the literature, most of which involve the engineering of global signaling or glucose uptake systems. While these engineering approaches have led to co-utilization phenotypes, the growth and glucose uptake rate of the strain are often decreased and necessitate additional intervention to increase utility.

An alternative approach is to engineer orthogonal, sugar-specific, signaling and utilization channels to enable defined sugar mixture co-utilization. This strategy circumvents the modification of global signaling systems that leads to hindrance of cell growth and does not change other sugar utilization channels. There is a large opportunity in this space to develop better methods for carbon catabolite repression relaxation that could outperform current strategies. The challenge of this approach is the need for orthogonal systems to be identified and the complete understanding of how they operate. Therefore, very few examples of carbon catabolite repression relaxation using these methods have been reported in the literature. As the understanding of sugar catabolism systems increases, however, more opportunities present themselves to pursue these strategies. A better solution to carbon source co-utilization and the relaxation of carbon catabolite repression will enable better and more efficient bioprocessing and further increase its ability to compete with petroleum-based methods.

5.2.4 Biosensors as Tools for High-Throughput Screening

High–throughput testing of large libraries has become imperative for the discovery of novel proteins and the optimization of bioprocesses. The theory of directed evolution necessitates the assembly of a large, diverse, mutant protein library that can be tested for efficacy and subsequently enriched by maintaining beneficial mutations. This revolutionizes the design-build-test metabolic engineering cycle that informs the development of novel systems. The design step often requires rational design to initially develop novel pathway assemblies. However, as the cycle progresses, the results from testing inform further design decisions. Recently, the toolbox available to metabolic engineers for the build part of the cycle has been enriched to enable the generation of large libraries of diverse protein mutants. Micro-titer plates and automation do not provide enough testing power to keep up with the large libraries that are capable of being generated during the build phase. However, these conventional methods are still important due to the accurate and quantitative results they provide. The development of high-throughput testing has the power to eliminate the bottleneck that is present in the test phase of the cycle. The application of biosensors to high-throughput screening and sorting methods enables the cursory analysis of millions of mutants in a short period of time to identify candidates for intensified analysis. Biosensor-based screening will transform the design-test-build cycle by complementing conventional screening methods to increase the design space available to metabolic engineers.

References

1. Hornsey, I.S., *A history of beer and brewing*. Vol. 34. 2003: Royal Society of Chemistry.
2. Cohen, S.N., et al., *Construction of biologically functional bacterial plasmids in vitro.* Proceedings of the National Academy of Sciences, 1973. **70**(11): p. 3240-3244.
3. Bailey, J.E., *Toward a science of metabolic engineering.* Science, 1991. **252**(5013): p. 1668-1675.
4. Stephanopoulos, G., *Metabolic fluxes and metabolic engineering.* Metabolic engineering, 1999. **1**(1): p. 1-11.
5. Bulter, T., J.R. Bernstein, and J.C. Liao, *A perspective of metabolic engineering strategies: moving up the systems hierarchy.* Biotechnology and bioengineering, 2003. **84**(7): p. 815-821.
6. Wheeler, D.L., et al., *Database resources of the national center for biotechnology information.* Nucleic acids research, 2007. **36**(suppl_1): p. D13-D21.
7. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic acids research, 2000. **28**(1): p. 27-30.
8. UniProt, C., *UniProt: a worldwide hub of protein knowledge.* Nucleic acids research, 2019. **47**(D1): p. D506-D515.
9. Salis, H.M., E.A. Mirsky, and C.A. Voigt, *Automated design of synthetic ribosome binding sites to control protein expression.* Nature Biotechnology, 2009. **27**(10): p. 946-950.
10. Seo, S.W., et al., *Synthetic biology: Tools to design microbes for the production of chemicals and fuels.* Biotechnology Advances, 2013. **31**(6): p. 811-817.
11. Lee, J.W., et al., *Microbial production of building block chemicals and polymers.* Current Opinion in Biotechnology, 2011. **22**(6): p. 758-767.
12. Casini, A., et al., *A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology.* J Am Chem Soc, 2018. **140**(12): p. 4302-4316.
13. Hill, J., et al., *Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels.* Proceedings of the National Academy of Sciences, 2006. **103**(30): p. 11206.
14. Erickson, B., Nelson, and P. Winters, *Perspective on opportunities in industrial biotechnology in renewable chemicals.* Biotechnology Journal, 2012. **7**(2): p. 176-185.
15. Chen, G.-Q., *New challenges and opportunities for industrial biotechnology.* Microbial Cell Factories, 2012. **11**(1): p. 111.
16. Zheng, Y., et al., *Evaluation of Different Biomass Materials as Feedstock for Fermentable Sugar Production*, in *Applied Biochemistry and Biotecnology*. 2007, Humana Press. p. 423-435.
17. Shallom, D. and Y. Shoham, *Microbial hemicellulases.* Current Opinion in Microbiology, 2003. **6**(3): p. 219-228.
18. Liu, N., S. Santala, and G. Stephanopoulos, *Mixed carbon substrates: a necessary nuisance or a missed opportunity?* Current Opinion in Biotechnology, 2020. **62**: p. 15-21.
19. Görke, B. and J. Stülke, *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.* Nature Reviews Microbiology, 2008. **6**(8): p. 613-624.
20. Gosset, G., et al., *Transcriptome Analysis of Crp-Dependent Catabolite Control of Gene Expression in Escherichia coli.* Journal of Bacteriology, 2004. **186**(11): p. 3516.

21. Deutscher, J., C. Francke, and P.W. Postma, *How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria.* Microbiology and Molecular Biology Reviews, 2006. **70**(4): p. 939.
22. Aidelberg, G., et al., *Hierarchy of non-glucose sugars in Escherichia coli.* BMC Systems Biology, 2014. **8**(1).
23. Ammar, E.M., X. Wang, and C.V. Rao, *Regulation of metabolism in Escherichia coli during growth on mixtures of the non-glucose sugars: arabinose, lactose, and xylose.* Scientific Reports, 2018. **8**(1): p. 609.
24. Okano, H., et al., *Regulation underlying hierarchical and simultaneous utilization of carbon substrates by flux sensors in Escherichia coli.* Nature Microbiology, 2019.
25. Wang, X., et al., *Growth strategy of microbes on mixed carbon sources.* Nature Communications, 2019. **10**(1): p. 1279.
26. Kremling, A., et al., *Understanding carbon catabolite repression in Escherichia coli using quantitative models.* Trends in Microbiology, 2015. **23**(2): p. 99-109.
27. Choe, M., et al., *Structural insight into glucose repression of the mannitol operon.* Scientific Reports, 2019. **9**(1): p. 13930.
28. Balderas-Hernandez, V.E., et al., *Improvement of hydrogen production by metabolic engineering of Escherichia coli: Modification on both the PTS system and central carbon metabolism.* International Journal of Hydrogen Energy, 2019.
29. Park, H.-C., et al., *Production of d-ribose by metabolically engineered Escherichia coli.* Process Biochemistry, 2017. **52**: p. 73-77.
30. Zhang, Y., et al., *Production of d-Xylonate from Corn Cob Hydrolysate by a Metabolically Engineered Escherichia coli Strain.* ACS Sustainable Chemistry & Engineering, 2019. **7**(2): p. 2160-2168.
31. Liang, Q., et al., *Comparison of individual component deletions in a glucose-specific phosphotransferase system revealed their different applications.* Scientific Reports, 2015. **5**(1): p. 13200.
32. Bilal, M., et al., *Systematically engineering Escherichia coli for enhanced shikimate biosynthesis co-utilizing glycerol and glucose.* Biofuels, Bioproducts and Biorefining, 2018. **12**(3): p. 348-361.
33. Aguilar, C., et al., *Analysis of differentially upregulated proteins in ptsHIcrr− and rppH− mutants in Escherichia coli during an adaptive laboratory evolution experiment.* Applied Microbiology and Biotechnology, 2018. **102**(23): p. 10193-10208.
34. Kurgan, G., et al., *Parallel experimental evolution reveals a novel repressive control of GalP on xylose fermentation in Escherichia coli.* Biotechnology and Bioengineering, 2019. **116**(8): p. 2074-2086.
35. Sun, J., et al., *Improved ethanol productivity from lignocellulosic hydrolysates by Escherichia coli with regulated glucose utilization.* Microbial Cell Factories, 2018. **17**(1): p. 66.
36. Gao, C., et al., *Programmable biomolecular switches for rewiring flux in Escherichia coli.* Nature Communications, 2019. **10**(1): p. 3751.
37. Kim, J., et al., *Systems Metabolic Engineering of Escherichia coli Improves Coconversion of Lignocellulose-Derived Sugars.* Biotechnology Journal, 2019. **14**(9): p. 1800441.
38. Chappell, T.C. and N.U. Nair, *Co-utilization of hexoses by a microconsortium of sugar-specific E. coli strains.* Biotechnology and Bioengineering, 2017. **114**(10): p. 2309-2318.

39. Wang, L., et al., *Simultaneous fermentation of biomass-derived sugars to ethanol by a co-culture of an engineered Escherichia coli and Saccharomyces cerevisiae.* Bioresource Technology, 2019. **273**: p. 269-276.
40. Zhao, C., et al., *Design and development of a "Y-shaped" microbial consortium capable of simultaneously utilizing biomass sugars for efficient production of butanol.* Metabolic Engineering, 2019. **55**: p. 111-119.
41. Zhang, H., et al., *Engineering Escherichia coli coculture systems for the production of biochemical products.* Proceedings of the National Academy of Sciences, 2015. **112**(27): p. 8266.
42. Saini, M., et al., *Synthetic Consortium of Escherichia coli for n-Butanol Production by Fermentation of the Glucose–Xylose Mixture.* Journal of Agricultural and Food Chemistry, 2017. **65**(46): p. 10040-10047.
43. Harman, J., K. McKenney, and A. Peterkofsky, *Structure-function analysis of three cAMP-independent forms of the cAMP receptor protein.* The Journal of biological chemistry, 1987. **261**: p. 16332-9.
44. Khankal, R., et al., *Transcriptional effects of CRP* expression in Escherichia coli.* Journal of Biological Engineering, 2009. **3**(1): p. 13.
45. Yuan, X., et al., *Combination of the CRP mutation and ptsG deletion in Escherichia coli to efficiently synthesize xylitol from corncob hydrolysates.* Applied Microbiology and Biotechnology, 2020.
46. Wang, X., E.-B. Goh, and H.R. Beller, *Engineering E. coli for simultaneous glucose–xylose utilization during methyl ketone production.* Microbial Cell Factories, 2018. **17**(1): p. 12.
47. Tsai, M.-J., et al., *PredCRP: predicting and analysing the regulatory roles of CRP from its binding sites in Escherichia coli.* Scientific Reports, 2018. **8**(1): p. 951.
48. Frendorf, P.O., et al., *Mutations in the Global Transcription Factor CRP/CAP: Insights from Experimental Evolution and Deep Sequencing.* Computational and Structural Biotechnology Journal, 2019. **17**: p. 730-736.
49. Hollinshead, W.D., et al., *Examining Escherichia coli glycolytic pathways, catabolite repression, and metabolite channeling using Δpfk mutants.* Biotechnology for Biofuels, 2016. **9**(1): p. 212.
50. Shiue, E., I.M. Brockman, and K.L.J. Prather, *Improving product yields on D-glucose in Escherichia coli via knockout of pgi and zwf and feeding of supplemental carbon sources.* Biotechnology and Bioengineering, 2015. **112**(3): p. 579-587.
51. Woo, J.E., et al., *Metabolic Engineering of Escherichia coli for the Production of Hyaluronic Acid From Glucose and Galactose.* Frontiers in bioengineering and biotechnology, 2019. **7**: p. 351-351.
52. Kadoya, R., et al., *Enhanced production of lactate-based polyesters in Escherichia coli from a mixture of glucose and xylose by Mlc-mediated catabolite derepression.* Journal of Bioscience and Bioengineering, 2018. **125**(4): p. 365-370.
53. Lloyd, C., et al., *The Small Protein SgrT Controls Transport Activity of the Glucose-Specific Phosphotransferase System.* Journal of Bacteriology, 2017. **199**: p. JB.00869-16.
54. Vinuselvi, P. and S.K. Lee, *Engineered Escherichia coli capable of co-utilization of cellobiose and xylose.* Enzyme and Microbial Technology, 2012. **50**(1): p. 1-4.

55. Sabri, S., L.K. Nielsen, and C.E. Vickers, *Molecular Control of Sucrose Utilization in Escherichia coli W, an Efficient Sucrose-Utilizing Strain.* Applied and Environmental Microbiology, 2013. **79**(2): p. 478.
56. Koirala, S., X. Wang, and C.V. Rao, *Reciprocal Regulation of L-Arabinose and D-Xylose Metabolism in Escherichia coli.* Journal of Bacteriology, 2016. **198**(3): p. 386.
57. Kim, S.M., et al., *Simultaneous utilization of glucose and xylose via novel mechanisms in engineered Escherichia coli.* Metabolic Engineering, 2015. **30**: p. 141-148.
58. Sievert, C., et al., *Experimental evolution reveals an effective avenue to release catabolite repression via mutations in XylR.* Proceedings of the National Academy of Sciences, 2017. **114**(28): p. 7349.
59. Martinez, R., et al., *The XylR variant (R121C and P363S) releases arabinose-induced catabolite repression on xylose fermentation and enhances coutilization of lignocellulosic sugar mixtures.* Biotechnology and Bioengineering, 2019. **116**(12): p. 3476-3481.
60. Huo, G., et al., *Metabolic engineering of an E. coli ndh knockout strain for PHB production from mixed glucose–xylose feedstock.* Journal of Chemical Technology & Biotechnology, 2017. **92**(10): p. 2739-2745.
61. Ren, C., et al., *An evolved xylose transporter from Zymomonas mobilis enhances sugar transport in Escherichia coli.* Microbial Cell Factories, 2009. **8**(1): p. 66.
62. Gawand, P., et al., *Novel approach to engineer strains for simultaneous sugar utilization.* Metabolic Engineering, 2013. **20**: p. 63-72.
63. Lim, H.G., S.W. Seo, and G.Y. Jung, *Engineered Escherichia coli for simultaneous utilization of galactose and glucose.* Bioresource Technology, 2013. **135**: p. 564-567.
64. Novak, K., et al., *Characterizing the effect of expression of an acetyl-CoA synthetase insensitive to acetylation on co-utilization of glucose and acetate in batch and continuous cultures of E. coli W.* Microbial Cell Factories, 2018. **17**(1): p. 109.
65. Alva, A., et al., *New insights into transport capability of sugars and its impact on growth from novel mutants of Escherichia coli.* Applied Microbiology and Biotechnology, 2020. **104**(4): p. 1463-1479.
66. Zhang, F. and J. Keasling, *Biosensors and their applications in microbial metabolic engineering.* Trends in Microbiology, 2011. **19**(7): p. 323-329.
67. Ravikumar, S., et al., *Engineered microbial biosensors based on bacterial two-component systems as synthetic biotechnology platforms in bioremediation and biorefinery.* Microbial cell factories, 2017. **16**(1): p. 62-62.
68. Mehdizadeh Aghdam, E., M.S. Hejazi, and A. Barzegar, *Riboswitches: From living biosensors to novel targets of antibiotics.* Gene, 2016. **592**(2): p. 244-259.
69. Mahr, R. and J. Frunzke, *Transcription factor-based biosensors in biotechnology: current state and future prospects.* Applied Microbiology and Biotechnology, 2015. **100**(1): p. 79-90.
70. Ravikumar, S., et al., *Construction of Copper Removing Bacteria Through the Integration of Two-Component System and Cell Surface Display.* Applied Biochemistry and Biotechnology, 2011. **165**(7): p. 1674-1681.
71. Ravikumar, S., et al., *A study on the dynamics of the zraP gene expression profile and its application to the construction of zinc adsorption bacteria.* Bioprocess and Biosystems Engineering, 2011. **34**(9): p. 1119.
72. Levskaya, A., et al., *Engineering Escherichia coli to see light.* Nature, 2005. **438**(7067): p. 441-442.

73. Sudarsan, N., et al., *An mRNA structure in bacteria that controls gene expression by binding lysine.* (0890-9369 (Print)).
74. Ott, E., J. Stolz, and M. Mack, *The RFN riboswitch of Bacillus subtilis is a target for the antibiotic roseoflavin produced by Streptomyces davawensis.* RNA Biology, 2009. **6**(3): p. 276-280.
75. Thore, S., M. Leibundgut, and N. Ban, *Structure of the Eukaryotic Thiamine Pyrophosphate Riboswitch with Its Regulatory Ligand.* Science, 2006. **312**(5777): p. 1208.
76. Beisel, C.L. and C.D. Smolke, *Design principles for riboswitch function.* PLoS computational biology, 2009. **5**(4): p. e1000363-e1000363.
77. Espah Borujeni, A., et al., *Automated physics-based design of synthetic riboswitches from diverse RNA aptamers.* Nucleic Acids Research, 2015. **44**(1): p. 1-13.
78. Dixon, N., et al., *Reengineering orthogonally selective riboswitches.* Proceedings of the National Academy of Sciences, 2010. **107**(7): p. 2830.
79. Topp, S., et al., *Synthetic Riboswitches That Induce Gene Expression in Diverse Bacterial Species.* Applied and Environmental Microbiology, 2010. **76**(23): p. 7881.
80. Slusarczyk, A.L., A. Lin, and R. Weiss, *Foundations for the design and implementation of synthetic genetic circuits.* Nature Reviews Genetics, 2012. **13**(6): p. 406-420.
81. Dietrich, J.A., et al., *Transcription Factor-Based Screens and Synthetic Selections for Microbial Small-Molecule Biosynthesis.* ACS Synthetic Biology, 2013. **2**(1): p. 47-58.
82. Zhang, F., J.M. Carothers, and J.D. Keasling, *Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids.* Nature Biotechnology, 2012. **30**(4): p. 354-359.
83. Tang, S.-Y. and P.C. Cirino, *Design and Application of a Mevalonate-Responsive Regulatory Protein.* Angewandte Chemie International Edition, 2011. **50**(5): p. 1084-1086.
84. Combs, S.A., et al., *Small-molecule ligand docking into comparative models with Rosetta.* Nature protocols, 2013. **8**(7): p. 1277.
85. Doong, S.J., A. Gupta, and K.L.J. Prather, *Layered dynamic regulation for improving metabolic pathway productivity in Escherichia coli.* Proceedings of the National Academy of Sciences, 2018. **115**(12): p. 2964.
86. Dinh, C.V. and K.L.J. Prather, *Development of an autonomous and bifunctional quorum-sensing circuit for metabolic flux control in engineered Escherichia coli.* Proceedings of the National Academy of Sciences, 2019. **116**(51): p. 25562.
87. Arnold, F.H., *Design by Directed Evolution.* Accounts of Chemical Research, 1998. **31**(3): p. 125-131.
88. Abatemarco, J., A. Hill, and H.S. Alper, *Expanding the metabolic engineering toolbox with directed evolution.* Biotechnology Journal, 2013. **8**(12): p. 1397-1410.
89. Wang, Q., S.-Y. Tang, and S. Yang, *Genetic biosensors for small-molecule products: Design and applications in high-throughput screening.* Frontiers of Chemical Science and Engineering, 2017. **11**(1): p. 15-26.
90. Furubayashi, M., et al., *A High-Throughput Colorimetric Screening Assay for Terpene Synthase Activity Based on Substrate Consumption.* PLOS ONE, 2014. **9**(3): p. e93317.
91. Santos, C.N.S. and G. Stephanopoulos, *Melanin-Based High-Throughput Screen for l-Tyrosine Production in Escherichia coli.* Applied and Environmental Microbiology, 2008. **74**(4): p. 1190.

92. Alper, H., et al., *Engineering Yeast Transcription Machinery for Improved Ethanol Tolerance and Production.* Science, 2006. **314**(5805): p. 1565.
93. Schallmey, M., et al., *Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors.* Current Opinion in Biotechnology, 2014. **26**: p. 148-154.
94. Zhang, X., et al., *Integration of ARTP mutagenesis with biosensor-mediated high-throughput screening to improve l-serine yield in Corynebacterium glutamicum.* Applied Microbiology and Biotechnology, 2018. **102**(14): p. 5939-5951.
95. Chen, W., et al., *Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis.* Metabolic Engineering, 2015. **30**: p. 149-155.
96. Wu, J., et al., *Design and application of a lactulose biosensor.* Scientific Reports, 2017. **7**(1): p. 45994.
97. Tang, S.-Y., et al., *Screening for enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed triacetic acid lactone reporter.* Journal of the American Chemical Society, 2013. **135**(27): p. 10099-10103.
98. Thomas, J., *Reflections on the topic of solar fuels.* Energy & Environmental Science, 2014. **7**: p. 19.
99. Narodoslawsky, M., A. Niederl-Schmidinger, and L. Halasz, *Utilising renewable resources economically: new challenges and chances for process development.* Journal of Cleaner Production, 2008. **16**(2): p. 164-170.
100. Girotto, F., L. Alibardi, and R. Cossu, *Food waste generation and industrial uses: A review.* Waste Management, 2015. **45**: p. 32-41.
101. Papargyropoulou, E., et al., *The food waste hierarchy as a framework for the management of food surplus and food waste.* Journal of Cleaner Production, 2014. **76**: p. 106-115.
102. Mirabella, N., V. Castellani, and S. Sala, *Current options for the valorization of food manufacturing waste: a review.* Journal of Cleaner Production, 2014. **65**: p. 28-41.
103. Elbeshbishy, E., et al., *Single and combined effect of various pretreatment methods for biohydrogen production from food waste.* International Journal of Hydrogen Energy, 2011. **36**(17): p. 11379-11387.
104. Moon, H.C., et al., *Enzymatic hydrolysis of food waste and ethanol fermentation.* International Journal of Energy Research, 2009. **33**(2): p. 164-172.
105. Liguori, R. and V. Faraco, *Biological processes for advancing lignocellulosic waste biorefinery by advocating circular economy.* Bioresource Technology, 2016. **215**: p. 13-20.
106. Zhang, J., A. Babtie, and G. Stephanopoulos, *Metabolic engineering: enabling technology of a bio-based economy.* Current Opinion in Chemical Engineering, 2012. **1**(4): p. 355-362.
107. Yu, C., et al., *Metabolic engineering of Escherichia coli for biotechnological production of high-value organic acids and alcohols.* Applied Microbiology and Biotechnology, 2011. **89**(3): p. 573-583.
108. Li, X., Y. Chen, and J. Nielsen, *Harnessing xylose pathways for biofuels production.* Current Opinion in Biotechnology, 2019. **57**: p. 56-65.
109. Habe, H., et al., *Biotechnological production of d-glyceric acid and its application.* Applied Microbiology and Biotechnology, 2009. **84**(3): p. 445-452.

110. Sato, S., et al., *Comparative Study of Interfacial and Biological Properties in d-Glycerate-Derived Surfactants.* Journal of the American Oil Chemists' Society, 2017. **94**(11): p. 1393-1401.
111. Habe, H., et al., *Production of Glyceric Acid by Gluconobacter sp. NBRC3259 Using Raw Glycerol.* Bioscience, Biotechnology, and Biochemistry, 2009. **advpub**: p. 0907091558-0907091558.
112. Habe, H., et al., *Biotransformation of glycerol to d-glyceric acid by Acetobacter tropicalis.* Applied Microbiology and Biotechnology, 2009. **81**(6): p. 1033-1039.
113. Wang, B., *Production of D-Glyceric Acid by a Two-step Culture Strategy Based on Whole-cell Biocatalysis of Acetobacter tropicalis.* Chemical and Biochemical Engineering Quarterly, 2018. **32**: p. 135-140.
114. Habe, H., et al., *Microbial Production of Glyceric Acid, an Organic Acid That Can Be Mass Produced from Glycerol.* Applied and Environmental Microbiology, 2009. **75**(24): p. 7760-7766.
115. Kondamudi, N., et al., *Simultaneous production of glyceric acid and hydrogen from the photooxidation of crude glycerol using TiSi2.* Applied Catalysis B: Environmental, 2012. **126**: p. 180-185.
116. Carrettin, S., et al., *Selective oxidation of glycerol to glyceric acid using a gold catalyst in aqueous sodium hydroxide.* Chemical Communications, 2002(7): p. 696-697.
117. Gerstenbruch, S., et al., *Asymmetric synthesis of d-glyceric acid by an alditol oxidase and directed evolution for enhanced oxidative activity towards glycerol.* Applied Microbiology and Biotechnology, 2012. **96**(5): p. 1243-1252.
118. Müller-Maatsch, J., et al., *Pectin content and composition from different food waste streams.* Food Chemistry, 2016. **201**: p. 37-45.
119. Blumenthal, H.J. and D.C. Fish, *Bacterial conversion of D-glucarate to glycerate and pyruvate.* 1963.
120. Weickert, M.J., et al., *Optimization of heterologous protein production in Escherichia coli.* Current opinion in biotechnology, 1996. **7**(5): p. 494-499.
121. Beckwith, J.R. and D. Zipser, *The lactose operon*. Vol. 1. 1970: Cold Spring Harbor Laboratory.
122. Valdez-Cruz, N.A., et al., *Production of recombinant proteins in E. coli by the heat inducible expression system based on the phage lambda pL and/or pR promoters.* Microbial cell factories, 2010. **9**(1): p. 18.
123. Gupta, A., et al., *Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit.* Nature Biotechnology, 2017. **35**(3): p. 273-279.
124. Kang, Z., et al., *Construction of a stress-induced system in Escherichia coli for efficient polyhydroxyalkanoates production.* Applied microbiology and biotechnology, 2008. **79**(2): p. 203-208.
125. Sambrook, J., *Molecular cloning : a laboratory manual*. 2001: Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001.
126. Quan, J. and J. Tian, *Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries.* Nature Protocols, 2011. **6**(2): p. 242-251.
127. Engler, C. and S. Marillonnet, *Golden Gate Cloning*, in *DNA Cloning and Assembly Methods*, S. Valla and R. Lale, Editors. 2014, Humana Press: Totowa, NJ. p. 119-131.

128. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proceedings of the National Academy of Sciences, 2000. **97**(12): p. 6640.
129. Liu, N., K. Qiao, and G. Stephanopoulos, *13C Metabolic Flux Analysis of acetate conversion to lipids by Yarrowia lipolytica.* Metabolic engineering, 2016. **38**: p. 86-97.
130. Yoon, S.-H., et al., *Cloning and Characterization of Uronate Dehydrogenases from Two Pseudomonads and Agrobacterium tumefaciens Strain C58.* Journal of Bacteriology, 2009. **191**(5): p. 1565.
131. Ashiuchi, M. and H. Misono, *Biochemical evidence that Escherichia coli hyi (orf b0508, gip) gene encodes hydroxypyruvate isomerase11The nucleotide sequence will appear in the DDBJ/EMBL/GenBank databases with the accession number AB028043.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1999. **1435**(1): p. 153-159.
132. Ornston, M.K. and L.N. Ornston, *Two Forms of D-Glycerate Kinase in Escherichia coli.* Journal of Bacteriology, 1969. **97**(3): p. 1227.
133. Fujiwara, R., et al., *Metabolic engineering of Escherichia coli for shikimate pathway derivative production from glucose–xylose co-substrate.* Nature Communications, 2020. **11**(1): p. 279.
134. *Glycerate, TBDMS (Compound).* Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Glycerate_-TBDMS#section=Mass-Spectrometry.
135. Naik, S., et al., *Characterization of Canadian biomass for alternative renewable biofuel.* Renewable Energy, 2010. **35**(8): p. 1624-1631.
136. Stephanopoulos, G., *Metabolic engineering.* Biotechnology and Bioengineering, 1998. **58**(2-3): p. 119-120.
137. Green, J., et al., *Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches.* Current Opinion in Microbiology, 2014. **18**: p. 1-7.
138. Presper, K.A., et al., *Site-directed mutagenesis of the phosphocarrier protein. IIIGlc, a major signal-transducing protein in Escherichia coli.* Proceedings of the National Academy of Sciences, 1989. **86**(11): p. 4052-4055.
139. Zeng, G.Q., H. De Reuse, and A. Danchin, *Mutational analysis of the enzyme IIIGlc of the phosphoenolpyruvate phosphotransferase system in Escherichia coli.* Research in Microbiology, 1992. **143**(3): p. 251-261.
140. Saier, M.H., et al., *Permease-specific mutations in Salmonella typhimurium and Escherichia coli that release the glycerol, maltose, melibiose, and lactose transport systems from regulation by the phosphoenolpyruvate:sugar phosphotransferase system.* Journal of Bacteriology, 1978. **133**(3): p. 1358.
141. Titgemeyer, F., R.E. Mason, and M.H. Saier, Jr., *Regulation of the raffinose permease of Escherichia coli by the glucose-specific enzyme IIA of the phosphoenolpyruvate:sugar phosphotransferase system.* Journal of bacteriology, 1994. **176**(2): p. 543-546.
142. Saier, M.H. and M. Crasnier, *Inducer exclusion and the regulation of sugar transport.* Research in Microbiology, 1996. **147**(6): p. 482-489.
143. Ni, C., K.J. Fox, and K.L.J. Prather, *Substrate-activated expression of a biosynthetic pathway in Escherichia coli.* Biotechnology Journal, 2021. **n/a**(n/a): p. 2000433.

144. St-Pierre, F., et al., *One-Step Cloning and Chromosomal Integration of DNA.* ACS Synthetic Biology, 2013. **2**(9): p. 537-541.
145. Dietrich, J.A., A.E. McKee, and J.D. Keasling, *High-Throughput Metabolic Engineering: Advances in Small-Molecule Screening and Selection.* Annual Review of Biochemistry, 2010. **79**(1): p. 563-590.
146. Sjostrom, S.L., et al., *High-throughput screening for industrial enzyme production hosts by droplet microfluidics.* Lab on a Chip, 2014. **14**(4): p. 806-813.
147. UniProt. *UniProtKB - P0AA78 (EXUT_ECOLI)*. Available from: https://www.uniprot.org/uniprot/P0AA78.
148. Hatzikyriakou, M., *Specialty chemicals and contending with forces of change.* Journal of Business Chemistry, 2018. **15**(3): p. 94.
149. Clomburg, J.M. and R. Gonzalez, *Biofuel production in Escherichia coli: the role of metabolic engineering and synthetic biology.* Applied microbiology and biotechnology, 2010. **86**(2): p. 419-434.
150. Savage, D.F., J. Way, and P.A. Silver, *Defossiling fuel: how synthetic biology can transform biofuel production*. 2008, Citeseer.
151. Wyman, C.E., *Ethanol production from lignocellulosic biomass: overview.* Handbook on Bioethanol, 2018: p. 1-18.
152. Hamelinck, C.N., G.v. Hooijdonk, and A.P.C. Faaij, *Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term.* Biomass and Bioenergy, 2005. **28**(4): p. 384-410.
153. Bala, S., *Specialty Chemicals Integrated Across Multiple Vertical Applications: Global Markets*. 2020.
154. Botte, G.G., *Electrochemical manufacturing in the chemical industry.* The Electrochemical Society Interface, 2014. **23**(3): p. 49.
155. Danielson, N., et al., *Industrial Biotechnology—An Industry at an Inflection Point.* Industrial Biotechnology, 2020. **16**(6): p. 321-332.
156. Linthorst, J.A., *An overview: origins and development of green chemistry.* Foundations of chemistry, 2010. **12**(1): p. 55-68.
157. Le Feuvre, R.A. and N.S. Scrutton, *A living foundry for synthetic biological materials: a synthetic biology roadmap to new advanced materials.* Synthetic and systems biotechnology, 2018. **3**(2): p. 105-112.
158. Sheldon, R.A., *The road to biorenewables: carbohydrates to commodity chemicals.* ACS Sustainable Chemistry & Engineering, 2018. **6**(4): p. 4464-4480.
159. Clarke, L.J., *Synthetic biology–pathways to commercialisation.* Engineering Biology, 2019. **3**(1): p. 2-5.
160. Bueso, Y.F. and M. Tangney, *Synthetic biology in the driving seat of the bioeconomy.* Trends in biotechnology, 2017. **35**(5): p. 373-378.
161. Straathof, A.J.J., et al., *Grand research challenges for sustainable industrial biotechnology.* Trends in biotechnology, 2019. **37**(10): p. 1042-1050.
162. Schmidt, M., et al., *A priority paper for the societal and ethical aspects of synthetic biology.* Systems and synthetic biology, 2009. **3**(1): p. 3-7.
163. Dawid, H. and M. Reimann, *Evaluating market attractiveness: individual incentives versus industry profitability.* Computational Economics, 2005. **24**(4): p. 321-355.
164. Hax, A.C. and N.S. Majluf, *The use of the industry attractiveness-business strength matrix in strategic planning.* Interfaces, 1983. **13**(2): p. 54-71.

165. Wernerfelt, B. and C.A. Montgomery, *What is an attractive industry?* Management Science, 1986. **32**(10): p. 1223-1230.
166. Porter, M.E., *The five competitive forces that shape strategy.* Harvard business review, 2008. **86**(1): p. 78.
167. Ezekoye, O., A. Milutinovic, and T. Jan Simons, *Chemicals and capital markets: Back at the top.* McKinsey on Chemicals, 2018.
168. Aase, G. *A winning formula for specialty-chemical conglomerates*. 2021.
169. Glaschke, M., M. Moder, and C. Schmitz, *What to do When a Specialty Chemical Business Gets Commoditized*. 2017, McKinsey.
170. *World Economic Outlook Reports*. 2021; Available from: https://www.imf.org/en/Publications/WEO.
171. Waerder, B., S. Stinnes, and O. Erdenberger, *Design thinking as driver for innovation in the chemical industry.* Journal of Business Chemistry, 2017. **14**(2).
172. Dickson, D., *2021 Chemical Industry Outlook*. 2020.
173. Carpi, R., et al., *Pursuing purchasing excellence in chemicals*. 2016, McKinsey & Company.
174. Keller, W., *Raw Material Excellence in the chemical industry, a game changer turning cost into value.* Journal of Business Chemistry, 2017. **14**(3).
175. Caruso, D., *Staying Competitive in the Changing Chemical Manufacturing Marketplace*. 2008, Microsoft Dynamics.
176. Budde, F., *Chemicals' changing competitive landscape*. 2011, McKinsey & Company.
177. Valimaki, C., *The inevitability of Chemical Substitution*. 2017, Elsevier.
178. *Biorefinery Products: Global Markets*. 2021, BCC Publishing.
179. Check Hayden, E., *Synthetic-biology firms shift focus.* Nature News, 2014. **505**(7485): p. 598.
180. Carbonell, P., et al., *Mapping the patent landscape of synthetic biology for fine chemical production pathways.* Microbial biotechnology, 2016. **9**(5): p. 687-695.
181. Kircher, M., *Sustainability of biofuels and renewable chemicals production from biomass.* Current Opinion in Chemical Biology, 2015. **29**: p. 26-31.
182. Fiorentino, G., M. Ripa, and S. Ulgiati, *Chemicals from biomass: technological versus environmental feasibility. A review.* Biofuels, Bioproducts and Biorefining, 2017. **11**(1): p. 195-214.
183. Zeng, A.-P., *New bioproduction systems for chemicals and fuels: needs and new development.* Biotechnology advances, 2019. **37**(4): p. 508-518.
184. Markham, K.A. and H.S. Alper, *Synthetic biology for specialty chemicals.* Annual review of chemical and biomolecular engineering, 2015. **6**: p. 35-52.
185. Martinelli, L. and P.I. Nikel, *Breaking the state-of-the-art in the chemical industry with new-to-Nature products via synthetic microbiology.* Microbial biotechnology, 2019. **12**(2): p. 187.
186. Degli Esposti, M., et al., *The role of biotechnology in the transition from plastics to bioplastics: an opportunity to reconnect global growth with sustainability.* FEBS Open bio, 2021. **11**(4): p. 967-983.
187. Bomgardner, M., *Why the flavor and fragrance industry is embracing biotechnology*, in *Chemical and Engineering News*. 2021.
188. Li, X., et al., *Advances on (+)-nootkatone microbial biosynthesis and its related enzymes.* Journal of Industrial Microbiology and Biotechnology, 2021.

189. Paul, V., et al., *A comprehensive review on vanillin: its microbial synthesis, isolation and recovery.* Food Biotechnology, 2021. **35**(1): p. 22-49.
190. Choi, K.R., et al., *Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering.* Trends in biotechnology, 2019. **37**(8): p. 817-837.
191. Chubukov, V., et al., *Synthetic and systems biology for microbial production of commodity chemicals.* NPJ systems biology and applications, 2016. **2**(1): p. 1-11.
192. Si, T. and H. Zhao, *A brief overview of synthetic biology research programs and roadmap studies in the United States.* Synthetic and systems biotechnology, 2016. **1**(4): p. 258-264.
193. Mao, N., et al., *Future trends in synthetic biology in Asia.* Advanced Genetics, 2021. **2**(1): p. e10038.
194. Bolton, R. and R. Thomas, *Biohackers: the science, politics, and economics of synthetic biology.* Innovations: Technology, Governance, Globalization, 2014. **9**(1-2): p. 213-219.
195. Dana, G.V., et al., *Four steps to avoid a synthetic-biology disaster.* Nature, 2012. **483**(7387): p. 29-29.
196. Ribeiro, B. and P. Shapira, *Anticipating governance challenges in synthetic biology: Insights from biosynthetic menthol.* Technological forecasting and social change, 2019. **139**: p. 311-320.
197. Erickson, B., R. Singh, and P. Winters, *Synthetic biology: regulating industry uses of new biotechnologies.* Science, 2011. **333**(6047): p. 1254-1256.
198. Kelley, N.J., et al., *Engineering biology to address global problems: Synthetic biology markets, needs, and applications.* Industrial Biotechnology, 2014. **10**(3): p. 140-149.
199. Stephanopoulos, G., *Challenges in engineering microbes for biofuels production.* Science, 2007. **315**(5813): p. 801-804.
200. Clomburg, J.M., A.M. Crumbley, and R. Gonzalez, *Industrial biomanufacturing: the future of chemical production.* Science, 2017. **355**(6320).
201. Nicholas, J.W., *Bioinformatics for the synthetic biology of natural products: integrating across the Design–Build–Test cycle.* Natural product reports, 2016. **33**(8): p. 925-932.
202. Kurian, J.K., et al., *Feedstocks, logistics and pre-treatment processes for sustainable lignocellulosic biorefineries: A comprehensive review.* Renewable and Sustainable Energy Reviews, 2013. **25**: p. 205-219.
203. Lopes, M.S.G., *Engineering biological systems toward a sustainable bioeconomy.* Journal of Industrial Microbiology and Biotechnology, 2015. **42**(6): p. 813-838.
204. Brooks, S.M. and H.S. Alper, *Applications, challenges, and needs for employing synthetic biology beyond the lab.* Nature Communications, 2021. **12**(1): p. 1-16.
205. Maynard, A.D., *Navigating the fourth industrial revolution.* Nature nanotechnology, 2015. **10**(12): p. 1005-1006.
206. Baker, B., *Synthetic Biology and the Marketplace: Building the new bioeconomy.* BioScience, 2017. **67**(10): p. 877-883.
207. Shapira, P., S. Kwon, and J. Youtie, *Tracking the emergence of synthetic biology.* Scientometrics, 2017. **112**(3): p. 1439-1469.
208. *Pitchbook*. Available from: www.pitchbook.com.
209. Feldman, A. *Synthetic Biology Company Amyris Swapped Biofuels for Personal Care, Its Stock is up 10-fold over the Past Year*. 2021.

210. Birch, K. and D. Tyfield, *Theorizing the bioeconomy: biovalue, biocapital, bioeconomics or... what?* Science, Technology, & Human Values, 2013. **38**(3): p. 299-327.
211. Hytha, M. *"Biofacturing" Firm Zymergen Raises $500 Million in Expanded IPO*. 2021.
212. de Leon, R. *Bill Gates-backed Ginkgo Bioworks going public via $15 billion SPAC*. 2021.
213. *U.S. Department of Defense awards $87.5 million to create new Manufacturing Onnovation Institute sited at the University of Minnesota*. 2020.
214. Iles, A. and A.N. Martin, *Expanding bioplastics production: sustainable business innovation in the chemical industry.* Journal of Cleaner Production, 2013. **45**: p. 38-49.
215. Chung, H., et al., *Bio-based production of monomers and polymers by metabolically engineered microorganisms.* Current opinion in biotechnology, 2015. **36**: p. 73-84.
216. Schneiderman, D.K., et al., *Chemically recyclable biobased polyurethanes.* ACS Macro Letters, 2016. **5**(4): p. 515-518.
217. Aldor, I.S. and J.D. Keasling, *Process design for microbial plastic factories: metabolic engineering of polyhydroxyalkanoates.* Current opinion in biotechnology, 2003. **14**(5): p. 475-483.
218. John, G., et al., *Natural Monomers: A mine for functional and sustainable materials-occurrence, chemical modification and polymerization.* Progress in Polymer Science, 2019.
219. Bátori, V., et al., *Anaerobic degradation of bioplastics: A review.* Waste management, 2018. **80**: p. 406-413.
220. Shimao, M., *Biodegradation of plastics.* Current opinion in biotechnology, 2001. **12**(3): p. 242-247.
221. Ryan, C.A., S.L. Billington, and C.S. Criddle, *Assessment of models for anaerobic biodegradation of a model bioplastic: Poly (hydroxybutyrate-co-hydroxyvalerate).* Bioresource technology, 2017. **227**: p. 205-213.
222. Isikgor, F.H. and C.R. Becer, *Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers.* Polymer Chemistry, 2015. **6**(25): p. 4497-4559.
223. Brodin, M., et al., *Lignocellulosics as sustainable resources for production of bioplastics–A review.* Journal of Cleaner Production, 2017. **162**: p. 646-664.
224. Beckham, G.T., et al., *Opportunities and challenges in biological lignin valorization.* Current opinion in biotechnology, 2016. **42**: p. 40-53.