5-2021

Engineering Modularity of Ester Biosynthesis Across Biological Scales

Hyeongmin Seo
University of Tennessee, Knoxville, hseo5@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Hyeongmin Seo entitled "Engineering Modularity of Ester Biosynthesis Across Biological Scales." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Chemical Engineering.

Cong T. Trinh, Major Professor

We have read this dissertation and recommend its acceptance:

Eric T. Boder, Adam M. Guss, Constance Bailey, Gladys Alexandre

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Engineering Modularity of Ester Biosynthesis Across Biological Scales

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Hyeongmin Seo

May 2021

Dedicated to my wife,

Hanbyeol Joo

ACKNOWLEDGEMENTS

First, I immensely appreciate my primary advisor Dr. Cong T. Trinh who supported, inspired, and advised me during the PhD training. His insightful guidance exposed me to a wide range of techniques and ideas.

Second, I would like to thank committee members, Dr. Eric Boder, Dr. Adam M. Guss, Dr. Constance Bailey, and Dr. Alexandre Gladys for their unwavering support and insightful advices. Next, I appreciate many wonderful colleagues, Seunghyun Ryu, Brian Mendoza, Caleb Walker, Sergio Garcia, Jong-Won Lee, David Dooley, Niantai Li, and other Trinh lab members for fruitful discussion and collaborative works. I also appreciate my outstanding collaborators Dr. Heidi Schindel, Dr. Richard Giannone and Dr. Robert Hettich from Oak Ridge National Laboratory (ORNL), Dr. Charles Cai and Priyanka Singh from University of California Riverside. I am immensely thankful for Center for Bioenergy Innovation (CBI) colleagues, in particular Dr. Lee Lynd, Dr. Brain Davidson, Dr. Josh Michener, Dr. Gerald Tuskan, Dr. Janet Westpheling, Dr. Carrie Eckert, Dr. Lauren Riley, Dr. Yannick Bomble, for their assistance in many stages of my training. Special acknowledgment is required to my undergraduate mentees, Preston Nicely, Noah Dunlap, John Hill, Caleb Young, Marina Vlasyuk, Gillian Castro, and Anthony Huber. They always motivated and taught me to be a better person.

Lastly, I would like to acknowledge that the studied in this dissertation were funded by National Science Foundation (NSF), Joint Genome Institute (JGI), and Center for Bioenergy Innovation (CBI), the U.S. Department of Energy (DOE) Bioenergy Research centers funded by the Office of Biological and Environmental Research in the DOE Office of Science. I was very fortunate to have such a great research opportunity.

ABSTRACT

Metabolic engineering and synthetic biology enable controlled manipulation of whole-cell biocatalysts to produce valuable chemicals from renewable feedstocks in a rapid and efficient manner, helping reduce our reliance on the conventional petroleum-based chemical synthesis. However, strain engineering process is costly and time-consuming that developing economically competitive bioprocess at industrial scale is still challenging. To accelerate the strain engineering process, modular cell engineering has been proposed as an innovative approach that harnesses modularity of metabolism for designing microbial cell factories. It is important to understand biological modularity and to develop design principles for effective implementation of modular cell engineering. In this dissertation, the modularity of ester biosynthesis was engineered from the molecular to the microbial community levels. Specifically, three important features of modularity (i.e., robustness, efficiency, and compatibility) were engineered and quantitatively analyzed across different scales. At the molecular (enzymatic) level, thermostability and promiscuity of alcohol acyltransferases were engineered to develop a robust designer ester biosynthesis. At the metabolic network (cellular) level, metabolism of *Escherichia coli* was rewired to overproduce isoamyl acetate through metabolic engineering and synthetic biology strategies. Also, by harnessing the engineered robust alcohol acyltransferase, a non-model thermophilic bacterium *Clostridium thermocellum* was engineered to produce medium chain esters directly from recalcitrant lignocellulosic biomass at elevated temperatures. Finally, at the microbial community level, a syntrophic *E. coli* co-culture was engineered for isobutyl butyrate production from a mixture of glucose and xylose. The successful engineering of the modularity of ester biosynthesis not only sheds light into the modular design principles of biological systems, but also seeks to develop industrially relevant ester production platforms.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

Biotechnology has led to breakthroughs in many areas including pharmaceutical, agricultural, chemical, material, and environmental industries. Understanding living organisms is increasingly important to tackle many contemporary problems but living organisms are highly complex and diverse. How can we understand and effectively engineer biological systems for multiple purposes? This chapter is a review of current biotechnology for commodity chemical production, providing background information relevant to the work presented throughout in this dissertation.

1-1 Biotechnology for value-added chemical production

Human societies are facing unprecedent challenges such as pandemic[1], climate change and global warming[2], and environmental pollution[3]. Many of the contemporary problems have attributed to our reliance on petroleum[4, 5]. Therefore, it is one of the most urgent tasks to develop sustainable biochemical manufacturing for reducing our reliance on fossil derived energy and materials. White biotechnology, a branch of biotechnology, uses living cells such as yeasts, bacteria, and plants to synthesize products that are easily degradable, require less energy, and create less waste during their production[6]. White biotechnology has suggested an alternative way of synthesizing valuable chemicals from renewable resources that can potentially replace the current petroleum-dependent carbon emitting chemical manufacturing with carbon neutral or negative bioprocessing.

While current biotechnology enables engineering various living organisms from a simple single-cell microbe to multicellular organisms such as plants and animals. Microbes have been commonly deployed as hosts for biochemical production because of their relatively simple metabolisms and fast growth (e.g., minutes to hours for doubling) that makes them tractable in the

laboratory. Also, many microbes are generally recognized as safe, hence engineering them does not provoke ethical problems. Microbes live in almost everywhere: air, water, soil, food, animal intestines, plant roots, rocks, glaciers, hot springs, and even in deep-sea hydrothermal vents[7]. The wide variety of habitats reflects biochemical and metabolic diversity of microbes suggesting unlimited potential of microbial biotechnology.

Traditional microbial biotechnology for value-added chemical production largely focused on natural producers through classical bioprocess engineering. One of the most popular examples is penicillin production by *Penicillium notatum*. After Alexander Fleming's discovery of penicillin in 1928, bioprocess engineering improved production yield and recovery rate and successfully industrialized in the 1940s[8]. The commercialized penicillin has saved millions of human lives since World War II, and many penicillin-derived antibiotic drugs have been clinically used until now. The following decades increasingly utilized microorganisms to synthesize natural products of pharmaceutical interests such as antibiotics and anti-cancer drugs[9]. In this period, strain engineering was limited to mutagenesis and screening due to the lack of techniques for quantitative analysis of cellular metabolism.

In early 1990s, bioinformatics and mathematical modeling methods allowing quantitative analysis brought new insights into cellular metabolism, leading to the beginning of metabolic engineering and synthetic biology that enable specific genetic manipulations altering metabolic flux towards the product of interest[10]. The advent of metabolic engineering and synthetic biology could develop novel bioprocesses by rationally engineering cellular metabolism. Particularly, heterologous pathway engineering could expand bioproducts beyond the traditional natural products. Currently, not only natural products, but also a large space of valuable chemicals

including fuels, polymers, and building block molecules that have been manufactured from fossil feedstock can be synthesized by engineered microbes using renewable resources[11, 12].

1-2 Challenges of engineering biological systems

State-of-the-art systems metabolic engineering tools have improved overall bioprocess from strain engineering to fermentation and product recovery[13]. However, developing a cost competitive bioprocess to the established petroleum-based chemical processes is still challenging. A major barrier is the costly strain development process which typically requires more than a decade and over $50 million to meet the economic requirements for industrial scale production[14]. The primary reason of the costly process is rooted from the nature of living organisms comprising of complex regulation and interactions in metabolic networks. Cellular regulation machinery is a result of long-time evolution ensuring metabolic homeostasis under environmental fluctuations. Therefore, rewiring the metabolism for the overproduction of desired products requires extensive manipulation of regulation machinery and/or insertion of heterologous pathways.

Compatibility of heterologous pathways to a host cell is critical to achieve optimal phenotypes (e.g., efficient production of desired chemicals with stable cell viability). Engineering heterologous pathways often witnesses undesirable metabolic imbalance such as inefficient protein expression, metabolic burden, redox imbalances that reduce production efficiency with low titers, yields, and productivities. The problems are unsatisfactorily predictable that iterative trial-and-error optimization so called design-build-test-learn (DBTL) cycle is required to achieve desirable phenotypes. To accelerate the DBTL cycle, many multiplex engineering tools, high-throughput screening techniques, and bioinformatic methodologies have been developed. Examples of the multiplex engineering tools include multiplex automated genome engineering (MAGE)[15],

trackable multiplex recombineering (TRMR)[16], tracking combinatorial engineered libraries (TRACE)[17]. Recently, clustered regularly interspaced short palindromic repeats (CRISPR) genome editing technologies has been integrated to the multiplex genome engineering techniques, further accelerating the strain development process[18]. The techniques particularly contribute to rapid design and build of a large space of genotypes.

Testing a library of genotypic variants is the most time-consuming and laborious task that often requires high-throughput screening methods. When the desirable phenotypes are related to observable characteristics such as color, fluorescence, and size, the desired mutant cells can be easily acquired using colorimetric assays, fluorescence-activated cell sorting (FACS), spectrophotometer, and microfluidic sorting devices. However, in strain engineering, the desirable phenotypes are not easily observable. Therefore, lack of proper high-throughput methods becomes a critical bottleneck to acquiring the top performing mutants from a population with a large library of genotypes. Synthetic biology strategies can address the issues by coupling the desirable phenotype with expression of observable biological parts (e.g., reporter genes)[19]. Examples include transcription regulators[20, 21], or riboswitch mediated biosensors[22, 23], expressing fluorescence proteins in response to target molecules.

Another strategy of heterologous pathway engineering is metabolic coupling driven selection. A core principle of the strategy is to couple cell growth with synthesis of desired chemicals so that the phenotype can be funneled to maximize the chemical production by employing growth as a simple selection criterion[24, 25]. The RNA mediated biosensors can also couple the cell growth to a metabolite through the genetic circuits regulating expression of antibiotics resistance genes. However, the RNA mediated growth coupling has principal barriers, the possibility of “evolutionary escape” caused by mutations in the biosensor components (e.g.,

transcription factor, reporter gene, promoters), such that the biosensor output does not correspond to concentration of the target metabolites[26]. Cell growth coupling at metabolic pathway level without the specialized biological parts can avoid the unfavorable mutations. Modular cell (MODCELL) engineering is a representative selection-based strain engineering framework which aims to design an auxotrophic chassis cell containing core metabolic pathways that are necessary yet insufficient to support healthy cell growth without the production module[27]. The modular chassis cells can be designed for growth coupling to various biochemical production pathways through computational metabolic pathway modeling and simulations[28, 29]. Theoretically, the approach can bypass the iterative DBTL cycle because the best production module can be isolated through simple transfers of cell cultures under a controlled condition. The modular chassis cell can then select the most compatible heterologous pathway achieving the maximum production from the library of production modules[29]. For example, the modular cell engineering could select the best ethanol pathway enzymes by strong coupling between the production module and cell growth under anaerobic conditions[30]. However, it is critical to ensure that there is a significantly efficient and compatible module to be selected among the library pool to avoid undesirable adaptive evolution of the chassis cell over generations. Without appropriate constraint of a production module library, the cell could adapt to restore the cell growth without metabolic coupling to a production module[30].

Therefore, it is important to establish the design principles for engineering the modules not only for efficient DBTL cycle, but also for practical implementation of the modular cell engineering.

1-3 Biological modularity and modular design in biotechnology

Modular design is a proven engineering principle which has been applied to various manufactures. Many complex systems such as architectures, vehicles, factories, and computers are assembly of separable and exchangeable modules. The module is defined as a separable component, frequently one that is interchangeable with others, for assembly into units of differing size, complexity, or function. Flexible combinations of modules enable rapid construction of a system with improved robustness and efficiency, reducing manufacturing costs[31]. From the definition, modularity is a degree of flexibility determining how much the interfaces can be separated or combined to create a system. Therefore, the idea of modularity refers to a connected pattern where elements are grouped into physically or dynamically connected subsets[32]. Modular design is an engineering principle to subdivide a system into the modules, allowing redesign of the system by combining the modules in various ways.

Biology has been shaped in myriad phenotypes to survive better under diverse environments over millions of years. Although a natural organism is not designed, it comprises separable independent components that allow for the separate evolution of each component[33, 34]. In developmental biology, the modularity has become a central concept describing evolution[35, 36]. The biological module may evolve with limited perturbation to each other, enabling more effective and thorough search for fitness maxima over the evolution[2]. Biological modularity is observed across all scales of biology from DNA and RNA to ecosystem[37]. At molecular level, interactions between intracellular molecules (e.g., proteins, DNA, RNA) formulate a network contributing robust cellular metabolism against environmental perturbations. The most simplified example of the biological modularity at molecular level is an enzyme reaction, physical and chemical interactions between a ligand and a protein. The catalytic mechanism is often shared among the

proteins in a protein family, reflected by their related functions and similarities in structure. The shared features often result in substrate promiscuity of enzymes. At cellular level, many interaction between biomolecules such as DNA[38], RNA[39, 40], proteins[32], and metabolites[41] display modularity that is highly associated with cellular functions and pathways. Therefore, cellular metabolism is operated by the metabolic networks highly connected between the biomolecules and their interactions with one another. Metabolic pathways are sometimes highly conserved in many classes of living organisms. Modularity also exists at microbial community level represented by a network of individual organism interacting each other through consumption, symbiosis, or competition[42].

The modularity of biology brings a lot of insights to biotechnology. In biotechnology, modular design principle can be implemented to reconstruct biological systems rapidly and effectively by engineering and harnessing the modularity[27]. Many metabolic engineering and synthetic biology strategies have utilized the modularity in strain development process. Since metabolic pathways are linked together, deletion/down-regulation or overexpression of a single gene can disrupt or activate a whole pathway module without exhaustive genetic engineering[43, 44]. Engineering a global regulatory protein linked to multiple gene regulation allows multiplex phenotype engineering[45]. A synthetic oscillator can be generated by genetic circuits linked together by biomolecules[46]. A wide range of products can be synthesized by engineering modular polyketide synthase[47-49]. Many more biotechnology studies have implemented the modular design principle to understand and engineer biological systems although the principle might not be stated in the contexts. Because a large space of commodity chemicals can be derived from modular metabolic networks[11], modular design principle is particularly helpful to search for an optimal design of cellular metabolism for chemical production.

1-4 Modularity of ester biosynthesis

Natural formulation of ester is an excellent example representing the modularity across multiple scales (Figure 1-1). At the molecular level, chemistry of ester biosynthesis consists of three key modules: two chemical moieties, an alcohol and an acyl-CoA, that are condensed to esters by an alcohol acetyltransferase (AAT, EC 2.3.1.84). At the cellular level, metabolic pathways form exchangeable production modules providing alcohols and acyl-CoAs from reducing carbon sources, and intracellularly expressed AATs are combined with these modules to produce various esters. Finally, the ester production module is combined with other volatile metabolites to determine floral scents, and also interacts with other environmental modules such as pollinators and pathogens at a community level. The flexible and broad interaction creates a driving force for molecular evolution of the AATs.

In addition to their important roles in nature, esters are an important class of industrially important chemicals that can be used as flavors, fragrances, solvents, and drop-in fuels. Since esters are mainly manufactured from petroleum feedstocks through chemical processes, microbial ester biosynthesis using renewable and sustainable feedstocks (e.g., cellulosic biomass) is a promising alternative approach. Especially, the AAT-dependent pathway is thermodynamically favorable as compared to other chemical and biological pathways such as esterase/lipase pathway, and therefore microbial catalysts that harbor the AAT-dependent pathway can be best used for conversion of fermentable sugars to esters under mild aqueous conditions. For example, an engineered *Escherichia coli* overexpressing ATF1 from *Saccharomyces cerevisiae* produced isobutyl acetate up to 17.5 g/L from glucose[50], suggesting a feasible microbial catalyst for ester production. Also, engineered *E. coli* biocatalysts expressing the AATs originating from plants could upgrade waste organic acids to esters[51].

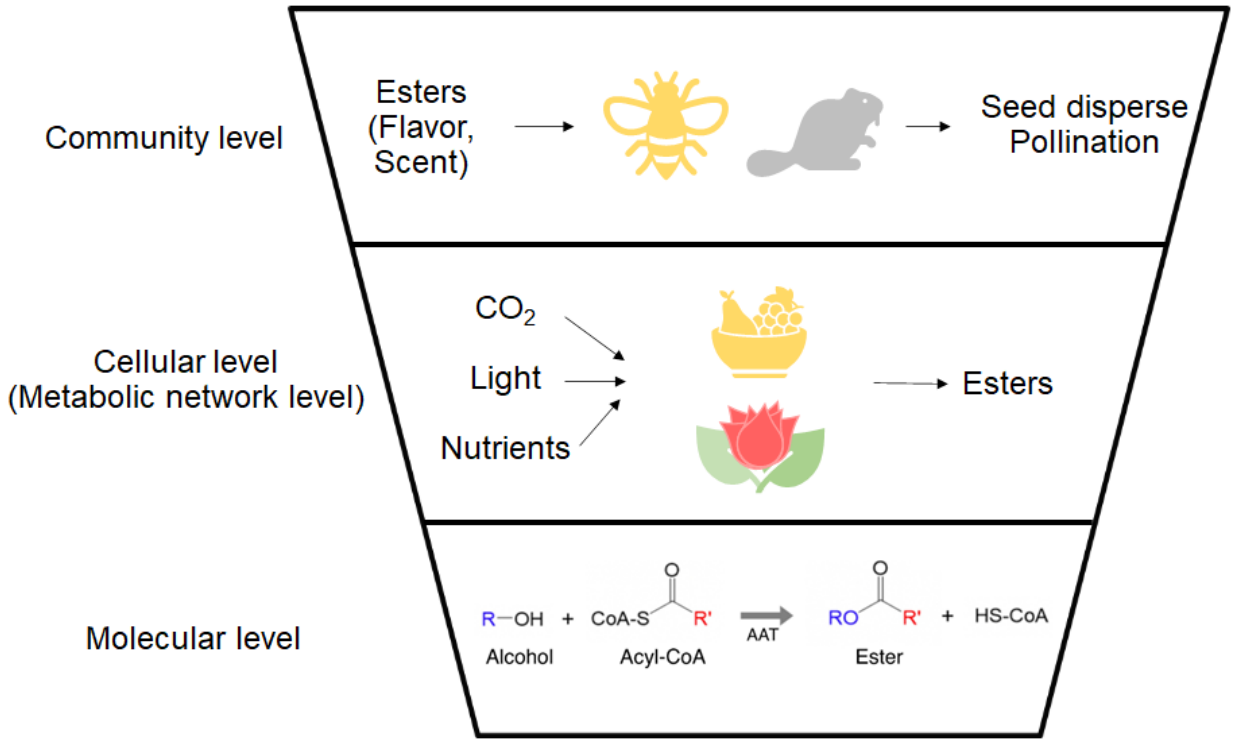


Figure 1-1: Ester formulation in nature and its modularity across different scales.

For efficient microbial ester biosynthesis, cellular metabolism should be rewired to overproduce a precursor alcohol and acyl-CoA. Alcohols longer than ethanol can be synthesized through the two recursive elongation pathways[52-54] (Figure 1-2A). In nature, the +1 recursive elongation pathway synthesizes 2-ketoisocaproate, the precursor of leucine, from 2-ketoisovalerate by LeuABCD operon. The elongation begins with the condensation of acetyl-CoA to 2-ketoisovalerate catalyzed by 2-isopropylmalate synthase (IPMS, LeuA). Although the wild type IPMS in *E. coli* is not capable of synthesizing long chain ketoacids, engineering the IPMS can improved substrate specificity towards ketoacids longer than 2-ketoisovalerate, activating the +1 recursive elongation pathway[54, 55]. The activated elongation module can be employed for longer chain alcohol synthesis upon decarboxylation and reduction by ketoacid decarboxylase and alcohol dehydrogenase[55]. The synthetic elongation pathway can produce ketoacids up to 2-ketooctanoate (C8) that can be converted into heptanol, a C7 chain length alcohol (Figure 1-2A).

Another alcohol production module is reverse β-oxidation cycle which is reversal of fatty acid oxidation metabolism. Since β-oxidation cycle is governed by *fad* and *ato* regulons activated by the presence of fatty acids[56], constitutive expression of the pathway is necessary to exploit the pathway for long-chain alcohol synthesis[52]. Here, specificity of alcohol dehydrogenase is critical for a controllable alcohol production through the reverse β-oxidation cycle because the cycle iteration is not regulated for a specific carbon length[52]. The alcohol production pathway modules can be exploited for ester biosynthesis upon enzymatic esterification by a compatible AAT. The reverse β-oxidation cycle necessitates synthesis of various length of acyl-CoAs, another key precursor for ester biosynthesis (Figure 1-2B). A well-studied bioprocess employing the acyl-CoA synthesis module is production of polyhydroxyalkanoate (PHA), which can serve as biodegradable plastic[57]. In the PHA production pathway, a partial reverse β-oxidation cycle synthesizes

acetoacetyl-CoA, enoyl-CoA, and 3-hydroxyacyl-CoA those can be subsequently polymerized into various PHAs having different physical properties of the bioplastic[57]. Because the +2 recursive elongation module can produce both alcohol and acyl-CoA, a programmable ester biosynthesis module will be critical to achieve a desirable metabolic balance. The modular alcohol and acyl-CoA pathways suggest that a wide variety of esters can be synthesized by harnessing the alcohol and acyl-CoA pathway modules and a compatible ester biosynthesis module.

The current ester biosynthesis system, however, encounters several engineering challenges such as limited application to various bacterial hosts and ester products. Particularly, most of the eukaryotic AATs characterized to date for ester biosynthesis exhibit low modularity due to limited host and product compatibility, and poor thermostability. For instance, a well characterized ATF1 from *S. cerevisiae* exhibits no activity against longer acyl-CoAs and therefore, products are confined to acetate esters. The limited range of ester biosynthesis necessitates laborious bioprospecting of the AATs for individual target ester production. Also, the eukaryotic AATs exhibit low thermostability, confining the ester production in temperatures lower than 42℃, although ester production at elevated temperatures can benefit downstream product separation and lower the overall production costs. Thus, engineering modularity of ester biosynthesis will enable rapid development of novel cell factories for scalable designer ester production.

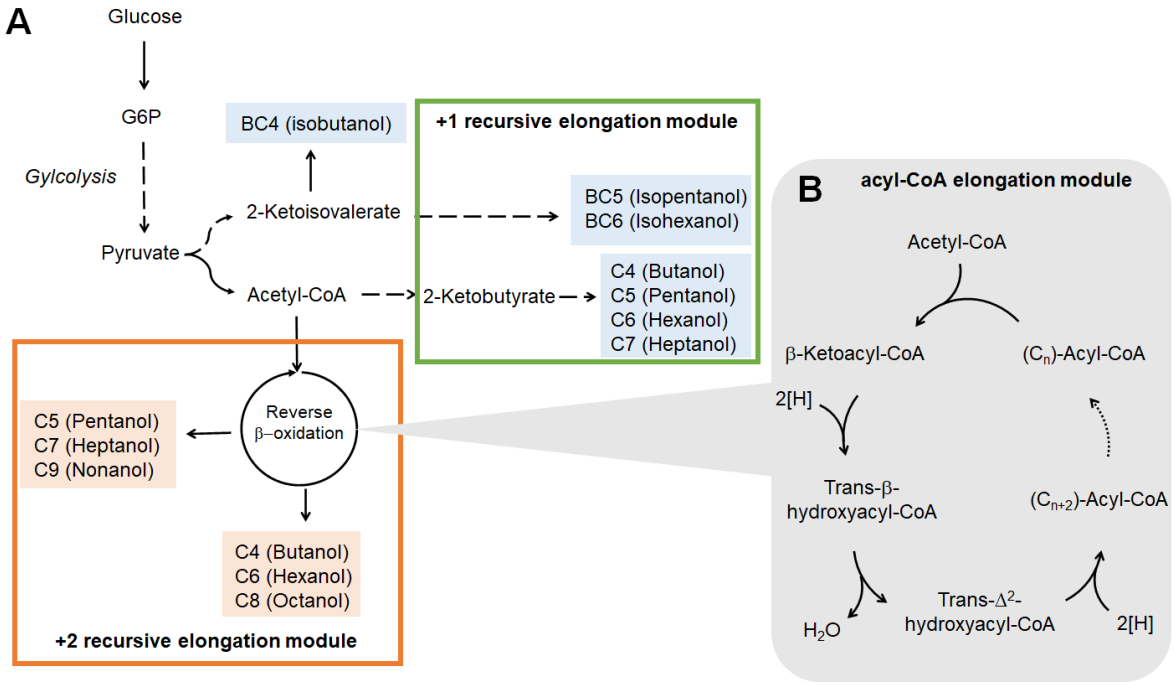


Figure 1-2: Representative alcohol and acyl-CoA biosynthesis pathways. The pathways can be subsidized as modules due to the modularity.

1-5 Outline of dissertation

This dissertation is my efforts to understand biological systems by engineering its modularity across scales. The focus of this dissertation is to demonstrate rapid development of designer ester production platforms by engineering modularity of biological systems at different scales (Figure 1-3).

There are six chapters in this dissertation. The Introduction chapter discusses engineering modularity in biotechnology for commodity chemical production. To demonstrate the concept, ester biosynthesis modules are created with high modularity exhibiting high compatibility, efficiency, and robustness at different scales throughout the following chapters. Chapter 2 is focused on engineering molecular level modularity of ester biosynthesis, and chapter 3 demonstrates engineering metabolic network level modularity for a model organism, a mesophilic gram-negative bacterium *Escherichia coli*. Chapter 4 expands the concept to a non-model organism, a thermophilic and cellulolytic gram-positive bacterium *Clostridium thermocellum*. Chapter 5 demonstrates engineering of modularity at a community level by designing *E. coli-E. coli* syntrophic co-culture systems. Finally, Chapter 6 discusses the significance of this dissertation. Most of the contents are slightly modified versions of published or submitted works in collaboration with Jong-Won Lee, Sergio Garcia, Preston Nicely, Noah Dunlap, Gillian Castro, Heidi Schindel, Richard Giannone, Adam Guss, Priyanka Singh, and others. They have been reproduced here with appropriate citation and/or the permission of the copyright holder.

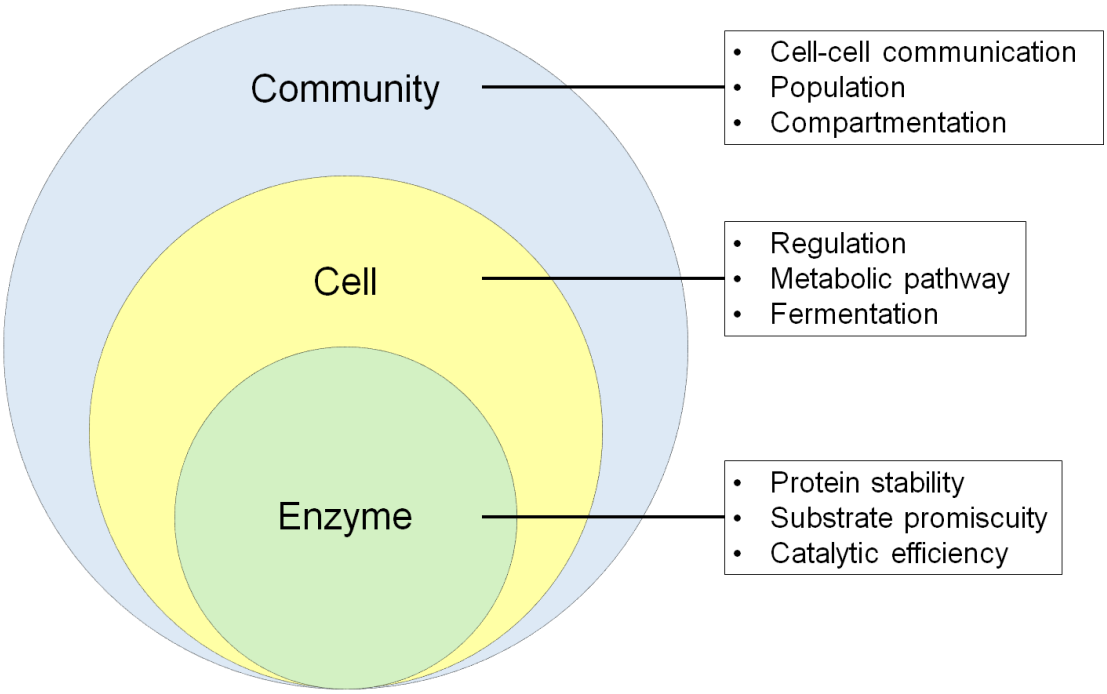


Figure 1-3: Engineering modularity of ester biosynthesis across the three biological scales.

Chapter 2. Engineering molecular level modularity of ester biosynthesis

This chapter focuses on engineering a robust and efficient AAT through bioprospecting and rational design of a promiscuous and thermostable chloramphenicol O-acetyltransferase (CAT, EC 2.3.1.28). Some of the results and discussion were slightly modified from published papers[58, 59].

At the molecular level, combinatorial interaction between alcohols and acyl-CoAs facilitated by an AAT can synthesize large space of various esters. The scientific question is: what is an "ideal" AAT to engineer modularity of ester biosynthesis across biological scales? I hypothesized that an ideal AAT having high compatibility, robustness, and efficiency can be generated by engineering promiscuity of the CAT.

2.1 Motivation

Enzymes are essential modules that facilitate interchangeable design of metabolic pathways to perform desirable functions across microbial hosts. Critical to the optimal design of heterologous pathways is the availability of robust and efficient enzymes that are compatible with various endogenous pathways and microbial hosts to enable combinatorial biosynthesis of desirable molecules. Engineering microbial systems to produce a diverse class of esters requires the broad enzymatic robustness, efficiency, and compatibility.

Industrially, short to medium chain (C_4-C_{14}) volatile esters have versatile applications as flavors[60], fragrances[50], solvents[61], and fuels[62, 63]. In nature, esters are formulated by an AAT that condenses an alcohol and an acyl-CoA in a thermodynamically favorable reaction, providing flavors and fragrances in ripening fruits and fermenting yeasts and having an ecological role in

pollination[64]. Inspired by nature, most of the metabolic engineering and synthetic biology strategies have deployed eukaryotic AATs originating from plants or yeasts for microbial biosynthesis of target esters[50, 61, 65]. However, these eukaryotic AATs lack sufficient robustness, efficiency, and compatibility for an interchangeable catalytic module as they commonly exhibit poor enzyme expression[66], solubility[67], and thermostability[68] in microbes, thus limiting optimal microbial production of esters. Although some AATs have been modified[69] or codon-optimized[50] to improve the host compatibility, efficient expression of these AATs has remained problematic in the heterologous hosts such as *Escherichia coli* likely due to different protein expression machineries and cellular environments[67]. In addition, limited knowledge on substrate profiles and specificities of AATs often requires laborious bioprospecting of AATs for individual target esters[50, 61, 63, 65, 66]. Therefore, developing robust and efficient AATs compatible with multiple pathways and microbial hosts is critical to expand biological routes for designer ester biosynthesis.

Chloramphenicol O-acetyltransferase (CAT, EC 2.3.1.28) is an antibiotic resistant enzyme that acetylates and detoxifies chloramphenicol, a potent drug that inhibits protein elongation[70]. Organisms become resistant to this drug by expressing CATs that recruit acetyl-CoA(s) as an acetyl donor to inactivate chloramphenicol. In nature, the *cat* gene is one of the most widespread genetic elements[71], expressing a functional enzyme in many organisms including plants[72], animals[73], and bacteria[74]. Interestingly, when being used as antibiotic selection in a recombinant *E. coli*, some CATs exhibited substrate promiscuity resulting in unexpected production of some esters including perillyl acetate[75] and isobutyl isobutyrate[50]. However, utility of CATs as an enzyme module in developing whole-cell biocatalysts for biosynthesis of designer esters was marginally studied due to low product yield and titer. It has been suggested that high enzyme stability and promiscuous activities help drive enzyme evolvability[76-78] to uncover novel functions

for *de novo* pathway engineering[79]. Thus, stability and promiscuity of CATs can be potentially exploited for an efficient ester biosynthesis platform. In this study, we identified and engineered promiscuity of CATs to function as robust and efficient AATs via screening of enzyme diversity and model-guided protein engineering to develop novel biomanufacturing microbial platforms for designer ester biosynthesis.

2.2 Materials and methods

2.2.1 Bacterial strains and plasmids

E. coli BL21(DE3) was used for protein expression and purification. Plasmids used in this study were listed in Table 2-1.

2.2.2 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich (MO, USA) and/or Thermo Fisher Scientific (MA, USA), unless specified elsewhere. For molecular cloning, restriction enzymes and T4 ligase were obtained from New England Biolabs (MA, USA). Phusion Hot Start II DNA polymerase was used for polymerase chain reaction (PCR).

2.2.3 Media and cultivation

For molecular cloning and protein expression, *E. coli* strains were grown in lysogeny broth (LB) containing appropriate antibiotics unless noted otherwise. For *in vivo* characterization of CATsa in *E. coli*, M9 hybrid medium[80] with 20 g/L glucose was used. Optical density (OD) was measured

Table 2-1: A list of plasmids and *E. coli* strains used in this study.

Name	Descriptions	Source
Plasmids		
pETDuet-1	pBR322 ori, Amp^R, lacI, T7lac promoter	Novagen
pET29a	pBR322 ori, Kan^R, lacI, T7lac promoter	Novagen
pET_CATsa	CATsa wild type encoding gene between BamHI, SacI site, pETDuet-1 backbone, 6X His-tag at N-terminus	This study
E. coli		
BL21(DE3)	F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)	Novagen
Top10	Host for molecular cloning, mcrA, Δ(mrr-hsdRMS-mcrBC), Phi80lacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(SmR), endA1, nupG	Invitrogen

by a spectrophotometer at 600 nm wavelength (Spectronic 200+, Thermo Fisher Scientific, MA, USA).

2.2.4 Multiple sequence alignment analysis

Multiple sequence alignment (MSA) analysis was performed using MEGA7[81]. Protein sequences were aligned by ClustalW[82] and visualized by ESPript 3.0 (http://espript.ibcp.fr)[83]. The key features in protein structures of 3U9F[84], 4CLA[85], and 2XAT[86] were extracted from CAT1_ECOLIX, CAT3_ECOLIX, and CAT4_PSEAE, respectively.

2.2.5 Molecular modeling and docking simulations

Swiss-Model[87] and the 'Builder' tools of the commercial software MOE (Molecular Operating Environment software, version 2019.01) were used to generate the three-dimensional (3D) structures of CATs, alcohols, acyl-CoAs, and their complexes, as described in *Seo et al*.[58]. To perform docking simulations in MOE, potential binding pocket was searched using the 'Site Finder' tool and selecting the best-scored site that is consistent with the reported catalytic sites [88]. Next, docking simulations were performed for acyl-CoA and alcohol with CATs using the induced fit protocol with the Triangle Matcher placement method and the London ΔG scoring function. The best-scored binding pose exhibiting the interaction between the residue and the substrate at root-mean-square-deviation (RMSD) < 2.3 Å was selected. To identify the potential residue candidates for mutagenesis of the acyl-CoA-alcohol-CAT complex, 'alanine scan' and 'residue scan' tools of MOE were used based on the ΔStability and/or ΔAffinity values calculated. Mutant candidates with small values of the ΔStability and/or ΔAffinity were chosen for experimental testing. To perform the protein contact analysis, 'Protein Contacts' tool of MOE was used.

2.2.6. Molecular cloning

Plasmids were constructed by the standard molecular cloning technique of ligase dependent method and/or Gibson assembly[89] using the primers listed in Table 2-2. The constructed plasmids were introduced into *E. coli* TOP10 by heat shock transformation. Colonies isolated on a selective plate were PCR screened and plasmid purified. The purified plasmids were verified via Sanger sequencing before being transformed into *E. coli* BL21 (DE3). Site-directed mutagenesis was performed using the QuickChange™ site-directed mutagenesis protocol with reduced overlap length[90] or Gibson assembly method[89].

2.2.7 *In vivo* characterization of CATsa and its variants in *E. coli*

For *in vivo* characterization of CATsa and its variants in *E. coli*, high-cell density cultures were performed with an addition of 2 g/L of various alcohols[91]. For *in-situ* extraction of esters, each tube was overlaid with 25% (v/v) hexadecane. To confirm the protein expression of CATsa and its variants, 1% (v/v) of stock cells were grown overnight at 37°C and 200 rpm in 15 mL culture tubes containing 5 mL of LB media and antibiotics. Then, 4% (v/v) of the overnight cultures were transferred into 1 mL of LB media containing antibiotics in a 24-well microplate. The cultures were grown at 37°C and 350 rpm using an incubating microplate shaker (Fisher Scientific, PA, USA) until OD reached to 0.4~0.6 and then induced by 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 hours with a Breathe-Easy Sealing Membrane to prevent evaporation and cross contamination (cat# 50-550-304, Research Products International Corp., IL, USA). The protein samples were obtained using the B-PER complete reagent (cat# 89822, Thermo Scientific, MA, USA), according to the manufacturer's instruction and analyzed by SDS-PAGE.

Table 2-2: A list of primers used in this study. The bold and underlined letters indicate restriction and site-directed mutation sites, respectively.

Primers	Primer sequence (5' to 3')
Site-saturation mutagenesis of CATsa F97	
CATsa_F BamHI	CTCT**GGATCC**AATGAACTTTAATAAAATTGATTTAG
CATsa_R SacI	CTCT**GAGCTC**TTATAAAAGCCAGTCATTAGGCCTA
F97R_F	TGATGGTGTATCTAAAACACGTTCTGGTATTTGGACTC
F97R_R	GTCCAAATACCAGAACGTGTTTTAGATACAC
F97K_F	TGATGGTGTATCTAAAACAAAATCTGGTATTTGGACTC
F97K_R	GTCCAAATACCAGATTTTGTTTTAGATACAC
F97L_F	TGATGGTGTATCTAAAACACTGTCTGGTATTTGGACTC
F97L_R	GTCCAAATACCAGACAGTGTTTTAGATACAC
F97W_F	TGATGGTGTATCTAAAACATGGTCTGGTATTTGGACTC
F97W_R	GTCCAAATACCAGACCATGTTTTAGATACAC
F97I_F	TGATGGTGTATCTAAAACAATTTCTGGTATTTGGACTC
F97I_R	GTCCAAATACCAGAAATTGTTTTAGATACAC
F97G_F	TGATGGTGTATCTAAAACAGGTTCTGGTATTTGGACTC
F97G_R	GTCCAAATACCAGAACCTGTTTTAGATACAC
CATsa_F	TGTTTAACTTTAAGAAGGAGATATACCAT
BB_CATsa_R	TGGATCCTGGCTGTGG
F97A_R	CATTCTTTACAGGAGTCCAAATACCAGAGGCTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97A_F	GCCTCTGGTATTTGGACTCCTGTA
F97N_R	CATTCTTTACAGGAGTCCAAATACCAGAATTTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97N_F	AATTCTGGTATTTGGACTCCTGTA
F97D_R	CATTCTTTACAGGAGTCCAAATACCAGAGTCTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97D_F	GACTCTGGTATTTGGACTCCTGTA
F97C_R	CATTCTTTACAGGAGTCCAAATACCAGAACATGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97C_F	TGTTCTGGTATTTGGACTCCTGTA
F97Q_R	CATTCTTTACAGGAGTCCAAATACCAGATTGTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97Q_F	CAATCTGGTATTTGGACTCCTGTA
F97E_R	CATTCTTTACAGGAGTCCAAATACCAGACTCTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97E_F	GAGTCTGGTATTTGGACTCCTGTA
F97H_R	CATTCTTTACAGGAGTCCAAATACCAGAGTGTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97H_F	CACTCTGGTATTTGGACTCCTGTA
F97M_R	CATTCTTTACAGGAGTCCAAATACCAGACATTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97M_F	ATGTCTGGTATTTGGACTCCTGTA
F97P_R	CATTCTTTACAGGAGTCCAAATACCAGATGGTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97P_F	CCATCTGGTATTTGGACTCCTGTA
F97S_R	CATTCTTTACAGGAGTCCAAATACCAGAGCTTGTTTTAGATACACC ATCAAAAATTGTAT

Table 2-2 Continued.

Primers	Primer sequence (5' to 3')
BB_F97S_F	AGCTCTGGTATTTGGACTCCTGTA
F97T_R	CATTCTTTACAGGAGTCCAAATACCAGACGTTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97T_F	ACGTCTGGTATTTGGACTCCTGTA
F97Y_R	CATTCTTTACAGGAGTCCAAATACCAGAGACTGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97Y_F	GTCTCTGGTATTTGGACTCCTGTA
F97V_R	CATTCTTTACAGGAGTCCAAATACCAGAGTATGTTTTAGATACACC ATCAAAAATTGTAT
BB_F97V_F	TACTCTGGTATTTGGACTCCTGTA
Y20W F	TTAATCAT TGG TTGAACCAACAAACGAC
Y20W R	GTTCAA CCA ATGATTAAATATCTCTTTTCTC
Y20H F	TTAATCAT CAT TTGAACCAACAAACGAC
Y20H R	GTTCAA ATG ATGATTAAATATCTCTTTTCTC
G194K F	GTTTGTGAT AAA TATCATGCAGGATTG
G194K R	TGCATGATA TTT ATCACAAACAGAATGATG
Y20D F	TTAATCAT GAT TTGAACCAACAAACGAC
Y20D R	GTTCAA ATC ATGATTAAATATCTCTTTTCTC
T88Y F	ACTTTAT TAT ATTTTTGATGGTGTATC
T88Y R	CAAAAAT ATA ATAAAGTGGCTCTAAC
T88D F	ACTTTAT GAT ATTTTTGATGGTGTATC
T88D R	CAAAAAT ATC ATAAAGTGGCTCTAAC
T88F F	ACTTTAT TTT ATTTTTGATGGTGTATC
T88F R	CAAAAAT AAA ATAAAGTGGCTCTAAC
Y20T F	TTAATCAT ACC TTGAACCAACAAACGAC
Y20T R	GTTCAA GGT ATGATTAAATATCTCTTTTCTC
Y20F F	TTAATCAT TTT TTGAACCAACAAACGAC
Y20F R	GTTCAA AAA ATGATTAAATATCTCTTTTCTC
G194R F	TGTGAT CGT TATCATGCAGGATTG
G194R R	TGATAA ACG TCACAAACAGAATGATGTA
T88L F	ACTTTAT CTG ATTTTTGATGGTGTATC
T88L R	CAAAAAT CAG ATAAAGTGGCTCTAAC
G194Y F	TGTGAT TAT TATCATGCAGGATTG
G194Y R	TGATAA ATA TCACAAACAGAATGATGTA
G194V F	TGTGAT GTT TATCATGCAGGATTG
G194V R	TGATAA AAC TCACAAACAGAATGATGTA
Y20L F	TTAATCAT CTG TTGAACCAACAAACGAC
Y20L R	GTTCAA CAG ATGATTAAATATCTCTTTTCTC
T88M F	ACTTTAT ATG ATTTTTGATGGTGTATC
T88M R	CAAAAAT CAT ATAAAGTGGCTCTAAC
L116R F	TATAC CGT TCTGATGTAGAGAAATA
L116R R	ATCAGA ACG GTATAAATCATAAAAC
L116K F	TATAC AAA TCTGATGTAGAGAAATA
L116K R	ATCAGA TTT GTATAAATCATAAAAC
F49Y F	CATTTAT TAT CTTAGTGACAAGGGTG
F49Y R	CTAAG ATA ATAAATGCAGGGTAAAATT
F49W F	CATTTAT TGG CTTAGTGACAAGGGTG

Table 2-2 Continued.

Primers	Primer sequence (5' to 3')
F49W R	CTAAG CCA ATAAATGCAGGGTAAAATT
G92R F	TTTGAT CGT GTATCTAAAACATTCTCTGG
G92R R	GATACA ACG TCAAAAATTGTATAAAGTGG
K48R F	GCATT CGT TTTCTTAGTGACAAGGGTGA
K48R R	AGAAA ACG AATGCAGGGTAAAATTTATATCC
Y112K F	AGTTT AAA GATTTATACCTTTCTG
Y112K R	AATC TTT AAACTCTTTGAAGTCAT
K48I F	GCATT AAT TTTCTTAGTGACAAGGGTGA
K48I R	AGAAA ATT AATGCAGGGTAAAATTTATATCC
T88I F	ACTTTAT ATT ATTTTTGATGGTGTATC
T88I R	CAAAAAT AAT ATAAAGTGGCTCTAAC
T88R F	ACTTTAT CGT ATTTTTGATGGTGTATC
T88R R	CAAAAAT ACG ATAAAGTGGCTCTAAC
S140L F	CTTTT CTG CTTTCTATTATTCCATGGAC
S140L R	AGAAAG CAG AAAAGCATTTTCAGGTATAGG
A138T F	GAAAAT ACC TTTTCTCTTTCTATTATTC
A138T R	AGAAAA GGT ATTTTCAGGTATAGGTG
CATec3 mutagenesis	
ec3 F97W F	AGACG TGG AGCGCGTTATCGTGCC
ec3 F97W R	GCGCT CCA CGTCTCTGTTTCCTGATG
ec3 Y20F F	AATTT TTT CGCCACAGACTGCCAT
ec3 Y20F R	TGGCG AAA AAATTCGAAATGTTCG
CATec1 mutagenesis	
ec1 F102W F	AAACC TGG AGCTCATTATGGAGTGAA
ec1 F102W R	GAGCT CCA GGTTTCCGTTTGTTC

2.2.8 Enzyme characterization

For enzyme expression, an overnight culture was inoculated with a 1:50 ratio in fresh LB medium containing 1 mM IPTG and antibiotics, followed by 18ºC overnight incubation (up to 20 hours) in a shaking incubator at 200 rpm. The induced cells were harvested by centrifugation at 4℃, and 4,700 x g for 10 minutes. The cell pellet was then washed once with Millipore water and resuspended in the B-PER complete reagent. After 30 min incubation at room temperature, the mixture was centrifuged at 17,000 x g for 2 minutes. The supernatant was collected and designated as crude extract. For his-tag purification, the crude extract was incubated with HisPur Ni-NTA superflow agarose in a batch as the manufacturer recommends. Then, the resin was washed with at least three volumes of wash buffer, consisting of 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 10 mM imidazole, and 0.1 mM EDTA. The resin bound proteins were eluted by 300 μL elution buffer containing 50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 300 mM imidazole, and 0.1 mM EDTA. The eluted sample was then desalted and concentrated via an Amicon filter column with 10 kDa molecular weight cutoff. Finally, the protein sample was suspended in 200 μL of 20 mM Tris-HCl buffer (pH 8.0). Protein concentration was measured by the Bradford assay[92] with bovine serum albumin (BSA) as the reference protein.

To measure protein melting temperature (Tm), a thermofluor assay was employed with SYPRO Orange[93]. About 10 to 250 μg of His-tag purified protein was mixed with 5x SYPRO Orange in a 50 μL final volume in a 96-well qPCR plate. The plate was sealed with PCR caps before running the assay. The StepOne real-time PCR machine (Applied Biosystems, CA, USA) was used to run the assay with the following parameters: ROX reporter, 1℃ increment per cycle, one-minute hold at every cycle, and temperature range from 20℃ to 98℃. The data were collected, exported, and manually processed to calculate Tm.

Reaction rate for each CAT was determined by a DTNB assay[94] in a 384-well plate. Total reaction volume was 50 μL with the reaction buffer comprising of 50 mM Tris-HCl (pH 8.0). Concentrations of acetyl-CoA (CoALA Biosciences, TX, USA) and alcohols were varied as specified in each experiment. Final enzyme concentrations of 0.05 μg/mL and 10 μg/mL were used for the reactions towards chloramphenicol and alcohols, respectively. Reaction kinetics were collected by measuring absorbance at 412 nm every minute for one hour at 50℃ in a microplate reader (Synergy HTX microplate reader, BioTek). The reaction rate was calculated using the extinction coefficient from a standard curve of free coenzyme A (MP Biomedicals, OH, USA) under the same condition. It should be noted that since the maximum operating temperature recommended for the plate reader is 50℃, the high throughput enzyme assay for CAT at elevated temperature was only performed to determine enzyme kinetics parameters.

The parameters of Michaelis-Menten rate law (eqn. 1) were calculated for each enzyme as follows. First, linear regression was performed on data collected from a microplate reader to identify initial reaction rates, y_i, at different initial substrate concentrations, s_i, where $i=$ [95n] is the number of data points collected. Then, these initial reaction rates and associated initial substrate concentrations for all replicates were simultaneously fit to the Michaelis-Menten model (eqn. 1) using robust non-linear regression (eqn. 2) with a soft-L1-loss estimator (eqn. 3) as implemented in the SciPy numerical computing library v1.2.0[96, 97].

$$v_i = \frac{v_{max} s_i}{K_M + s_i} \quad [1]$$

$$min_{k_m, v_{max}} \sum_{i=1}^{n} \rho((v_i(s_i, K_M, v_{max}) - y_i)^2) \quad [2]$$

$$\rho(z) = 2\left(\sqrt{1 + z}\right) - 1 \quad [3]$$

The least squares problem determines the parameters K_M and v_{max} by minimizing the difference between the model predicted reaction rates v_i and measured reaction rates y_i (eqn. 2). A

smoothing function $\rho(z)$ is used to make the least square problem resistant to outliers (eqn. 3). Due to the unbiased resistance to outliers and the avoidance of errors resulting from conventional linearization methods, robust non-linear regression provides the most precise parameter estimate for the Michaelis-Menten model[98].

2.3 Developing a robust alcohol acyltransferase for designer ester biosynthesis

2.3.1 Chloramphenicol acetyltransferase is a potential candidate for a robust alcohol acyltransferase

To develop a robust microbial ester production platform, a stable AAT is required. Unfortunately, the AATs known to date are isolated from mesophilic yeasts or plants, and none of them has been reported to be active at a temperature above 50℃. To explore any of the eukaryotic AATs are thermostable, we screened 12 AATs by measuring activity after heat incubation. The results showed that all the AATs were vulnerable to heat and lost the activity after one hour of incubation at 50℃, agreeing that the AATs have poor thermostability (Table 2-3). Recent studies inferred that another class of acetyltransferase, chloramphenicol acetyltransferase (CAT), might have the AAT activity producing short chain esters[50]. CAT is one of the widespread proteins implying a broad host compatibility. For example, the CAT from *Staphylococcus aureus* (CATsa) has also been used in genetic engineering of thermophilic microbes as a marker[99, 100], which suggest the enzyme can serve as an alternative thermostable AAT (Table 2-3, Figure 2-1).

We confirmed substrate promiscuity of CATsa by measuring conversion of various alcohols into acetate esters when feeding the alcohols to an *E. coli* overexpressing CATsa (Figure 2-2A).

Therefore, we chose CATsa to investigate their potential functions as a thermostable AAT because some thermostable CATs have been successfully used as a selection marker in thermophiles[100-104] and others have been shown to perform the acetylation for not only chloramphenicol but various alcohols like AATs[75, 105-107]. As a proof-of-study, we investigated CATsa, classified as Type A-9, from the plasmid pNW33N for a broad range of alcohol substrates as it has been widely used for genetic engineering in *C. thermocellum* at elevated temperature (≥ 50°C)[100-102].

We first conducted alcohol docking simulations using the homology model (Figure 2-2B). Remarkably, the model predicted binding affinities of short-to-medium chain length alcohols (e.g., ethanol, propanol, isopropanol, butanol, and isobutanol) and aromatic alcohols (e.g., benzyl alcohol and phenethyl alcohol) to the binding pocket. The change in the protein's Gibbs free energy upon the substrate binding was ordered as follows: 2-phenethyl alcohol > benzyl alcohol > isobutanol > butanol > propanol > ethanol > isopropanol (Figure 2-2A).

To quickly evaluate the *in silico* docking simulation results experimentally, we next performed *in vivo* characterization of a CATsa-overexpressing *E. coli* strain and screened for acetate esters production. Acetyl-CoA was derived from glycolysis while various alcohols were externally supplied to the media. Remarkably, the results exhibited the same trends of specificities of CATsa towards alcohols as predicted by the *in silico* docking simulation (Figure 2-2A). The CATsa-overexpressing *E. coli* produced all the expected acetate esters including ethyl acetate, propyl acetate, isopropyl acetate, butyl acetate, isobutyl acetate, benzyl acetate, and 2-phenethyl acetate at titers of 1.12 ± 0.07, 2.30 ± 0.28, 0.08 ± 0.02, 9.75 ± 1.57, 17.06 ± 6.04, 152.44 ± 29.50, and 955.27 ± 69.50 mg/L and specific ester production rates of 0.02 ± 0.00, 0.05 ± 0.01, 0.00 ± 0.00, 0.19 ± 0.03, 0.34 ± 0.12, 3.02 ± 0.57, and 19.27 ± 1.32 mg/gDCW/h, respectively. We observed that the specific ester production titers and rates are higher for aromatic alcohols than

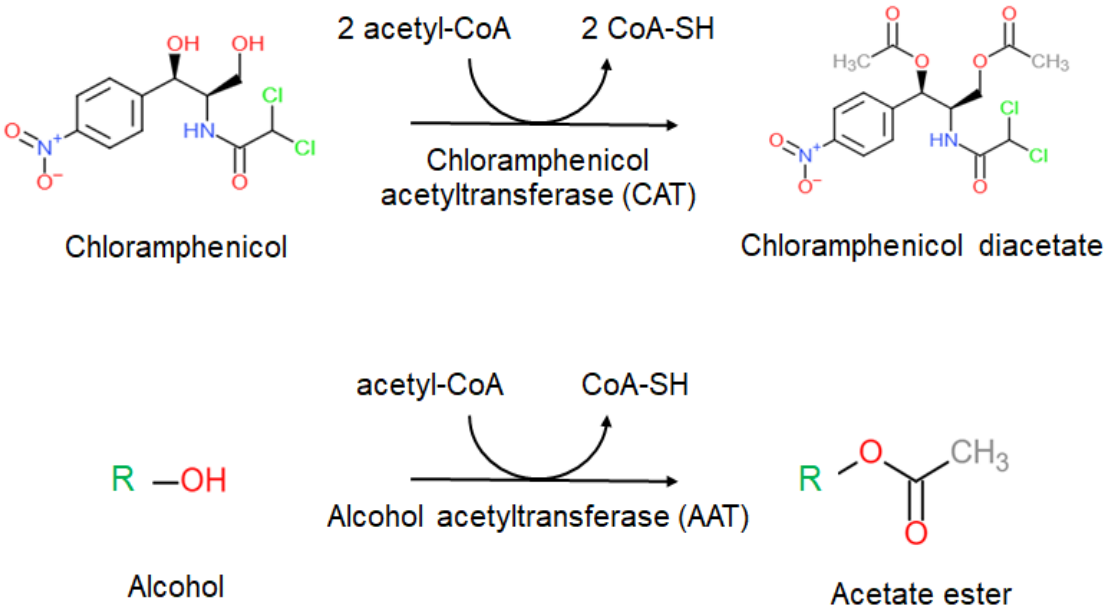


Figure 2-1: Acetylation by chloramphenicol acetyltransferase and alcohol acetyltransferase.

linear, short-chain alcohols likely because the hydrophobic binding pocket of CATsa has been evolved towards chloramphenicol [108], an aromatic antibiotic (Figure 2-2A). Specifically, the bulky binding pocket of CATsa likely contributes to more interaction with the aromatic substrates than the short, linear-chain alcohols.

Overall, thermostable CATs, (e.g., CATsa) can have broad range of substrate promiscuity towards linear, short-chain, and aromatic alcohols and hence can be harnessed as AATs for novel ester biosynthesis at elevated temperatures.

2.3.2 Rational design of CATsa to improve catalytic activity towards isobutanol and thermostability

Since the *in vivo* activity of CATsa is more than 50-fold higher for the aromatic alcohols than isobutanol, we asked whether its activity could be improved for isobutyl acetate biosynthesis. Using the *in silico* analysis, we started by examining whether any modification of the binding pocket of CATsa could improve the activity towards isobutanol. According to the homology model, the binding pocket consists of Tyr-20, Phe-27, Tyr-50, Thr-88, Ile-89, Phe-90, Phe-97, Ser-140, Leu-141, Ser-142, Ile-143, Ile-144, Pro-145, Trp-146, Phe-152, Leu-154, Ile-166, Ile-167, Thr-168, His-189, Asp-193, Gly-194, and Tyr-195, where the His189 and Asp193 are the catalytic sites (Figure 2-3A). Since chloramphenicol resistance is likely a strong selective pressure throughout evolution, we expected all CATs to exhibit a common binding pocket structure. Unsurprisingly, conserved sequences in the binding pocket were observed by protein sequence alignment of CATsa with other CATs of Type A. Especially, Pro-85 and Phe-97 were highly conserved in CATs of not only Type-A but also Type-B (Figure 2-3B).

Table 2-3: Bioprospecting of AATs by heat incubation at 50℃ for 1 hour. -: negative, +: positive

AATs	Origin	Activity after 1hr incubation at 50℃	Alcohol preference
ATF1	*Saccharomyces cerevisiae*	-	Heptanol, Isobutanol
KmATF	*Kluyveromyces marxianus*	-	Ethanol
CsAAT1	*Citrus sinensis* (Orange)	-	N.A.
CmAAT1	*Cucumis melo* (Muskmelon)	-	(E)-2-hexen-1-ol
CmAAT3	*Cucumis melo*	-	Benzyl Alcohol
CmAAT4	*Cucumis melo*	-	Benzyl Alcohol
CcAAT1	*Citrus clementina*	-	N.A.
CcAAT2	*Citrus clementina*	-	N.A.
VpAAT	*Vasconcellea pubescens*	-	Benzyl Alcohol
BanAAT	*Musa acuminate*	-	Cinnamyl Alcohol
MdAAT	*Golden delicious*	-	2-methyl butanol
CAT	pNW33N (*Staphylococcus sp.*)	+	Chloramphenicol

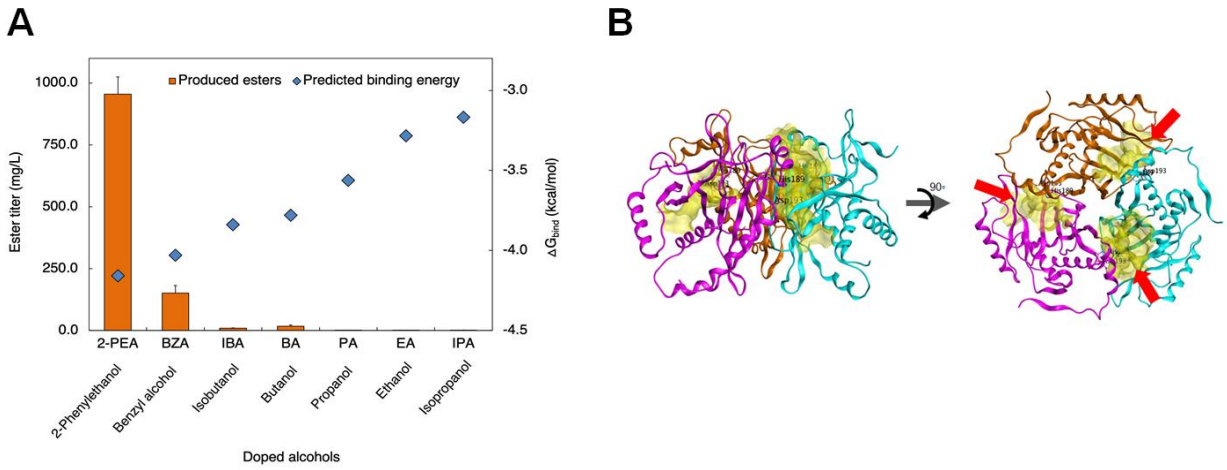


Figure 2-2: Broad substrate specificity of CATsa. **(A)** Comparison of predicated binding free energies of various alcohols to the binding pocket of CATsa and the titer of esters produced in *E. coli* carrying CATsa with an exogenous alcohol doping. **(B)** Structure of the CATsa homology model. The red arrows indicate the binding pockets formulated by the trimeric structure.

Based on the binding pocket identified, we performed docking simulation with alanine scans using the acetyl-CoA-isobutanol-CATsa complex and calculated ΔStability upon isobutanol binding (Figure 2-3C). Remarkably, the simulation suggested that Phe-97 residue is a promising target for mutagenesis. This residue is involved in the formation of a tunnel-like binding pocket[108]. Motivated by the analysis, Phe-97 was chosen for site saturated mutagenesis, and the variants were screened in *E. coli* for isobutyl acetate production by external supply of isobutanol.

Among the variants characterized, the F97W variant exhibited the best performance (Figure 2-4A), with the similar protein expression levels in *E. coli* (Figure 2-4B). As compared to the wild-type, the F97W variant enhanced the isobutyl acetate production by 4-fold. Subsequent *in silico* analysis showed that the mutation created a CH-π interaction between the hydrogen of isobutanol and the indole ring of F97W (Figure 2-4C). The model also indicated no change in distance between the isobutanol and active site (His-189) in F97W. Therefore, the CH-π interaction is likely responsible for the improved activity of F97W variant towards isobutyl acetate biosynthesis.

Next, we checked whether the F97W mutation affected thermal stability of the enzyme. We overexpressed and purified both the wildtype CATsa and CATsa F97W variant (Figure 2-5A). The SDS-PAGE analysis confirmed the expression and purification of the enzymes by bands with the expected monomer size (25.8 kDa). Thermofluor assay revealed that the F97W variant slightly lowered the wildtype melting point from 72℃ to 68.3℃ (Figure 2-5B).

Table 2-4 shows the *in vitro* enzymatic activities of both the wildtype CATsa and CATsa F97W at 50℃. The turnover number (k_{cat}) of CATsa F97W was two times higher than that of the wildtype. The increased turnover number of CATsa F97W led to 1.9-fold increase in enzymatic efficiency (k_{cat}/K_M, 4.08 ± 0.62, 1/M/sec) while the mutation did not result in significant change

in K_M. The improved enzymatic efficiency of CATsa F97W agrees with the enhanced isobutanol production observed in the *in vivo* characterization using the CATsa-overexpressing *E. coli* (Figure 2-4A).

Based on the rigidity of the binding pocket, we originally presumed that mutagenesis on the binding pocket would result in activity loss towards chloramphenicol. Surprisingly, CATsa F97W retained the activity towards chloramphenicol (Table 2-4). The F97W mutation decreased k_{cat} but also lowered K_M, resulting in a compensation effect. Turnover number of CATsa (k_{cat}, 202.97 ± 3.36, 1/sec) was similar to the previously reported value by Kobayashi *et al.*[109], but K_M (0.28 ± 0.02, mM) was about 1.75-fold higher. The difference might attribute to the experimental condition and analysis performed. Kobayashi *et al.* used chloramphenicol in a range of 0.05-0.2 mM for the assay and the Lineweaver-Burk method for analysis, while we used a 0-1.0 mM range with a nonlinear regression analysis method. Interestingly, affinity towards acetyl-CoA was independent of the alcohol co-substrates (Table 2-5), suggesting that the alcohol affinity would be likely the main bottleneck for microbial production of isobutyl acetate.

We further hypothesized that the thermostable CATsa could be repurposed to function as an AAT with higher catalytic efficiency towards a broad range of substrates and microbes including thermophiles. To test this hypothesis, we further aimed to engineer CATsa for enhanced melting temperature and catalytic efficiency towards smaller chain alcohols different from the native substrate chloramphenicol. To guide CAT engineering, we created a protein homology model of CATsa[58] from the available protein structures[108]. From the constructed homology model, we calculated its stability and affinity scores with small chain alcohols and acetyl-CoA required for ester biosynthesis.

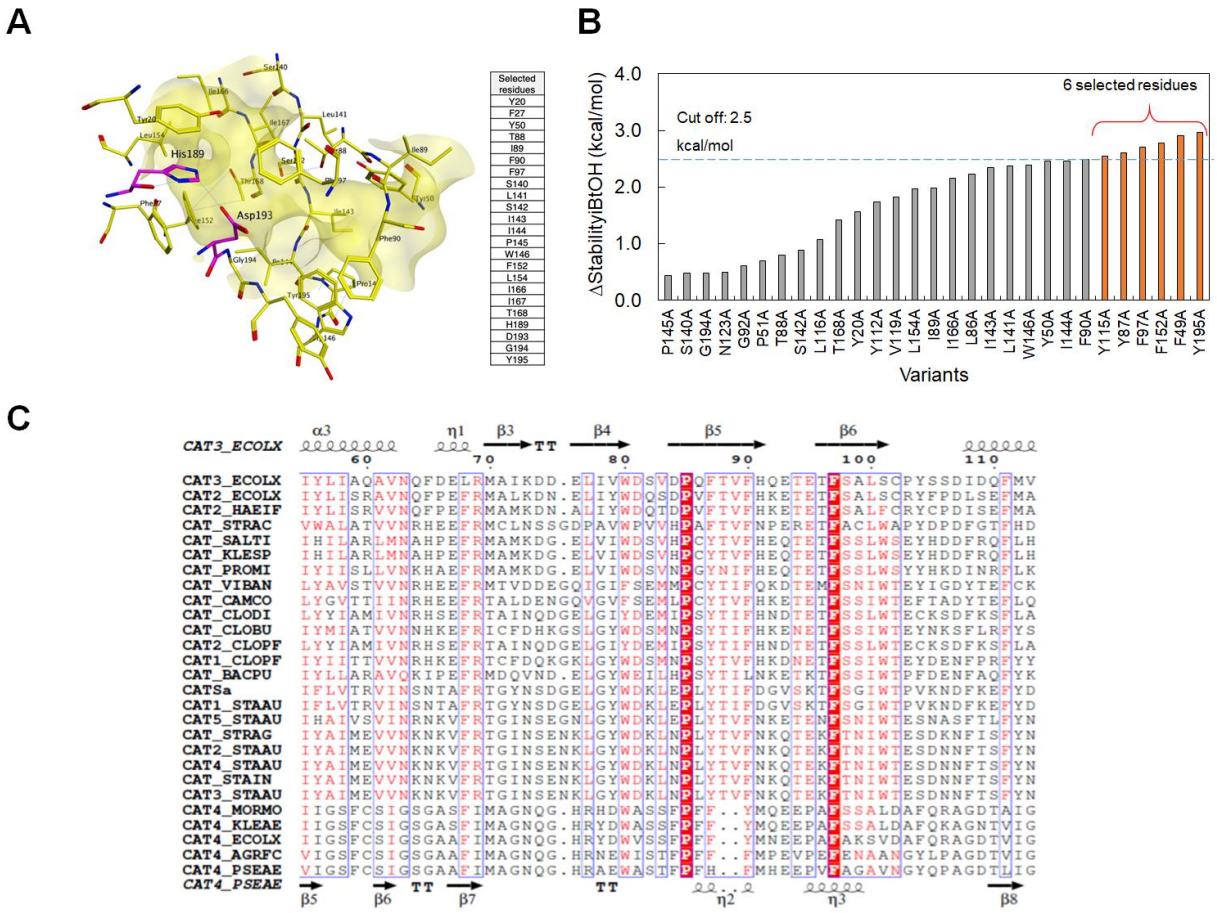


Figure 2-3: Docking simulation and consensus analysis of CATsa. **(A)** A binding pocket of CATsa and comprised amino acid residues. **(B)** Alanine scan of 23 identified residues in the binding pocket of the CATsa-isobutanol-acetyl-CoA complex. ΔStability (kcal/mol) represents the relative stability of each variant with respect to its wild type. Positive ΔStability values indicate that the structural stability of the protein-ligands complex is reduced if a specific residue is replaced with alanine, suggesting the significance of that specific residue in the protein-ligands complex. **(C)** Protein sequence alignment of CATsa with different CATs. The highly conserved regions, F97 and P85, are red highlighted.

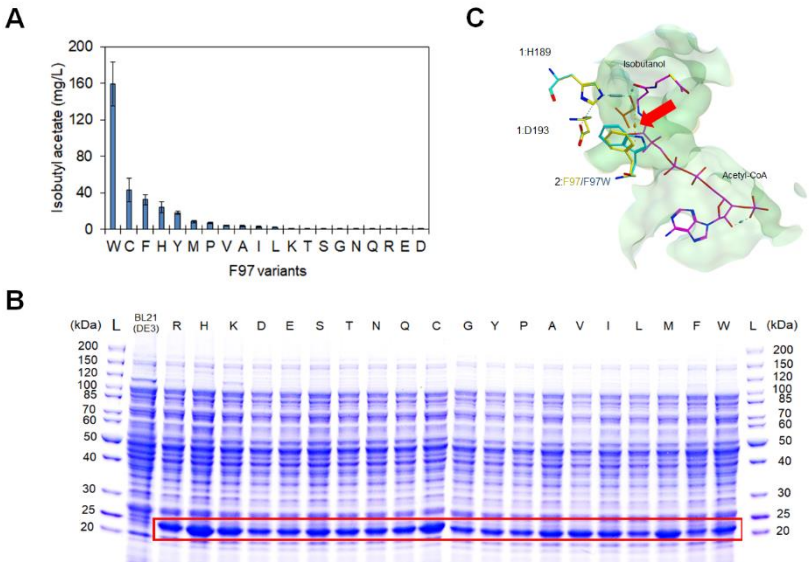


Figure 2-4: Mutagenesis study of CATsa to improve activity towards isobutanol. **(A)** SDS-PAGE analysis of *E. coli* crude extracts expressing CATsa F97 variants. The overexpressed CATsa F97 variants are shown in the red box. The alphabets annotate the amino acid variants. LD; protein ladder. **(B)** Screening of F97 site-saturated mutagenized variants for enhanced isobutyl acetate production in *E. coli*. The letters indicate amino acids substituting F in the wildtype CATsa. **(C)** Superposed binding pocket structure of the wildtype and CATsa F97W mutant. The red arrow indicates a CH-π interaction between the hydrogen of isobutanol and the indole ring of F97W.

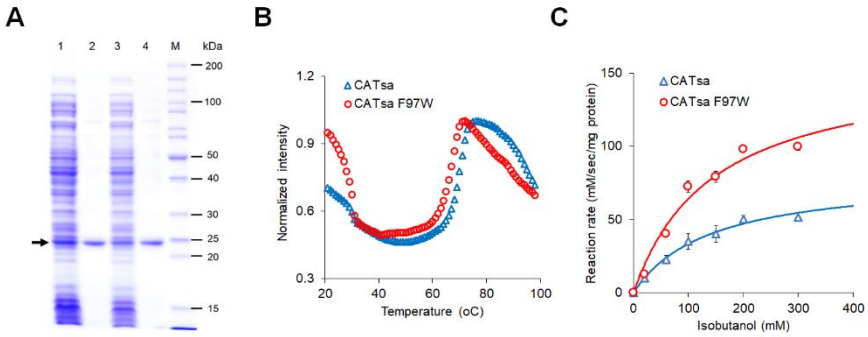


Figure 2-5: *In vitro* characterization of CATsa and CATsa F97W. **(A)** SDS-PAGE of purified CATsa WT and F97W. The black arrow indicates expected size of expressed target proteins, CATsa and CATsa F97W. 1; crude extract of IPTG induced *E. coli* BL21(DE3) harboring pET_CATsa, 2; His-tag purified CATsa, 3; crude extract of IPTG induced E. coli BL21(DE3) harboring pET_CATsa F97W, 4; His-tag purified CATsa F97W, M; protein ladder. **(B)** Melting curve of CATsa WT and F97W. The intensity was normalized by each maximum value. **(C)** Michaelis-Menten plots of CATsa wild type and F97W toward isobutanol at 50°C. 2 mM of acetyl-CoA was supplemented as the saturated cosubstrate. The error bars represent standard deviation of three biological replicates.

As a representative small alcohol, isobutanol was chosen for docking simulation and experimental characterization of CAT variants for thermostability and catalytic efficiency because this alcohol is much smaller and very different from chloramphenicol and can be efficiently produced by microorganisms[110]. Affinity and stability screening via *in silico* mutagenesis of selected CATsa binding pocket residues suggested some candidates potentially altering the isobutanol binding (Figure 2-6A, B). By experimentally constructing and characterizing the top 26 CATsa variants using an *in vivo* microbial screening assay[111], we identified the CATsa Y20F variant to exhibit the most significant increase in conversion of isobutanol to isobutyl acetate, about 43-fold higher than its wild-type enzyme (Table 2-6).

We next characterized robustness and efficiency of CATsa Y20F *in vitro* by measuring melting temperatures and k_{cat}/K_M values towards isobutanol. Since a previous study demonstrated that CATsa A138T increased thermostability[109], we also investigated whether the combinatorial mutagenesis CATsa variants enhanced the enzyme performance (Figure 2-6C). We found that CATsa Y20F improved the catalytic efficiency over the wild-type CATsa and CATsa F97W 5.0- and 2.5-fold, respectively, while the melting temperature slightly decreased from 71.2 to 69.3℃ (Figure 2-6D). Among the combinatorial mutagenesis, CATsa Y20F A138T exhibited both the highest melting temperature (76 ± 0.0℃) and catalytic efficiency towards isobutanol (10.3 ± 1.2, 1/M/s), respectively.

Unfortunately, the combinatorial mutation of F97W and Y20F did not further improve the catalytic efficiency and rather lowered the melting temperatures (Fig. 2-6D). It has been suggested that the hydrogen bonds between chloramphenicol, catalytic His-189, and Tyr-20 at the transition state are critical for high catalytic efficiency of the CAT, specifically by constraining the binding of chloramphenicol in the correct orientation[112].

Table 2-4: Kinetic parameters of the wildtype CATsa and mutant CATsa F97W.The reactions were performed at 50°C. The co-substrate, acetyl-CoA, was supplied at the saturated concentration of 2 mM. Melting temperature (Tm) of CATsa and CATsa F97W are 72.0 ± 0.8, and 68.3 ± 1.2°C, respectively.

	CATsa		**CATsa F97W**	
Substrates	Chloramphenicol	Isobutanol	Chloramphenicol	Isobutanol
K_M (mM)	0.28 ± 0.02	138.66 ± 28.92	0.18 ± 0.01	144.77 ± 23.65
k_{cat} (1/sec)	202.97 ± 3.36	0.30 ± 0.03	102.63 ± 2.04	0.59 ± 0.05
k_{cat}/K_M (1/M/sec)	$7.37 \pm 0.48 \times 10^5$	2.16 ± 0.45	$5.77 \pm 0.49 \times 10^5$	4.08 ± 0.62

Table 2-5: K_M values of CATsa and CATsa F97W towards acetyl-CoA.

	CATsa		CATsa F97W	
Co-substrates	Chloramphenicol	Isobutanol	Chloramphenicol	Isobutanol
K_M (mM)	0.08 ± 0.01	0.06 ± 0.01	0.09 ± 0.01	0.08 ± 0.02

Since the Y20F mutation retains the aromatic ring that contributes to tautomeric stabilization of His-189, the catalytic imidazole can still interact with the smaller chain alcohols stably and flexibly, likely contributing to the enhanced enzymatic activity observed for isobutyl acetate biosynthesis.

2.3.3 Bioprospecting CATs

The CAT is one of the most widespread enzymes in nature but only a few were discovered to have the AAT activity towards small alcohols. To explore whether high thermostability and alcohol promiscuity of CATs occurred in nature, we selected a library of 27 structurally distinctive type-A and type-B CATs[71] for characterization because influence of the protein structure differences on the small alcohol promiscuity was unknown. Since CATs are commonly found in various organisms with similar or identical protein sequences, we annotated the CAT origins based on the Uniprot and NCBI databases. We synthesized these CATs, expressed, purified, and characterized their melting temperatures and promiscuous activities towards isobutanol (Figure 2-7). Most of the CATs showed melting temperatures higher than 60℃ except CAT_GEO (T_m = 43.5℃) (Table 2-7). Among these CATs, CAT1_ECOLX (CATec1), CAT3_ECOLX (CATec3), CATsa, CAT_KLEPS (CATkl), CAT2_ECOLX (CATec2), CAT_HAEIF (CATha), and CAT_CLOBU (CATcb) exhibited the highest specific activities towards isobutanol at 50℃. Five out of these seven most isobutanol-active CATs were evolutionarily related, suggesting that their activities towards isobutanol might be influenced by their unique structural features. We further evaluated kinetic thermostability of these seven most isobutanol-active CATs by measuring their activity losses after one-hour incubation at elevated temperatures of 50, 55, 60, 65, and 70℃ (Figure 2-8A).

Table 2-6: Screening of CATsa variants for isobutyl ester synthesis in *E. coli*. Ethyl ester is a byproduct since *E. coli* produced ethanol endogenously. Each value represents mean ± 1 standard deviation (stdev) from at least three biological replicates.

CATsa variants	Ethyl acetate (mg/L)	Isobutyl acetate (mg/L)
WT	2.88 ± 0.24	4.16 ± 1.23
Y20W	2.12 ± 0.30	5.25 ± 3.51
Y20H	1.98 ± 0.45	3.75 ± 1.56
G194K	1.83 ± 0.36	1.33 ± 1.12
Y20D	1.98 ± 0.44	1.71 ± 1.44
T88Y	1.72 ± 0.17	1.72 ± 1.03
T88D	2.52 ± 0.83	1.81 ± 1.54
T88K	2.16 ± 0.66	1.14 ± 0.61
Y20T	1.92 ± 0.29	4.07 ± 0.95
Y20F	1.56 ± 0.15	181.80 ± 32.38
G194R	2.07 ± 0.59	4.09 ± 3.89
T88L	1.95 ± 0.16	0.58 ± 0.23
G194Y	2.06 ± 0.24	2.30 ± 0.69
G194V	2.03 ± 1.29	2.85 ± 2.47
Y20L	2.00 ± 0.45	12.88 ± 1.17
T88M	2.33 ± 0.83	3.52 ± 2.00
L116R	2.30 ± 0.75	7.99 ± 1.39
L116K	2.36 ± 0.11	12.62 ± 4.38
F49Y	2.13 ± 0.39	4.51 ± 2.92
F49W	2.28 ± 0.20	3.04 ± 2.60
G92R	2.75 ± 0.87	3.66 ± 2.50
K48R	2.20 ± 0.35	6.01 ± 0.86
Y112K	2.40 ± 0.52	2.85 ± 2.03
K48I	2.65 ± 1.36	9.11 ± 3.64
T88I	2.86 ± 0.72	3.43 ± 1.04
T88R	2.92 ± 1.23	4.54 ± 1.71
S140L	2.60 ± 1.08	4.46 ± 0.84

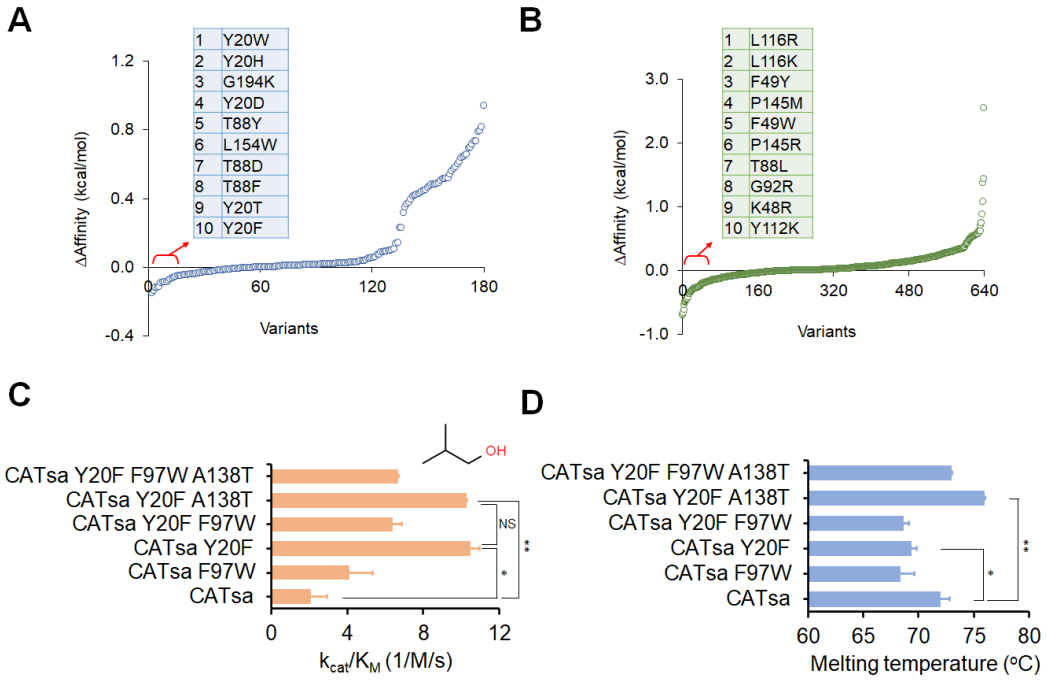


Figure 2-6: Protein engineering of CATsa to improve efficiency and robustness towards isobutanol. **(A)** Residue scan of binding pocket residues interacting with isobutanol of the CATSa-isobutanol-acetyl-CoA complex. Selected candidates are based on the ΔAffinity values. **(B)** Residue scan of residues interacting with acetyl-CoA of the CATsa-isobutanol-acetyl-CoA complex. Selected candidates are based on the ΔAffinity values. **(C)** Catalytic efficiencies of CATsa and its variants towards isobutanol. Student's t-test: *p-value = 1.20E-07 (df = 5), **p-value = 3.94E-04 (df = 8), NS = not significant (df = 6). **(D)** Melting temperatures of CATsa and its variants. Student's t-test: *p-value = 0.016 (df = 7), **p-value = 4.49E-05 (df = 8). In panels C-D, each value represents mean ± 1 stdev from at least three biological replicates. For each in vitro assay, 2 mM acetyl-CoA was used as a co-substrate.

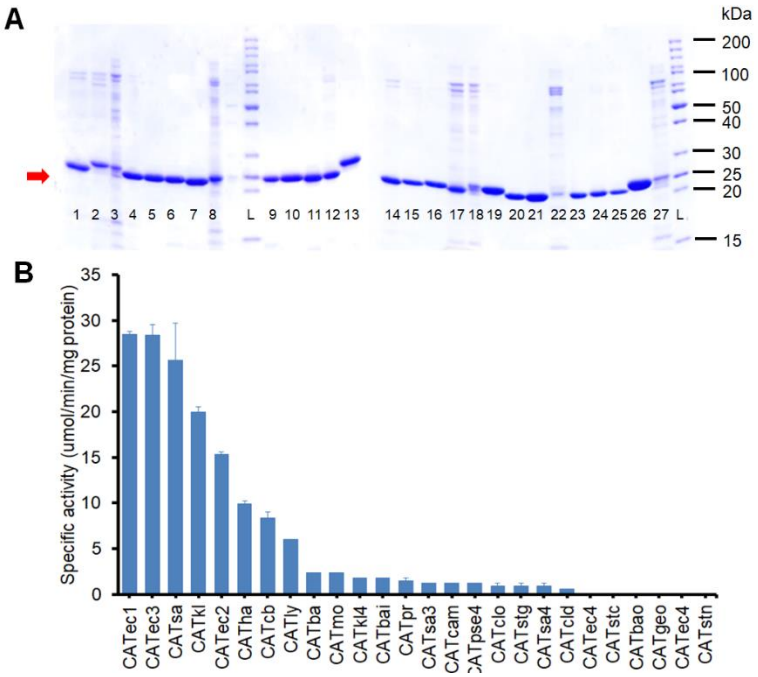


Figure 2-7: Screening of CATs for activity towards isobutanol. **(A)** SDS-PAGE of His-tag purified CATs. The protein samples were loaded at 500 ng per lane. From 1 to 27; CATsa, CAT_STRAC, CAT2_ECOLX, CAT_THEACI, CAT_BACI, CAT4_STAAU, CAT_BAOCI, CAT3_STAAU, CAT_PROMI, CAT_BACPU, CAT1_ECOLIX, CAT_KLEPS, CAT_CLOBU, CAT_GEO, CAT_CAMCO, CAT4_ECOLIX, CAT4_PSEAE, CAT_CLODI, CAT_LYSI, CAT4_MORMO, CAT1_CLOPF, CAT5_STAAU, CAT3_ECOLX, CAT4_KLEAE, CAT_STRAG, CAT2_HAEIF, and CAT_VIBAN. The red arrow indicates the average size of the CATs (24.9kDa). CAT_STAIN was not included because of unsuccessful expression. **(B)** Specific activity of CATs towards 100 mM isobutanol (mean ± 1 stdev from two biological replicates).

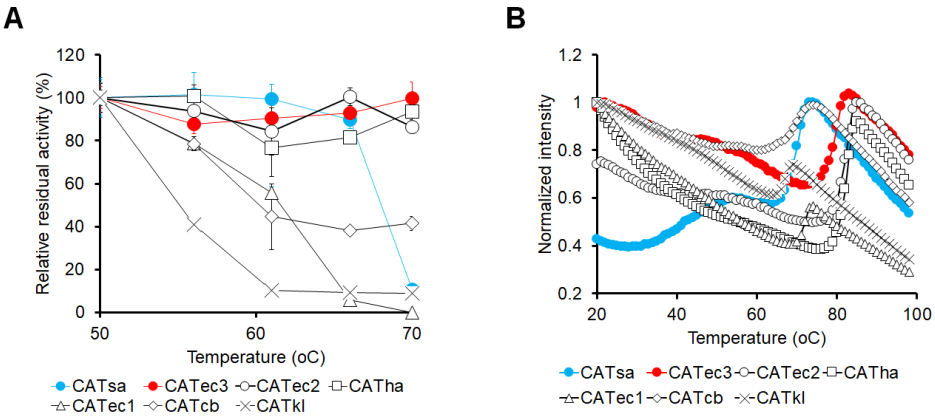


Figure 2-8: Thermostability of the seven selected CATs. **(A)** Kinetic thermostability of the selected CATs. Residual activity after the heat incubation at 50°C was used for the normalization. Each value represents mean ± 1 stdev from three biological replicates. **(B)** Melting curves of the seven selected CATs. The signal intensity was normalized by each maximum value.

Table 2-7: Melting temperatures of a set of 28 chloramphenicol acetyltransferases (CATs). Each value represents mean ± 1 stdev. Note that CAT1_ECOLX and CAT3_ECOLX are also denoted as CATec1 and CATec3, respectively.

No	Type	CATs	Origins	Uniprot (or NCBI) protein entry	T_m (°C)	Replicates
1	A	CATsa	*Staphylococcus aureus*	WP_001010387.1	71.2 ± 1.1	6
2	A	CAT_STRAC	*Streptomyces acrimycini*	P20074	77.5 ± 4.3	4
3	A	CAT2_ECOLX	*Escherichia coli*	P22615	81.6 ± 0.8	11
4	A	CAT_THEACI	*Thrmoactinomyces sp.*	WP_037996077.1	81.0 ± 0.0	2
5	A	CAT_BACI	*Bacillus sp.*	A0A3N9QXS5	70.7 ± 0.9	3
6	A	CAT4_STAAU	*Staphylococcus aureus*	P36882	71.0 ± 0.0	2
7	A	CAT_BAOCI	*Bacillus oceanisediminis*	A0A160M771	65.0 ± 0.0	3
8	A	CAT_CLOBU	*Clostridium butyricum*	Q02736	70.0 ± 0.0	3
9	A	CAT3_STAAU	*Staphylococcus aureus*	P06135	68.7 ± 2.4	3
10	A	CAT5_STAAU	*Staphylococcus aureus*	P36883	82.0 ± 0.0	2
11	A	CAT_BACPU	*Bacillus pumilus*	P00487	72.0 ± 0.0	2
12	A	CAT1_ECOLX	*Escherichia coli*	P62577	71.2 ± 0.7	6
13	A	CAT_KLEPS	*Klebsiella sp.*	P58777	67.7 ± 0.9	3
14	A	CAT_GEO	*Geomicrobium sp.*	WP_042419141.1	43.3 ± 0.8	4
15	B	CAT4_ECOLX	*Escherichia coli*	P26838	71.8 ± 1.3	12
16	B	CAT4_PSEAE	*Pseudomonas aeruginosa*	P26841	71.0 ± 1.6	4
17	B	CAT4_MORMO	*Morganella morganii*	P50869	57.0 ± 0.0	2
18	A	CAT_LYSI	*Lysinibacillus boronitolerans*	A0A0A3IEC4	76.0 ± 0.0	2
19	A	CAT_VIBAN	*Vibrio anguillarum*	P49417	71.5 ± 0.5	2
20	A	CAT1_CLOPF	*Clostridium perfringens*	P26825	65.0 ± 0.0	2
21	A	CAT_PROMI	*Proteus mirabilis*	P07641	81.0 ± 0.0	2
22	A	CAT3_ECOLX	*Escherichia coli*	P00484	80.2 ± 1.5	6
23	B	CAT4_KLEAE	*Klebsiella aerogenes*	P50868	70.5 ± 0.9	4
24	A	CAT_STRAG	*Streptococcus agalactiae*	Q03058	66.3 ± 5.8	4
25	A	CAT2_HAEIF	*Haemophilus influenzae*	P22616	82.0 ± 1.8	6
26	A	CAT_CAMCO	*Campylobacter coli*	P22782	69.0 ± 0.0	3
27	A	CAT_CLODI	*Clostridioides difficile*	P11504	n.a.	2
28	A	CAT_STAIN	*Staphylococcus intermedius*	P25309	n.a.	2

Remarkably, CATec3 and CATec2 could retain more than 95% of the activity at 70℃ while the residual activity of CATsa rapidly decreased. The high kinetic thermostability of CATec3 agrees with a previous study[113]. We also found that CATec1, previously employed to demonstrate some ester biosynthesis in *E. coli*[50], is less thermostable and efficient than CATec3 (Table 2-7). Particularly, CATec1 showed relatively poor kinetic thermostability although it showed similar thermodynamic stability as compared to CATsa (Figure 2-8B).

Overall, CATec3 was the most thermostable CAT with the highest promiscuous activity towards isobutanol (Figure 2-9A).

2.3.4 Engineering alcohol promiscuity of CATec3

Discovery of high thermostability and promiscuity of CATec3 towards the small alcohol (i.e., isobutanol) prompted us to investigate whether the beneficial mutations discovered in CATsa could be introduced into CATec3 to further improve its performance. We constructed CATec3 variants harboring single and combinatorial mutations of F97W and Y20F. Since the homolog residue of A138 does not exist in CATec3, its mutation was not constructed and characterized. Interestingly, the results showed that the CATec3 Y20F variant improved not only the catalytic efficiency (13.0 ± 0.2, 1/M/s), about 3.3-fold higher than wild-type, but also the melting temperature (87.5 ± 0.5℃) (Figure 2-9B, C). Among all the CATs characterized, CATec3 Y20F is the most thermostable and isobutanol-active (Table 2-8). Motivated by the improved catalytic efficiency introduced by the Y20F mutation towards various alcohols, we further examined whether CATec3 Y20F exhibited enhanced compatibility towards a library of 16 linear, branched, saturated, unsaturated, and aromatic alcohols with acetyl-CoA as a co-substrate (Table 2-9). We found that CATec3 Y20F was more efficient towards bulky and long-chain alcohols that are more

hydrophobic than short-chain alcohols, likely due to a stronger binding affinity. As compared to the wild-type CATec3 against a representative set of six alcohols that can be naturally synthesized by organisms, we found that the CATec3 Y20F variant exhibited higher catalytic efficiency towards butanol by 4.0-fold (Figure 2-10A), pentanol by 8.8-fold (Figure 2-10B), benzyl alcohol by 6.9-fold (Figure 2-10C), and phenylethyl alcohol by 6.2-fold (Figure 2-10D), respectively.

In addition, we found that CATec3 Y20F could catalyze acetylation of fatty alcohols (i.e., octanol and decanol) to produce long-chain wax esters for drop-in biodiesel applications (Figure 2-11A, B). Note that kinetic parameters were not reported for these fatty alcohols due to low substrate solubility. The alcohol compatibility of CATec3 Y20F expanded from ethanol to terpenoid alcohols such as geraniol and nerol. Due to high K_M value (>1M) towards ethanol (Figure 2-11C), CATec3 Y20F is more favorably applied for biosynthesis of higher-chain alcohol esters. This characteristic is potentially beneficial to produce designer esters rather than ethyl esters in organisms since ethanol is a common fermentative byproduct that can act as a competitive substrate. In comparison to CATsa Y20F A138T, CATec3 Y20F displayed higher activity towards not only isobutanol, but also most of other alcohols (Figure 2-12). It is noteworthy that these engineered CATs exhibited different alcohol specificities. For example, CATsa Y20F A138T was relatively more specific to phenylethyl alcohol than terpenoid alcohols as compared to CATec3 Y20F (Figure 2-12).

2.3.5 Acyl-CoA promiscuity of CATec3

We next examined whether CATec3 Y20F was also compatible with longer-chain acyl-CoAs.

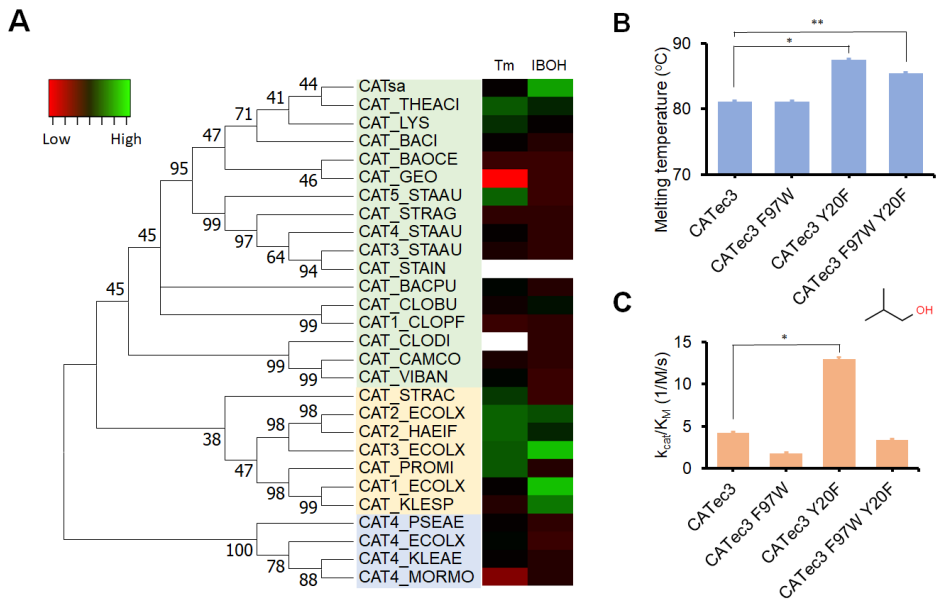


Figure 2-9: Bioprospecting and protein engineering of CATs. **(A)** A phylogenetic tree of 28 CATs and heat map of their melting temperatures (Tm) and activities towards isobutanol (IBOH). The numbers in the phylogenetic tree represent bootstrapping values (%) from 1000 bootstrap replicates. **(B)** Melting temperatures of CAT3_ECOLX (CATec3) variants. Student's t-test: *p-value = 4.34E-08 (df = 13), **p-value = 3.43E-06 (df = 13). Each value represents mean ± 1 stdev from at least three biological replicates. **(C)** Catalytic efficiencies of CATec3 and its variants towards isobutanol. Student's t-test: *p-value = 4.46E-05 (df = 6). Each value represents mean ± 1 stdev from at least three biological replicates.

Table 2-8: Melting temperatures and kinetic parameters of CAT variants towards isobutanol. The co-substrate, acetyl-CoA, was supplemented at the saturated concentration of 2 mM. Kinetic parameters were determined at 50°C. Each value represents mean ± 1 stdev from at least three biological replicates.

CATs	**Tm (°C)**	**k_{cat} (1/s)**	**K_M (mM)**	**k_{cat}/K_M (1/M/s)**
CATec1	71.2 ± 0.7	0.4 ± 0.1	375.9 ± 121.7	1.0 ± 0.4
CATec1 F102W	68.0 ± 0.0	1.2 ± 0.1	259.3 ± 59.2	4.7 ± 1.2
CATsa	71.2 ± 1.1	0.3 ± 0.05	164.1 ± 29.6	2.1 ± 0.4
CATsa F97W	68.3 ± 1.2	0.6 ± 0.05	144.8 ± 23.7	4.1 ± 0.7
CATsa Y20F	69.3 ± 0.5	1.5 ± 0.06	146.3 ± 13.8	10.5 ± 1.1
CATsa Y20F F97W	68.7 ± 0.5	1.4 ± 0.1	212.3 ± 32.4	6.4 ± 1.1
CATsa Y20F A138T	76.0 ± 1.2	2.0 ± 0.09	189.2 ± 19.8	10.3 ± 1.2
CATsa Y20F F97W A138T	73.0 ± 0.6	1.4 ± 0.06	205.7 ± 17.8	6.7 ± 0.6
CATec3	80.2 ± 1.5	0.7 ± 0.1	171.4 ± 24.6	4.2 ± 0.1
CATec3 F97W	81.2 ± 0.4	0.5 ±0.1	250.2 ± 84.3	1.8 ± 0.1
CATec3 Y20F	87.5 ± 0.5	2.3 ± 0.1	180.1 ± 21.7	13.0 ± 0.2
CATec3 Y20F F97W	85.5 ± 0.5	0.53 ± 0.1	156.7 ± 58.5	3.4 ± 0.1

Table 2-9: Catalytic efficiency of CATec3 Y20F towards multiple alcohol substrates. The catalytic efficiency was measured from the kinetic reactions performed at 50°C. The co-substrate, acetyl-CoA, was supplemented at the saturated concentration of 2 mM. The values represent average ± standard deviation from at least three biological replicates. *Calculation of the parameters against methanol was not statistically practical due to the low affinity. # 10% of DMSO was supplemented due to low substrate solubility.

Alcohol substrates	k_{cat} (1/s)	K_M (mM)	k_{cat}/K_M (1/M/s)
Ethanol*	0.8 ± 1.0	1232.5 ± 1986.3	0.6 ± 1.3
Butanol	1.9 ± 0.4	195.7 ± 87.3	10.5 ± 5.1
Isobutanol	2.3 ± 0.1	180.1 ± 21.7	13.0 ± 1.7
Prenol	2.7 ± 0.3	101.0 ± 21.6	26.4 ± 6.6
Furfuryl alcohol	1.1 ± 0.1	37.1 ± 9.8	29.1 ± 8.4
Pentanol	3.9 ± 1.0	64.7 ± 33.6	59.1 ± 19.3
Isoamyl alcohol	2.8 ± 0.3	59.9 ± 26.6	46.0 ± 11.9
Benzyl alcohol	1.6 ± 0.1	12.2 ± 1.6	130.3 ± 17.8
Hexanol	6.4 ± 0.4	29.2 ± 4.0	219.5 ± 33.5
3-cis-Hexen-1-ol	2.3 ± 0.2	6.4 ± 2.2	360.1 ± 129.4
Phenylethyl alcohol	6.1 ± 0.1	10.6 ± 0.4	577.1 ± 26.2
3-Methoxybenzyl alcohol[#]	18.8 ± 0.9	10.0 ± 1.2	1,891.6 ± 248.8
Geraniol[#]	8.0 ± 0.2	4.1 ± 0.4	1,975.7 ± 186.8
Citronellol[#]	7.7 ± 0.2	3.2 ± 0.3	2,400.9 ± 247.9
Nerol[#]	2.2 ± 0.04	4.1± 0.8	4,119.5 ± 753.0
Chloramphenicol	105.5 ± 6.9	0.2 ± 0.05	491,901.5 ± 115,286.2

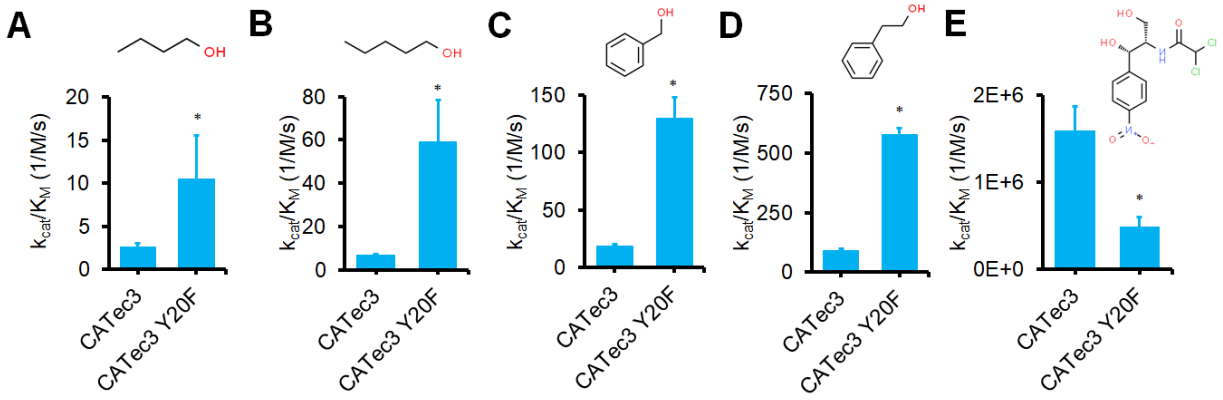


Figure 2-10: Comparison of catalytic efficiencies of CATec3 and CATec3 Y20F towards **(A)** butanol (Student's t-test: *p-value = 8.22E-04 (df = 5), **(B)** pentanol (Student's t-test: *p-value = 2.77E-06, df = 5), **(C)** benzyl alcohol (*p-value = 4.70E-03, df = 3), **(D)** 2-phenylethyl alcohol (Student's t-test: *p-value = 1.45E-04, df = 2, and **(E)** chloramphenicol (Student's t-test: *p-value = 1.22E-04, df = 5). Each value represents mean ± 1 stdev from at least three biological replicates.

We analyzed the relative activities of CATec3 Y20F against a set of 10 linear, branched, and aromatic acyl-CoAs that can be synthesized in organisms together with isobutanol as a co-substrate (Fig. 2-13A). CATec3 Y20F has the highest activity towards the native substrate acetyl-CoA, which is the most abundant and critical precursor metabolite for cell biosynthesis. As compared to acetyl-CoA, CATec3 Y20F achieved 46%, 28%, 15%, 12%, 11%, and 9% of activity towards isobutyl-CoA, propionyl-CoA, butyryl-CoA, valeryl-CoA, phenylethyl-CoA, and isovaleryl-CoA, respectively. No activity was detected against linear fatty acyl-CoAs longer than valeryl-CoA. Interestingly, CATec3 Y20F also exhibited activity towards an uncommon lactyl-CoA for lactate ester biosynthesis. Since lactyl-CoA is not commercially available for *in vitro* assay, we determined the activity *in vivo* by using a recombinant *E. coli* co-expressing CATec3 Y20F and a propionyl-CoA transferase (PCT) derived from different microbes including *Thermus thermophilus* (PCTtt) that transfers CoA from acetyl-CoA to lactate[114, 115] (Figure 2-13B). By co-feeding the recombinant *E. coli* with 2 g/L of isoamyl alcohol and lactate, isoamyl lactate could be produced at a level of about 66.6 mg/L (Figure 2-13C), which is at least 2.5-fold higher than the use of the eukaryotic AATs reported previously[61]. Since PCTtt is derived from a thermophile, the lactate ester biosynthesis pathway is likely thermostable and compatible with thermophilic hosts. Like eukaryotic AATs, CATs are also known to have weak thioesterase activities that hydrolyze acyl-CoAs; however, their turnover rates (k_{cat}) are significantly lower for thioesterases than for alcohol acetyltransferases[116]. Taken together, we have shown unequivocally that promiscuity of CATs can be engineered for *de novo* thermostable AATs for esterification of small alcohols. The engineered CATec3 Y20F likely has a higher modularity than other eukaryotic AATs because it is a robust and efficient ester biosynthesis module compatible with a wide range of alcohols and acyl-CoA substrates.

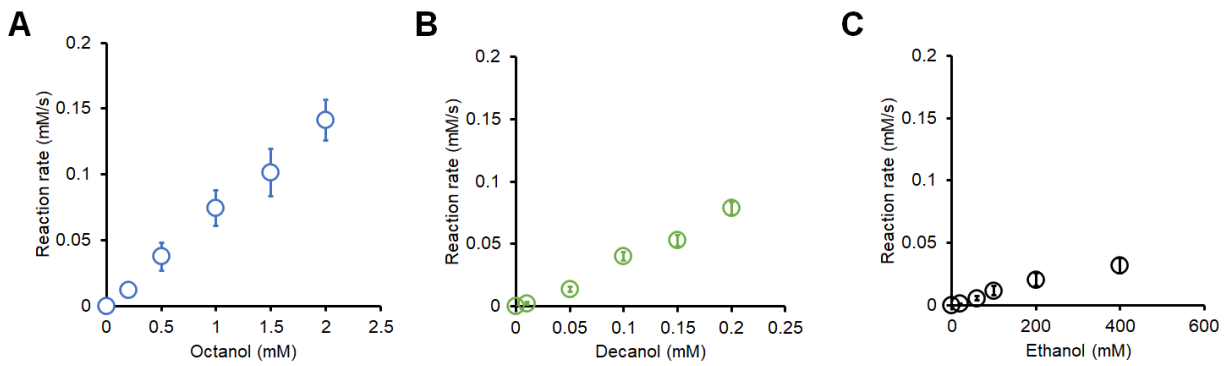


Figure 2-11: Michaelis-Menten plots of CATec3 Y20F towards octanol, decanol, and ethanol. **(A)** octanol, **(B)** decanol, and **(C)** ethanol. Each data represents mean ± 1 stdev from at least three biological replicates.

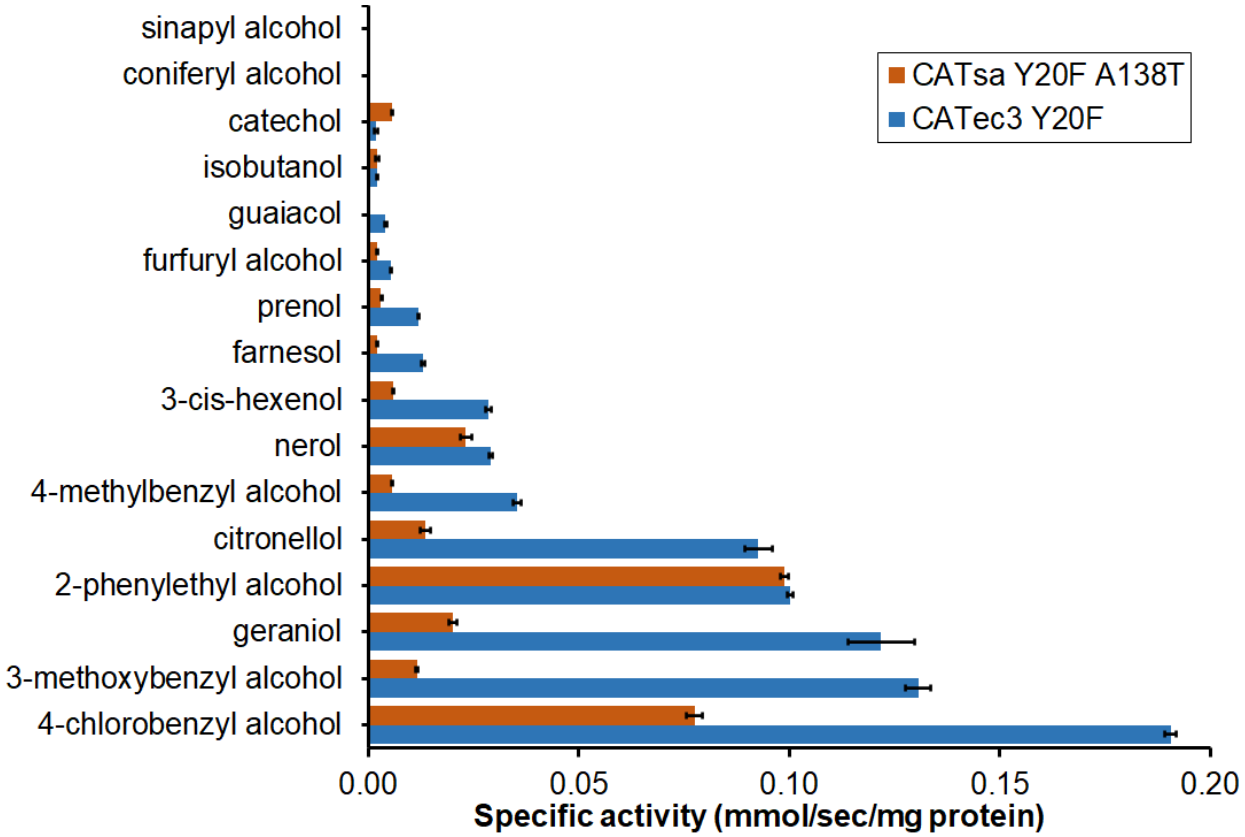


Figure 2-12: Specific activity of CATec3 Y20F and CATsa Y20F A138T towards various alcohols. The specific activity was measured at 20mM of alcohol substrates (mean ± 1stdev from three biological replicates).

2.4 Conclusion

Robust and efficient AATs are essential enzyme modules to expand biological routes for sustainable and renewable production of designer bioesters with useful applications. Although metabolic engineering and synthetic biology strategies expanded ester biosynthesis of short-chain volatile esters in well-studied microbes such as *E. coli*[50, 65, 66, 117], harnessing broader range of microbial metabolic networks for an efficient ester production has been largely limited due to the lack of compatible AATs known to have poor solubility and thermostability (e.g., even for ATF1 of *S. cerevisiae*, one of the best performing AATs known to date[118-120]). This chapter addressed the AAT problems by elucidating the promiscuity of CATs and exploiting this unique property to create novel biomanufacturing microbial platforms for designer ester biosynthesis.

Prior to our study, promiscuous activity of CATs was accidentally discovered from an unexpected production of esters by the engineered alcohol-producing recombinant *E. coli* strains [50, 75]; however, CATs were not efficient enough to convert alcohols to esters and hence utility of CATs as a catalyst was underestimated. Furthermore, fundamental understanding of CATs' enzymatic characteristics for designer ester biosynthesis was lacking. To fill this knowledge gap, we identified and engineered the promiscuity of CATs to function as robust and efficient AATs compatible with multiple alcohols and acyl-CoAs through screening of CAT diversity, model-guided protein engineering, and comprehensive enzyme characterization. A single mutation of Y20F, weakening the interaction between chloramphenicol, catalytic H189, and Y20 at the transition state, can enhance its promiscuity. In addition to discover the Y20F role, we engineered F97W and A138T for improving promiscuity of CATs towards smaller alcohols and thermostability, but the combinatorial mutations did not result in beneficial effects synergistically.

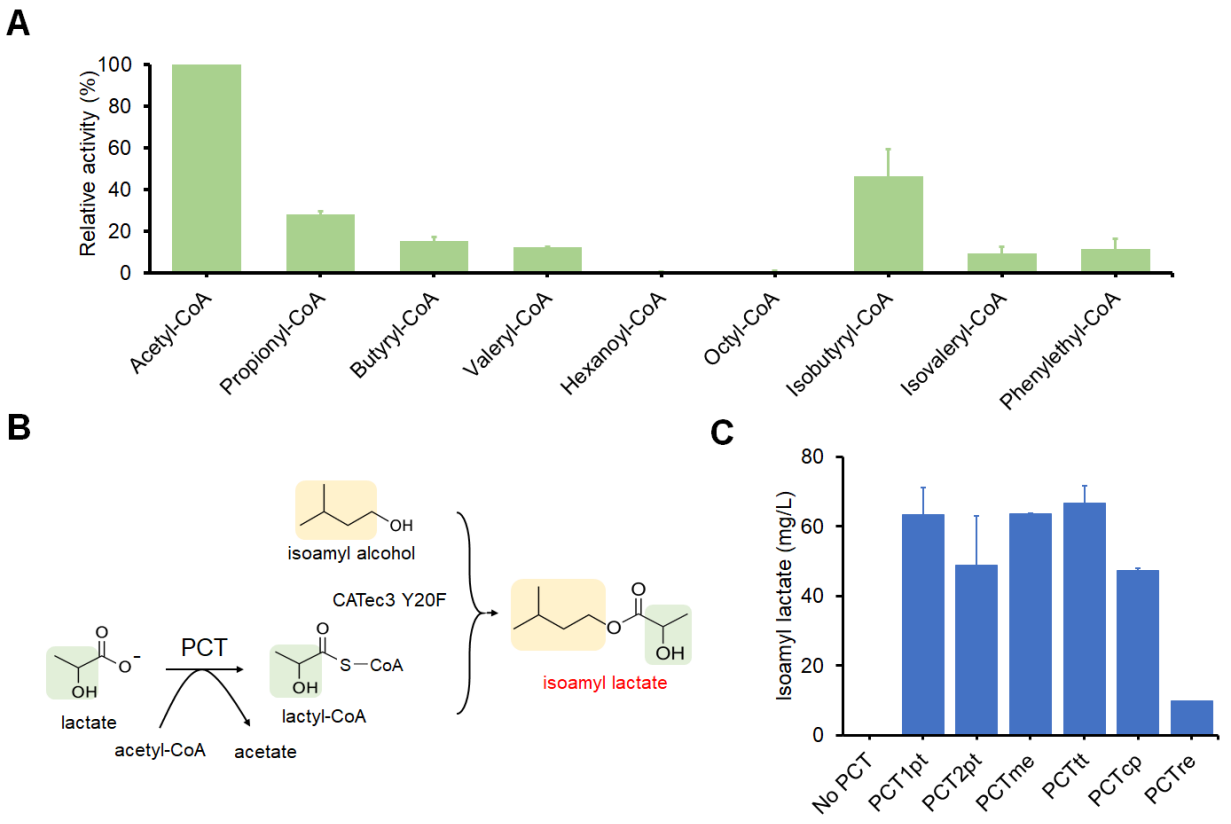


Figure 2-13: Substrate promiscuity of CATec3 Y20F towards acyl-CoAs. **(A)** Relative activity of CATec3 Y20F towards acyl-CoAs as compared to acetyl-CoA. 400mM of isobutanol was supplemented as the co-substrate. Each value represents mean ± 1 stdev from six biological replicates. **(B)** Microbial biosynthesis of isoamyl lactate by co-feeding the recombinant E. coli co-expressing CATec3 Y20F and propionate CoA transferases (PCTs) with isoamyl alcohol and lactate. **(C)** Effect of expressing various PCTs on isoamyl lactate production after 48 h. The initial M9 medium contained 10 g/L glucose, 5 g/L yeast extract, 2 g/L of isoamyl alcohol, 2 g/L lactate, and 0.1mM of IPTG to induce the protein expression. 1 mL of hexadecane was overlaid to extract the isoamyl lactate produced during the fermentation. Each value represents mean ± 1 stdev from three biological replicates. Origin of PCTs: PCTpt, *Pelotomaculum thermopropionicum*; PCTme, *Megasphera elsdenii*; PCTtt, *Thermus thermophilus*, PCTcp, *Clostridium propionicum*; PCTre, *Ralstonia eutropha*.

We found that CATsa Y20F A138T mutant achieved both the improved catalytic efficiency and thermostability while CATec3 Y20F was the best performer among the library of variants characterized. We believe that additional rounds of mutagenesis of the best enzyme CATec3 Y20F with a larger library screening could potentially identify other beneficial mutants that can further improve its catalytic efficiency towards a target substrate, which is an interesting future research topic. Overall, the repurposed CAT is a useful platform to synthesize designer bioesters in multiple microorganisms (Fig. 2-14). Theoretically, at least 168 esters can be synthesized by CATec3 Y20F in combination of the 21 alcohols and 8 acyl-CoAs. Since expressing the functional CAT does not require post-translational modifications and/or localization, a wider range of organisms can employ the repurposed CAT for ester production.

In summary, this study created a new direction for programmable ester production by repurposing CATs rather than bioprospecting AATs for designer ester biosynthesis. The CAT can serve a useful platform of the protein engineering towards a specific target alcohol and/or acyl-CoA. While this study presented a robust and efficient platform for diverse classes of ester biosynthesis, further optimization is required to control specificity. Relative short amino acid length and well identified protein structures of CATs as compared to other eukaryotic AATs are favorable for the protein engineering. The high stability and promiscuity of CATs also offer a better protein evolvability[76-78] than other AATs. Metabolic engineering strategies can also facilitate a specific ester production by controlling metabolic flux towards targeted alcohol and acyl-CoA precursors.

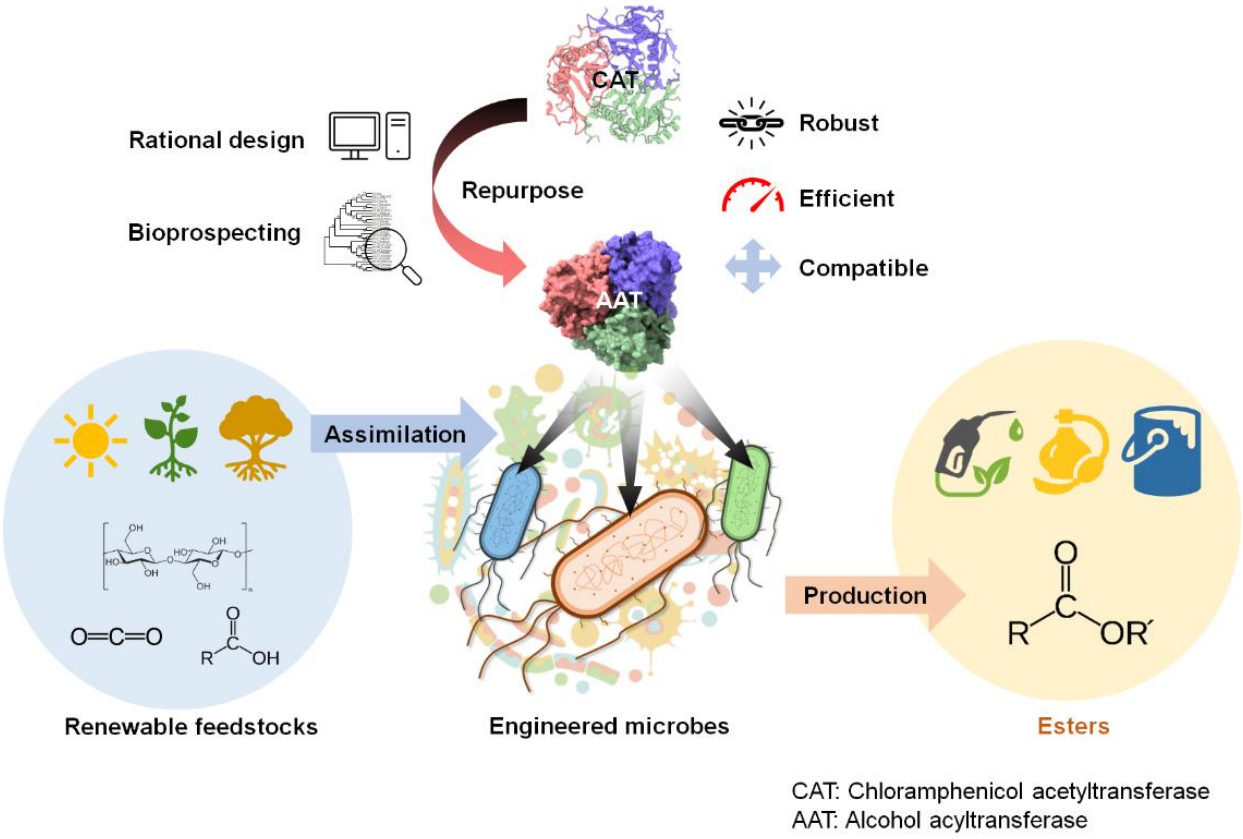


Figure 2-14: Repurposing chloramphenicol acetyltransferase (CAT) for an efficient and robust designer ester biosynthesis. The repurposed CAT can serve as a robust, efficient, and promiscuous AAT that is compatible with multiple microbial hosts, expanding biological routes for various ester production.

Chapter 3. Engineering metabolic network/cellular level modularity of ester biosynthesis in a model microorganism

This chapter focuses on metabolic engineering of *E. coli* to engineer modularity of ester biosynthesis at metabolic network/cellular level. The engineered CATec3 Y20F was expressed in *E. coli* for conversion of various alcohols using intracellular acetyl-CoA, demonstrating compatibility of the engineered ester biosynthesis module to the metabolic network. Further, engineered isoamyl acetate production was demonstrated by manipulating precursor pathway modules. Parts of the studies are under review for publication[59].

3.1 Motivation

Cellular system comprises thousands of different enzymes that build up a metabolic network. The metabolic pathways catalyze sequential metabolic reactions converting sugars into alcohols and acyl-CoAs, the two key precursors for ester biosynthesis. Therefore, expressing a functional AAT in a cell introduces a heterologous ester biosynthesis pathway on top of the endogenous metabolism.

Since the engineered CATec3 Y20F is robust and compatible with various substrates (Table 2-9), metabolic engineering strategies using the CATec3 Y20F can harness metabolic network of a microorganism to produce esters. At metabolic network/cellular level, ideal phenotypes are efficient production (e.g., titer, yield, productivity), robustness (e.g., cell growth, stable fermentation), and metabolic compatibility (e.g., metabolic balance, soluble protein expression).

Model organisms are relatively well characterized, and many genetic engineering tools and parts (e.g., plasmids, promoters, and terminators) are available, suitable as platform cell factories for proof-of-principle. *Escherichia coli* is one of the most well studied a facultative anaerobic and mesophilic gram-negative microorganism that has been largely deployed as a workhorse industrial host. Especially, many favorable characteristics (e.g., fast growth, broad sugar assimilation capability, trackable genetic elements) make the *E. coli* appropriate for engineering modularity of ester biosynthesis at cellular level.

3.2 Materials and methods

3.2.1 Bacterial strains and plasmids

E. coli BL21(DE3) was used as a host strain for alcohol conversion and fermentative ester production. The strains and plasmids used in this chapter were listed in Table 3-1.

3.2.2 Media and cultivation

E. coli strains were grown in lysogeny broth (LB) medium or M9 hybrid medium[80] containing glucose as a carbon source and 5 g/L yeast extract supplemented with 100 μg/mL ampicillin and/or 50 μg/mL kanamycin when appropriate.

3.2.3 Recombineering

E. coli gene deletions were carried out using recombineering[121]. A temperature sensitive low-copy plasmid that contains exo, bet and gam in their native phage operon, pL, under λ CI repressor

Table 3-1: A list of plasmids and strains used in this study.

Name	Descriptions	Source
Plasmids		
pET29a	pBR322 ori, Kan^R, lacI, T7lac promoter	Novagen
pET_CATec3 Y20F	CATec3 Y20F encoding gene in pET29a, 6X His-tag at C-terminus	This study
pACYCDuet-1	p15A ori, Cm^R, lacI, T7lac promoter	Novagen
pCDFDuet-1	CloDF ori, Sm^R, lacI, T7lac promoter	Novagen
pRSFDuet-1	RSF ori, Kan^R, lacI, T7lac promoter	Novagen
pET23a	pBR322 ori, Amp^R, T7 promoter	Novagen
pSIM6	pSC101 $repA^{ts}$, Gam, Beta, Exo under the control of a temperature sensitive promoter	[122]
pCP20	$repA101^{ts}$, Cm^R, Amp^R, FLP recombinase under the control of a temperature senstivie protmoer.	[123]
pHM46	pCDFDuet-1::*alsS*::*kivD*	[124]
pHM47	pET23a::*ilvC, ilvD, yqhD*	[124]
pHS121	pRSFDuet-1::*alsS*::*kivD*	This study
pHS122	pACYCDuet-1::*alsS*::*kivD*	This study
pHS133	pRSFDuet-1:: *alsS, catec3 Y20F*::*kivD*	This study
pHS144	pCDFDuet-1::*LeuABCD*	This study
pHS155	pCDFDuet-1:: *LeuABCD:: catec3 Y20F*	This study

Table 3-1 Continued.

Name	Descriptions	Source
E. coli		
Top10	Host for molecular cloning, mcrA, Δ(mrr-hsdRMS-mcrBC), Phi80lacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(SmR), endA1, nupG	Invitrogen
BL21(DE3)	F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)	Novagen
HSEC01	BL21(DE3) harboring pET_CATec3 Y20F	This study
HSEC0501	BL21(DE3) harboring pHM47 and pHS122	This study
HSEC0502	BL21(DE3) harboring pHM46 and pHS47	This study
HSEC0503	BL21(DE3) harboring pHM47 and pHS121	This study
HSEC0201	BL21(DE3) Δ*adhE*, Δ*dld*	This study
HSEC0302	BL21(DE3) Δ*adhE*, Δ*dld*, Δ*ldhA*	This study
HSEC0403	BL21(DE3) Δ*adhE*, Δ*dld*, Δ*ldhA*, Δ*ackA-pta*	This study
HSEC0504	BL21(DE3) Δ*adhE*, Δ*dld*, Δ*ldhA*, Δ*ackA-pta*, Δ*ilvE*	This study
HSEC0605	BL21(DE3) Δ*adhE*, Δ*dld*, Δ*ldhA*, Δ*ackA-pta*, Δ*ilvE*, Δ*tyrB*	This study
HSEC1006	BL21(DE3) harboring pHM47, pHS133, and pHS144	This study
HSEC1207	HSEC0201 harboring pHM47, pHS133, and pHS144	This study
HSEC1208	HSEC0201 harboring pHM47, pHS133_kivD V461A, F381L, and pHS144	This study
HSEC1209	HSEC0201 harboring pHM47, pHS133, and pHS144_LeuA G462D	This study
HSEC1210	HSEC0201 harboring pHM47, pHS133_kivD V461A, F381L, and pHS144_LeuA G462D	This study
HSEC1614	HSEC0605 harboring pHM47, pHS133_kivD V461A, F381L, and pHS155_LeuA G462D	This study

control (pSIM6) was used to induce homologous recombination of double strand DNA into the genome[122]. Briefly, *E. coli* strains harboring pSIM6 was cultured in 3 mL LB medium at 30℃ overnight. The grown cells were transferred to fresh 20 mL LB medium in a 250 mL flask with 1% inoculum size (200 μL) and cultured for 2-3 hours in a water bath shaking incubator at 200 rpm and 32℃. At OD_{600nm} of 0.4~0.6, the cell culture flask was transferred to a preheated 42℃ water bath shaking incubator at 200 rpm for 15 mins. Them the cells were immediately cooled down in ice for 10 mins and centrifuged at 4,700 rpm for 10 mins. The cell pellets were washed twice with 50 mL ice-cool sterile Millipore water and then suspended in the 200 μL ice-cool sterile Millipore water. 80 μL of the concentrated cells were mixed with ~100 ng of linear double-stranded DNA containing FRT-Kan-FRT cassette, amplified by PCR. Then, the cells were transferred to an ice-chilled 1-mm gap electroporation cuvette (BTX Harvard Apparatus, MA, USA) followed by an exponential decay pulse with 1.8 kV, 350 Ω, and 25 μF, which gave usual pulse duration of 4.5-6.0 ms. The cells were immediately mixed with 700 μL LB medium and recovered in a shaking incubator at 30℃ for two hours. The recovered cells were plated on LB solid medium with 25 μg/mL kanamycin and incubated at 30℃ for 2 days. Successful gene deletion was confirmed by colony PCR using multiple combinations of primers specifically binding at upstream and/or downstream of the target location and/or Kan in the FRT-Kan-FRT cassette. The kanamycin resistance marker was subsequently disrupted by FLP mediated recombination of FRT by pCP20[125].

3.2.4 Plasmid construction

Plasmids were constructed by ligation-dependent cloning and/or Gibson DNA assembly.

Briefly, DNA fragments were amplified using the Phusion DNA polymerase (cat# F530S, Thermo Fisher Scientific, MA, USA) and then purified by DNA purification and gel extraction kits (Omega Biotek, GA, USA). For the ligation-dependent cloning, the vectors and inserts were digested by restriction enzymes and ligated together using a T4 DNA ligase. In the case of Gibson DNA assembly cloning, the purified DNA fragments of the vector and insert were mixed together with the Gibson master mix[126] and assembled at 50℃ for one hour. Using the DNA mixtures, *E. coli* TOP10 was transformed by heat-shock transformation and selected on LB agar plates (15 g/L agar) with appropriate antibiotics. All the constructed plasmids were checked by PCR amplification and/or restriction enzyme digestion, and Sanger sequencing. The primers used in this chapter were listed in Table 3-2.

3.2.5 Conversion of alcohols to esters in *E. coli*.

For batch cultures, tube-scale alcohol conversions were performed in 4 mL M9 medium containing 10 g/L glucose with addition of 1 mL of hexadecane for *in situ* extraction at 37℃. 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the initial medium to induce expression of CATec3 Y20F. Alcohols were supplemented in the initial medium, and the product yield and titer were measured by esters collected in hexadecane layers and residual alcohols in culture media. Samples were analyzed after 24h for experiments with supplementation of single alcohols and after 12h, 24h, and 48h for the experiments simulating rose ester profiles. For fedbatch cultures designed to achieve high-level conversion of alcohols (i.e., isoamyl alcohol, phenylethyl alcohol), cells were grown microaerobically in a 125 mL screw capped shake flask with a working volume of 20 mL M9 medium containing 25 g/L glucose and 10 mL hexadecane.

Table 3-2: A list of primers used in this study. The bold and underlined letters indicate restriction and site-directed mutation sites, respectively.

Primer name	Primer sequence (5' to 3')	Description
HS677	CTCT **GTCGAC** TTAAGAAGGAGATATAGATATGAATTATACTAA	CAT cloning at MCS1 forward
HS678	CTCTAA **GCGGCCGC** TTACTTCAATTTCGAATTGCAGAGC	CAT cloning at MCS1 reverse
HS683	CTCT **GAGCTC** TATGAGCCAGCAAGTCATTATTTTCGATAC	LeuABCD cloning forward
HS684	CTCTAA **CTCGAG** TTAATTCATAAACGCAGGTTGTTTTGC	LeuABCD cloning reverse
HS802	CTCTAA **GCGGCCGC** TTAATTCATAAACGCAGGTTGTTTTGC	LeuABCD cloning at MCS1 reverse
HS785	AGTATAAGAAGGAGATATACATATGAATTATACT AAATTCGATG	CAT clonging at MCS2 forward
HS786	TTGCTCAGCGGTGGCAGCAGCCTAGGTTAATTTA CTTCAATTTCGAATTGC	CAT clonging at MCS2 reverse
HS701	ATGCGCTG <u>GAC</u> CAGGTGGATATCGTCGC	LeuA G462D mutation forward
HS702	CACCTG <u>GTC</u> CAGCGCATCTTTACCGTGG	LeuA G462D mutation reverse
HS703	TTATACA <u>GCG</u> GAAAGAGAAATTCATGG	kivD V461A mutation forward
HS704	TCTTTC <u>CGC</u> TGTATAACCATCATTATTG	kivD V461A mutation reverse
HS705	GGACATCA <u>CTG</u> TTTGGCGCTTCATCA	kivD F381L mutation forward
HS706	CGCCAAA <u>CAG</u> TGATGTCCCTTGTTCAG	kivD F381L mutation reverse
HS117	ATTTACTAAAAAAGTTTAACATTATCAGGAGAGC ATTATG GTGTAGGCTGGAGCTGCTTC	adhE deletion forward
HS118	GCCCAGAAGGGGCCGTTTATGTTGCCAGACAGCG CTACTGA CATATGAATATCCTCCTTA	adhE deletion reverse
HS372	CTATACTCTCGTATTCGAGCAGATG	adhE deletion check upstream
HS373	GGCCGTTTATATTGCCAGACAG	adhE deletion check downstream
HS31	TTTGTGATATTTTTTCGCCACCACAAGGAGTGGA AAATGGTGTAGGCTGGAGCTGCTTC	dld deletion forward
HS32	TAAGTGAATTCGGATGGCGATACTCTGCCATCCG TAATTTCATATGAATATCCTCCTTA	dld deletion reverse

Table 3-2 Continued.

Primer name	Primer sequence (5’ to 3’)	Description
HS39	CAGTTTATTGTCTGAATTTTCAAAATA	dld deletion check upstream
HS40	AGCTATAAAAAACAAAAAGCCGC	dld deletion check downstream
HS863	GCTTAAATGTGATTCAACATCACTGGAGAAAGTC TTATGGTGTAGGCTGGAGCTGCTTC	ldhA deletion forward
HS864	TCCCCTGCAACCCAGGGGAGCTGATTCAGATAAT CCCCAATCATATGAATATCCTCCTTA	ldhA deletion reverse
HS865	CCCGAGCGTCATCAGCAGCG	ldhA deletion check upstream
HS866	GGTCATTGCCAGCCCTTTGGCTG	ldhA deletion check downstream
HS713	CAAATCCGCGCCTGAGCGCAAAAGGAATATAAA AATGGTGTAGGCTGGAGCTGCTTC	ilvE deletion forward
HS714	AATGGGACGGTGCGTGCCGTCCCATTTTTTGTATC ATATGAATATCCTCCTTA	ilvE deletion reverse
HS717	AGTCAGTTAAATAAACTG	ilvE deletion check upstream
HS718	GCCATGGGTGGTGGTGGC	ilvE deletion check downstream
HS715	TAACCACCTGCCCGTAAACCTGGAGAACCATCGC GTGGTGTAGGCTGGAGCTGCTTC	tyrB deletion forward
HS716	GCTGGGTAGCTCCAGCCTGCTTTCCTGCATTACA CATATGAATATCCTCCTTA	tyrB deletion reverse
HS720	GTTGCTAATTGCCGTTC	tyrB deletion check upstream
HS873	TAGAACGATGGCATCAAAA	tyrB deletion check downstream

A volume of 25-50 μL of the alcohols (≥ 98% purity) were added to the culture at 6h, 9h, 12h, 15h, and 24h time points with a working concentration of 2 g/L per addition.

3.2.6 Fermentative isoamyl acetate production

For fermentative isoamyl acetate production, cells were microaerobically cultured in a 125 mL screw capped shake flask with a working volume of 20 mL M9 medium containing 40 g/L glucose and 5 g/L yeast extract. For production performance comparison of the engineered strains, cells were cultured at 37℃ for 48 hours. 10 mL of hexadecane (50% v/v) were overlaid to extract the isoamyl acetate during the fermentation. For the glucose fed-batch fermentation, cells were first cultured to OD 0.4~0.6 at 37℃ followed by the 0.1mM IPTG induction and subsequently cultured at 30℃. pH was adjusted by intermittent addition of 5M KOH.

3.2.7 Analytical methods

HPLC system (Shimadzu Inc., MD, USA) was used to quantify alcohols and sugars. 800 μL of samples were centrifuged at 17,000 x g for 3 minutes followed by filtering through 0.2 micron filters. The samples were run with 5 mM H_2SO_4 at 0.6 mL/min on an Aminex HPX-87H (Biorad Inc., CA, USA) column at 50℃. Refractive index detector (RID) and ultra-violet detector (UVD) at 220 nm were used to determine concentrations of sugars, organic acids, and alcohols.

Esters were measured by GC (HP 6890, Agilent, CA, USA) equipped with a MS (HP 5973, Agilent, CA, USA). For the GC system, the Zebron ZB-5 (Phenomenex, CA, USA) capillary column (30 m x 0.25 mm x 0.25 μm) was used to separate analytes, and helium was used as the carrier with a flow rate of 0.5 mL/min. The oven temperature program was set as follows: 50℃ initial temperature, 1℃/min ramp up to 58℃, 25℃/min ramp up to 235℃, 50℃/min ramp up to

300℃, and 2-minutes bake-out at 300ºC. 1 μL of sampled hexadecane layer was injected into the column in the splitless mode with an injector temperature of 280℃. For the MS system, selected ion mode (SIM) was used to detect and quantify esters with the following parameters: (i) ethyl acetate, m/z 45.00 and 61.00 from 4.2 to 4.6 minute retention time (RT), (ii) isopropyl acetate, m/z 45 and 102 from 4.7 to 5.0 minute RT, (iii) propyl acetate, m/z 59 and 73 from 5.2 to 5.8 minute RT, (iv) ethyl isobutyrate, m/z 73 and 116 from 6.1 to 6.6 minute RT, (v) isobutyl acetate, m/z 61 and 101 from 6.6 to 7.6 minute RT, (vi) butyl acetate, m/z 61 and 116 from 7.7 to 9.2 minute RT, (vii) isobutyl isobutyrate, m/z 89 and 129 from 10.1 to 12.5 minute RT, (viii) benzyl acetate, m/z 108 and 150 from 13.1 to 13.8 minute RT, and (ix) 2-phenethyl acetate, m/z 104 and 121 from 13.8 to 15.5 minute RT. As an internal standard, 10 mg/L n-decane were added in initial hexadecane layer and detected with m/z 85, 99, and 113 from 12 to 15-minute retention time range. The esters were identified by RT and quantified by the peak areas and standard curves. Standard curves were determined by using pure esters diluted into hexadecane at concentrations of 0.01 g/L, 0.05 g/L, 0.1 g/L, 0.5 g/L, and 1 g/L.

3.3 Efficient alcohol conversion by *E. coli* whole-cell biocatalyst

3.3.1 Microbial conversion of alcohols to esters

Microbial conversion presents a renewable and sustainable route to synthesize chemicals. Specifically, utilizing AATs for microbial ester production is most desirable because acyl-CoAs are very expensive for cell-free ester synthesis. Even though CATec3 Y20F is compatible with a broad range of alcohols and acyl-CoAs as presented in Chapter 2, its performance in whole-cell biocatalyst is unknown due to complicated reaction environments (e.g., intracellular enzyme abundance, precursor availability, intracellular pH, and temperature). Here, we characterized how

efficient and compatible the engineered CATec3 Y20F is to enable biosynthesis of designer esters in an *E. coli* whole-cell biocatalyst, a workhorse for industrial biotechnology.

3.3.2 Harnessing the engineered CAT for designer ester biosynthesis in *E. coli*.

To demonstrate the biosynthesis of designer acetate esters, we engineered HSEC01, a recombinant *E. coli* BL21 (DE3) harboring CATec3 Y20F under the control of a T7lac promoter that uses its native metabolism to convert fermentable sugars into the precursor acetyl-CoA and characterized the engineered strain in the growth media supplemented with representative linear, branched, saturated, unsaturated, and/or aromatic alcohols and with hexadecane for *in situ* ester extraction (Figure 3-1A). The result showed that the recombinant CATec3 Y20F-expressing *E. coli* produced all the expected acetate esters with relatively high conversion efficiency. In batch cultures, conversion of all alcohols to their respective acetate esters achieved more than 50% (mol/mol) yield within 24 h (Figure 3-1B). Noticeably, yields of phenylethyl and geranyl acetate reached up to 80% (mol/mol). The recombinant *E. coli* produced and secreted the acetate esters at final titers of 2.6 g/L butyl acetate, 2.3 g/L of isobutyl acetate, 3.1 g/L pentyl acetate, 2.9 g/L isoamyl acetate, 2.6 g/L 3-hexenyl acetate, 0.9 g/L benzyl acetate, 1.2 g/L 2-phenylethyl acetate, and 0.3 g/L geranyl acetate (Figure 3-1C). To further increase ester production, we performed fed-batch fermentation using the representative branched isoamyl alcohol and aromatic phenylethyl alcohol. With the intermittent feeding of up to 10 g/L alcohols, the recombinant *E. coli* produced the expected esters at a relatively high efficiency, achieving titers of 13.9 g/L and 10.7 g/L and yields of 95% (mol/mol) and 80% (mol/mol) for isoamyl acetate and phenylethyl acetate, respectively (Figure 3-2A). We did not observe any noticeable growth inhibition at this high level of ester production in the fed batch mode due to the *in situ* ester extraction with hexadecane (Figure 3-2B,

C), even though both alcohols and esters are known to be toxic to microbial health at relatively low concentrations (< 2 g/L)[127]. The turnover numbers of CATec3 Y20F were similar or better for biosynthesis of isobutyl acetate and isoamyl acetate as compared to a widely used AAT from *Saccharomyces cerevisiae* (ATF1), in engineered *E. coli* strains[50, 66, 117, 128]. Unlike ATF1, CATec3 Y20F is thermostable and compatible with broader range of alcohols and acyl-CoAs. As the heterologous pathways for higher alcohols have been metabolically engineered in recombinant organisms (e.g., *E. coli*)[55, 110, 129], designer bioesters can be produced by using the engineered CAT and either co-feeding fermentable sugars and alcohols as demonstrated here or via natural fermentative processes which produce alcohols natively.

3.3.3 Naturally reconstituting an ester profile of roses in *E. coli*

To demonstrate the high efficiency and substrate compatibility of the engineered CAT for designer ester biosynthesis in the *E. coli* whole-cell biocatalyst, we explored whether CATec3 Y20F could be harnessed as an efficient AAT to create ester profiles exhibited by floral plants. In nature, eukaryotic AATs in floral plants formulate a mixture of volatile esters that contribute to their beneficial interactions with the environment. For example, roses (*Rosa hybrida*) utilize AATs, regulated within their developing stages[130], to synthesize and emit volatile esters attracting pollinators together with other aromatic molecules (Figure 3-2A). To mimic an ester profile of *R. hybrida*, we fed HSEC01 with a mixture of alcohols at a total working concentration of 1 g/L, consisting of 0.2 g/L hexanol, 0.2 g/L 0.15 g/L 3-cis-hexen-1-ol, benzyl alcohol, 0.15 g/L phenylethyl alcohol, 0.1 g/L geraniol, 0.1 g/L nerol, and 0.1 g/L citronellol at a mid-log phase (OD_{600nm}~1.0).

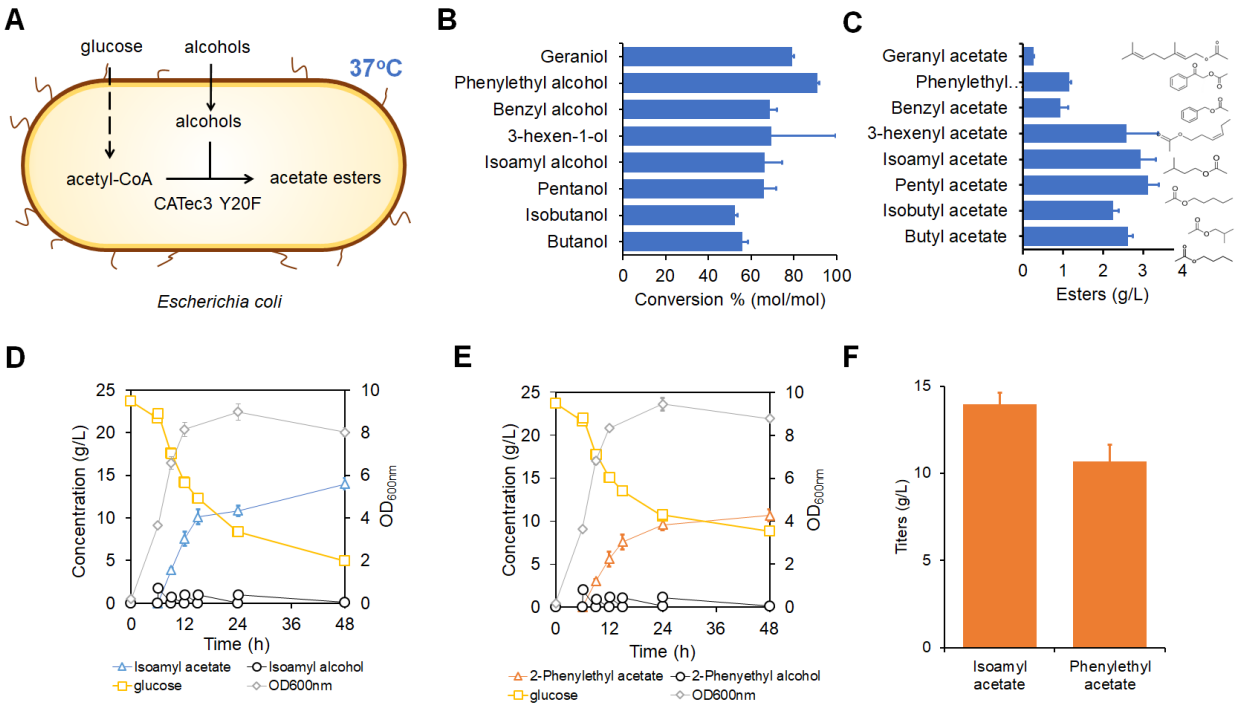


Figure 3-1: Alcohol conversion by *E. coli* overexpressing CATec3 Y20F. **(A)** Microbial ester biosynthesis by a CATec3 Y20F-expressing *E. coli* at 37°C. **(B)** Alcohol conversion efficiencies and **(C)** Final titer of acetate esters from alcohol conversion by *E. coli*. Each value represent mean ± 1 stdev from three biological replicates. Supplemented alcohol concentrations were 3 g/L for butanol, isobutanol, pentanol, and isoamyl alcohol, 1 g/L for benzyl alcohol, and phenylethyl alcohol, and 0.3 g/L for geraniol. **(D)** Fed-batch conversion of isoamyl alcohol to isoamyl acetate. Each value represent mean ± 1 stdev from three biological replicates. **(E)** Fed-batch conversion of 2-phenylethyl alcohol to 2-phenylethyl acetate. Each value represent mean ± 1 stdev from three biological replicates. **(F)** Titers of isoamyl acetate and 2-phenylethyl acetate by the CATec3 Y20F-expressing *E. coli* BL21(DE3) strain HSEC01 from the alcohol feeding conversion after 48 hours.

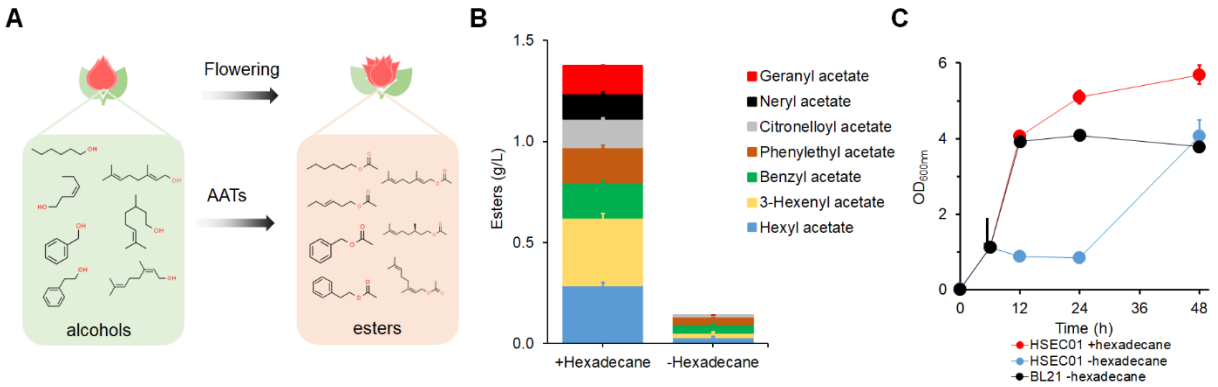


Figure 3-2: Reconstituting an ester profile of roses. **(A)** Ester biosynthesis in developing rose petals. Roses (*Rosa hybrida*) synthesize volatile esters by alcohol acyltransferases (AATs), contributing to its unique aroma. **(B)** Simulated ester profile of roses by HSEC01 harboring CATec3 Y20F. Each value represents mean ± 1 stdev from three biological replicates. **(C)** Cell growth during the alcohol mixture Each value represents mean ± 1 stdev from three biological replicates. The arrow indicates the point when the alcohol mixture was added to the medium.

The recombinant *E. coli* could rapidly and completely convert the alcohol mixture into the desirable acetate ester profiles with a yield of 97.1 ± 0.7 % (mol/mol) and a titer of ~1.5 g/L within 12 h (Figure 3-2B). For high production of esters, the *in situ* extraction using hexadecane overlay is very critical to mitigate the toxicity of esters (Figure 3-2C).

3.4 Engineered isoamyl acetate production in *E. coli*

3.4.1 Microbial isoamyl acetate production

Motivated by the efficient conversion of various alcohols to acetate esters, we next questioned whether the *E. coli* metabolism can be engineered to overproduce an acetate ester from glucose fermentation without the alcohol doping. As a proof of concept, we engineered isoamyl acetate production pathway that is a branched medium chain length ester.

Isoamyl acetate, also known as isopentyl acetate or 3-methyl-1-butyl acetate, is a versatile ester with broad industrial applications as solvents and/or flavoring agents in food and beverage, cosmetic, textile, aerospace, and pharmaceutical industries. Also, an isomer of isoamyl acetate, ethyl valerate, is fully compatible for blending with gasoline or diesel[62], suggesting potential application of isoamyl acetate as drop-in biofuel. Global market of isoamyl acetate was about 5 billion USD in 2019, and it is expected to continuously grow by 2026[131, 132]. Like other ester manufacturing, isoamyl acetate is synthesized by Fisher esterification of isoamyl alcohol and acetic acid in the presence of strong acids. The process is thermodynamically unfavorable that necessitates a large amount of alcohol and strong acid to push forward the reversible reaction. Also, the water should be removed during the process to prevent retardation of the reaction. In contrast, alcohol acyltransferase (AAT) dependent biological process synthesizes isoamyl acetate

by condensing an isoamyl alcohol and acetyl-CoA. Because the reaction is thermodynamically favorable, esters can be synthesized through fermentation under mild aqueous condition.

Recently, metabolic engineering strategies employing the AATs enabled high production of short to medium chain esters such as ethyl acetate [69], butyl acetate [133], and isobutyl acetate [66]. However, most of the AATs have poor substrate specificity[50, 65, 66, 134] and therefore, a long chain ester pathway involving intermediate alcohol pathways can potentially produce undesired short chain ester byproducts. Especially, isoamyl acetate pathway is so similar with isobutyl acetate pathway that requires precise metabolic control to reduce byproducts and enhance titer and yield.

Microbial isoamyl acetate production was actively studied in early 2000s with highlights on relationship between intracellular CoA and acetyl-CoA pool with isoamyl acetate production[135-138]. However, due to the limited knowledge on the AAT and pathway engineering, production titer was relatively low (<250 mg/L). Progresses of metabolic engineering could achieve high levels of isoamyl alcohol production[139, 140], a key precursor of isoamyl acetate, suggesting that bridging the alcohol production platform with an appropriate ester biosynthesis platform can potentially produce isoamyl acetate at high titer. A critical is to manipulate pathway specificity within multiple gene overexpression for an efficient isoamyl acetate production.

In this study, we engineered *E. coli* to produce isoamyl acetate from glucose fermentation. We investigated specificity control by engineering isobutanol and isoamyl alcohol pathway modules to achieve a high titers and specificity of isoamyl acetate production. The optimized strain produced isoamyl acetate at titers of 8g/L with reduced isobutyl acetate byproduct. The flux shift and amplification by stepwise engineering suggests metabolic bottlenecks in engineered isoamyl acetate production.

3.4.2 Optimization of isobutanol production module

Isoamyl acetate production pathway starts from pyruvate (Figure 3-3). Pyruvate is converted to 2-ketoisovalerate (2-KIV) which is elongated to 2-ketoisocaproate (2-KIC) via the L-leucine biosynthesis pathway facilitated by LeuABCD operon. The Ehrilch pathway driven by ketoacid decarboxylase (KDC) converts the 2-KIC to isovaleraldehyde which is reduced to isoamyl alcohol by alcohol dehydrogenases. Finally, the isoamyl alcohol is condensed with acetyl-CoA to isoamyl acetate via the AAT pathway. To maximize the flux, strong metabolic drive from pyruvate to 2-KIV is important. Since the pathway has been well studied for isobutanol production[141], we first aimed to construct an efficient isobutanol pathway that can be latter rewired to isoamyl acetate production.

It has been suggested that heterologous expression of AlsS (2-acetolactate synthase) from *Bacillus subtilis* and KivD (2-ketoisovalerate decarboxylase) from *Lactococcus lactis* is critical for high level isobutanol production[141]. We constructed three different plasmids harboring *alsS* and *kivD* with different plasmid copy numbers from a previously reported plasmid pHM46[124], and introduced to wild type BL21 (DE3). For a sufficient flux, *ilvC* (ketoacid reductoisomerase), *ilvD* (Dihydroxy-acid dehydratase), and *yqhD* (alcohol dehydrogenase) were overexpressed under the control of T7 promoter (Figure 3-4A). Then, isobutanol production by the recombinant *E. coli* strains were measured to identify the most efficient combination of the production modules (Figure 3-4B). Interestingly, the higher copy number of *alsS* and *kivD* achieved the higher isobutanol titer. HSEC0503 harboring the pRSFDuet-1 backbone plasmid with >100 plasmid copy number (PCN) produced 5.9 g/L of isobutanol within 48 hours, yet *alsS* and *kivD* in a relatively low copy number plasmid (pACYCDuet-1) produced only 0.03 g/L of isobutanol. Therefore, we started from the production modules with *alsS* and *kivD* in the high copy number plasmid.

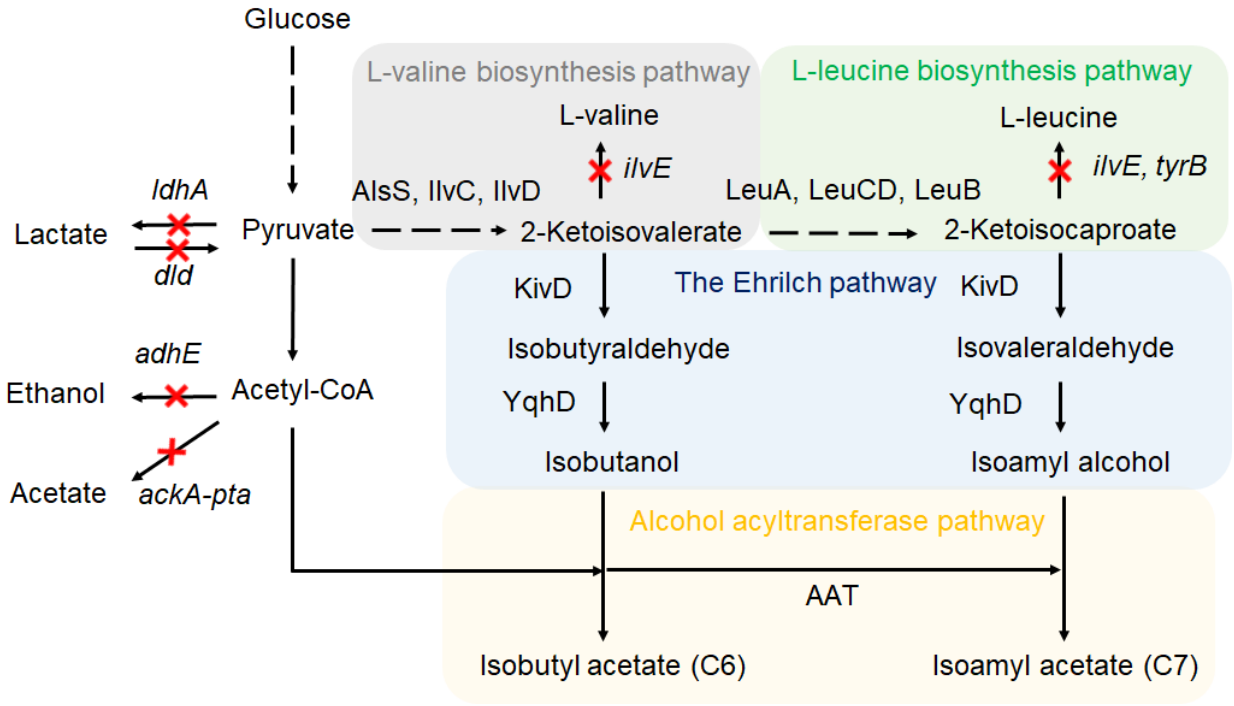
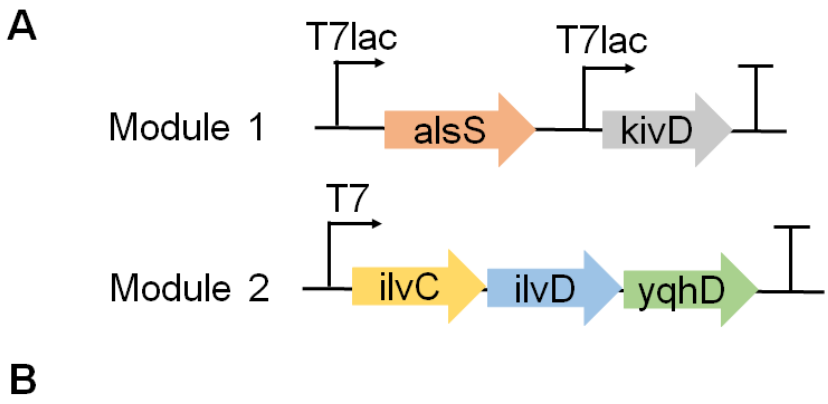
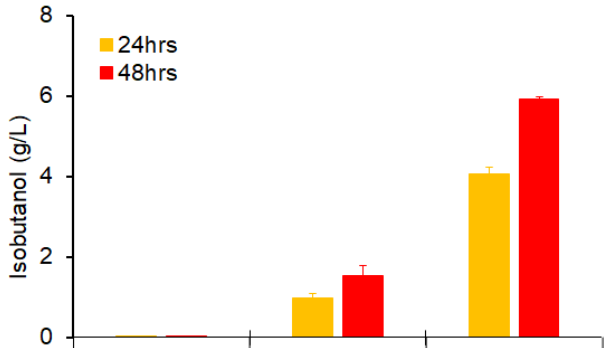


Figure 3-3: A map of metabolic pathway for isoamyl acetate production. Gene deletion targets are presented as the red cross marks.



B



	HSEC0501	HSEC0502	HSEC0503
Module 1	pACYC	pCDF	pRSF
Module 2	pET	pET	pET
Final OD	4.61 ± 0.02	5.72 ± 0.50	3.76 ± 0.09

Figure 3-4: Optimization of isobutanol production module. **(A)** Operon architecture of production modules, **(B)** Isobutanol production by *E. coli* BL21(DE3) harboring the production modules. Plasmid copy number of pRSF, pCDF, pET, and pACYC are >100, 20~40, 15~20, and 10~12, respectively.

3.4.3 Push and pull engineering

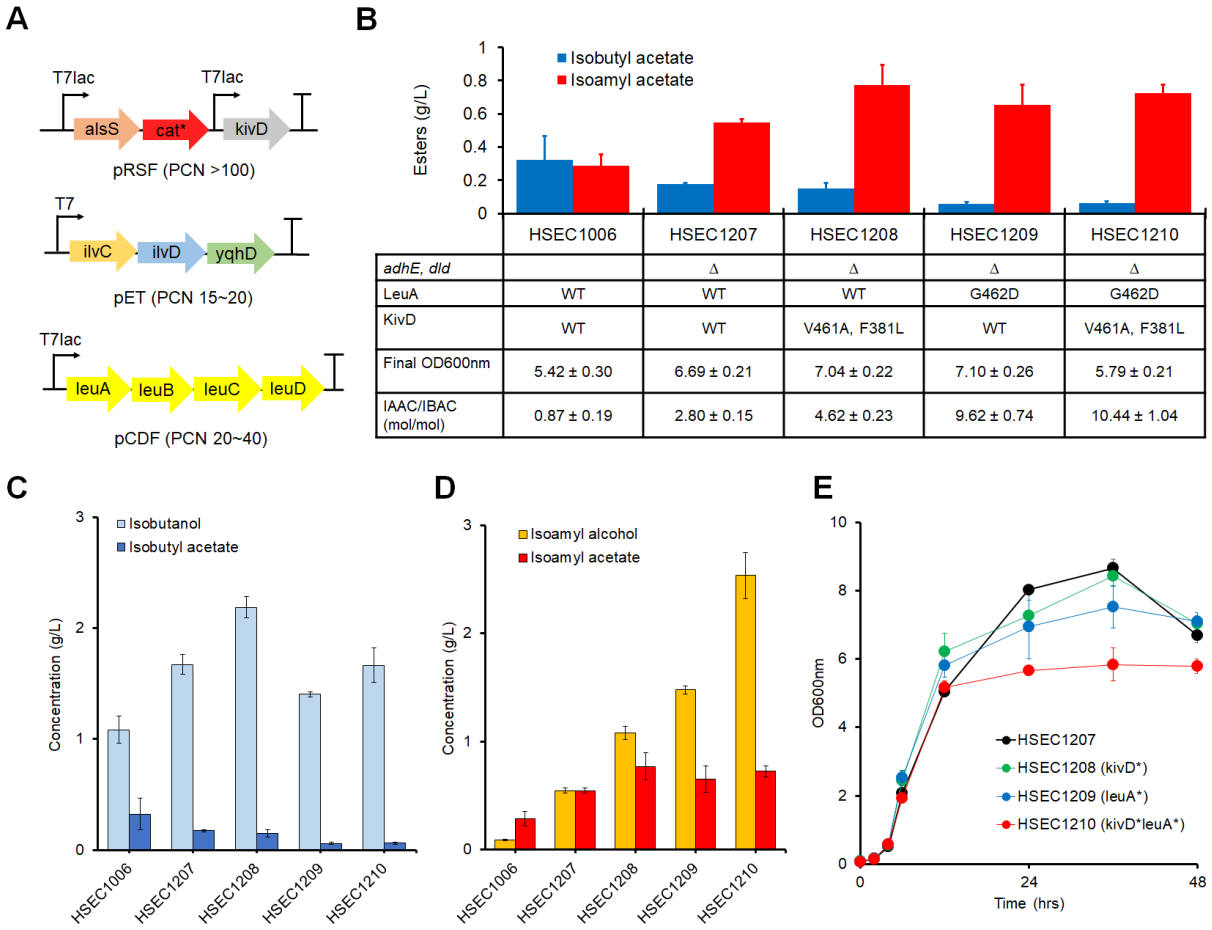
When metabolic flux is balanced, combination of deletion or downregulation of competitive upstream pathways and overexpression of metabolite-forming upstream pathways can achieve large flux amplification[142]. Starting from the isobutanol production module, we employed the "push-and-pull" metabolic engineering strategy to shift the flux towards isoamyl acetate.

Because KivD, YqhD are promiscuous, overexpression of LeuABCD can rewire the isobutanol pathway to isoamyl alcohol, which can be converted to isoamyl acetate by expressing an AAT (Figure 3-3). We constructed a *leuABCD* operon in pCDFDuet-1 plasmid to complete the isoamyl alcohol production pathway (Figure 3-5A). Theoretically, expressing an AAT on top of the isoamyl alcohol production modules can produce isoamyl acetate. However, due to the high similarity of intermediate metabolites of the isobutyl acetate and isoamyl acetate pathways, choosing an isoamyl alcohol specific AAT was important for specific isoamyl acetate production. Previously, we engineered a chloramphenicol acetyltransferase CATec3 Y20F to produce a broad range of esters in microbes[59]. The use of CATec3 Y20F capitalizes 3.5-folds higher catalytic efficiency (k_{cat}/K_M) towards isoamyl alcohol than isobutanol (Table 2-9). It should be noted that other AATs have similar catalytic efficiency towards isobutyl acetate and isoamyl acetate[66]. Because the acetate esters are toxic to cell growth, hexadecane was overlaid on the culture to extract the produced esters *in situ*. Wild type BL21(DE3) expressing the isoamyl acetate production modules (HSEC1006) produced 0.29 g/L isoamyl acetate and 0.32 g/L isobutyl acetate.

We hypothesized that improving flux to isoamyl alcohol is required to leverage the isoamyl acetate titers and specificity. Therefore, we implemented three strategies: 1) deletion of pyruvate and acetyl-CoA pool related genes, *adhE* (aldehyde-alcohol dehydrogenase) and *dld* (quinone-dependent D-lactate dehydrogenase), 2) expression of a feedback insensitive LeuA G462D

mutant[143], and 3) expression of longer chain keto-acid specific KivD V461A, F381L[55] (Figure 3-5B). The engineered strains produced less isobutyl acetate down to 0.07 g/L while isoamyl acetate titers reached to 0.8 g/L (Figure 3-5B). Although the engineering strategies could improve the isoamyl acetate specificity, isoamyl acetate titers were not further improved. Interestingly, more than 1 g/L of isobutanol and isoamyl alcohol were accumulated that likely inhibited HSEC1210 growth (Figure 3-5C, D, E). Especially, 2.5-3 g/L level of isoamyl alcohol inhibits 50-80% of cell viability[139, 144]. Therefore, we hypothesized that AAT activity should be improved to prevent the alcohol accumulation.

Gene arrangement in an operon affects expression level of the individual genes[145]. Because CATec3 Y20F was cloned at the downstream of *alsS* (Figure 3-5A), expression level of CATec3 Y20F could be compromised. To improve the CATec3 Y20F, we simply cloned additional monocistronic CATec3 Y20F operon in the pCDF plasmid under the control of T7lac promoter (Figure 3-6A). The additional expression of CATec3 Y20F significantly reduced isoamyl alcohol accumulation and improved isoamyl acetate production (Figure 3-6B, C), suggesting that high expression level of the AAT is important to avoid the accumulation of the toxic precursor alcohols. Although HSEC1311 produced 4.1g/L of isoamyl acetate, glucose consumption was significantly inhibited after 36 hours of the culture (Figure 3-6D), suggesting that there is another bottleneck(s) causing the fermentation cessation. The fermentation cessation was similarly observed from an engineered *E. coli* strain producing isoamyl alcohol (3-methyl-1-butanol), reasoning toxicity of isoamyl alcohol[140]. However, our system did not accumulate isoamyl alcohol higher than 0.2 g/L (Figure 3-6B), and isoamyl acetate was efficiently extracted to hexadecane layer during the fermentation at 94 ± 2% (w/w) efficiency.



	HSEC1006	HSEC1207	HSEC1208	HSEC1209	HSEC1210
adhE, dld		Δ	Δ	Δ	Δ
LeuA	WT	WT	WT	G462D	G462D
KivD	WT	WT	V461A, F381L	WT	V461A, F381L
Final OD600nm	5.42 ± 0.30	6.69 ± 0.21	7.04 ± 0.22	7.10 ± 0.26	5.79 ± 0.21
IAAC/IBAC (mol/mol)	0.87 ± 0.19	2.80 ± 0.15	4.62 ± 0.23	9.62 ± 0.74	10.44 ± 1.04

Figure 3-5: Engineered strain comparison. **(A)** Architecture of synthetic operons to produce isoamyl acetate. PCN: plasmid copy number. **(B)** Titers of isobutyl acetate and isoamyl acetate by engineered E. coli strains. **(C-E)** Characterization of the engineered strains. **(C)** Isobutanol and isobutyl acetate titers, **(D)** Isoamyl alcohol and isoamyl acetate titers, **(E)** Cell growth. The star (*) denotes mutant leuA and kivD. All data are presented as means ± stdev of three biological replicates.

We hypothesized that the growth condition might be the reason of the fermentation cessation because the pH after 48 hours of the culture was as low as 5.5. Therefore, we next deleted genes relevant to the organic acid formulation, *ldhA* (L-lactate dehydrogenase) and *ackA-pta* (An operon of acetate kinase and phosphate acetyltransferase), which have been commonly deleted to reduce the lactate and acetate formulation[15, 146]. To improve ketoacid availability, we also subsequently deleted *ilvE* (branched chain amino acid aminotransferase) and *tyrB* (aromatic-amino-acid aminotransferase)[139].

Unfortunately, additional deletions did not significantly improve isoamyl acetate titers and yields (Figure 3-7A, B), suggesting that the upstream competitive pathways were not major bottlenecks of the system. The production yield reached up to 0.21 (g/g) that corresponds to 59% of the theoretical maximum yield (0.36 g/g). The yield is higher than a previously reported isoamyl alcohol producing *E. coli* that reported 0.11 (g/g) corresponding to 33% of the theoretical maximum yield[140]. Theoretically, one mole of isoamyl acetate requires two moles of glucose that supplies two acetyl-CoAs and one isoamyl alcohol. The pathway necessitates production of five mole excess NAD(P)H that could hamper the fermentation under the microaerobic condition without an appropriately coupled electron sink pathway.

3.4.4 Glucose fed-batch fermentation

The comparable production yield prompted us to investigate glucose-fed batch fermentation with pH control for higher isoamyl acetate production titer. We employed HSEC1614 (BL21 (DE3) Δ*adhE* Δ*dld* Δ*ldhA* Δ*ackA-pta* Δ*ilvE* Δ*tyrB* harboring the isoamyl acetate production modules) to produce isoamyl acetate over 120 hours of glucose fed-batch fermentation.

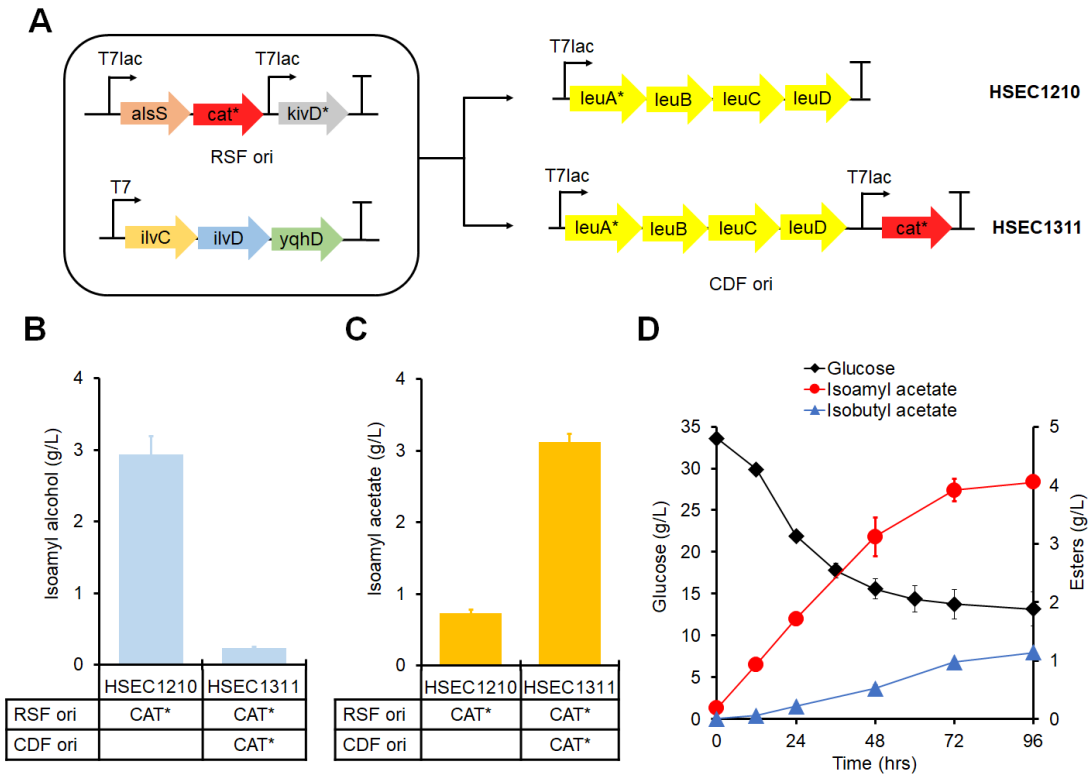


Figure 3-6: Increasing the AAT activity by additional clone of CATec3 Y20F. **(A)** A diagram of production module combination in HSEC1210 and HSEC1311. **(B)** Accumulated isoamyl alcohol concentration in the medium. **(C)** Produced isoamyl acetate in the hexadecane layer. **(D)** A time dependent profile of glucose, isoamyl acetate, and isobutyl acetate of HSEC1311 culture.

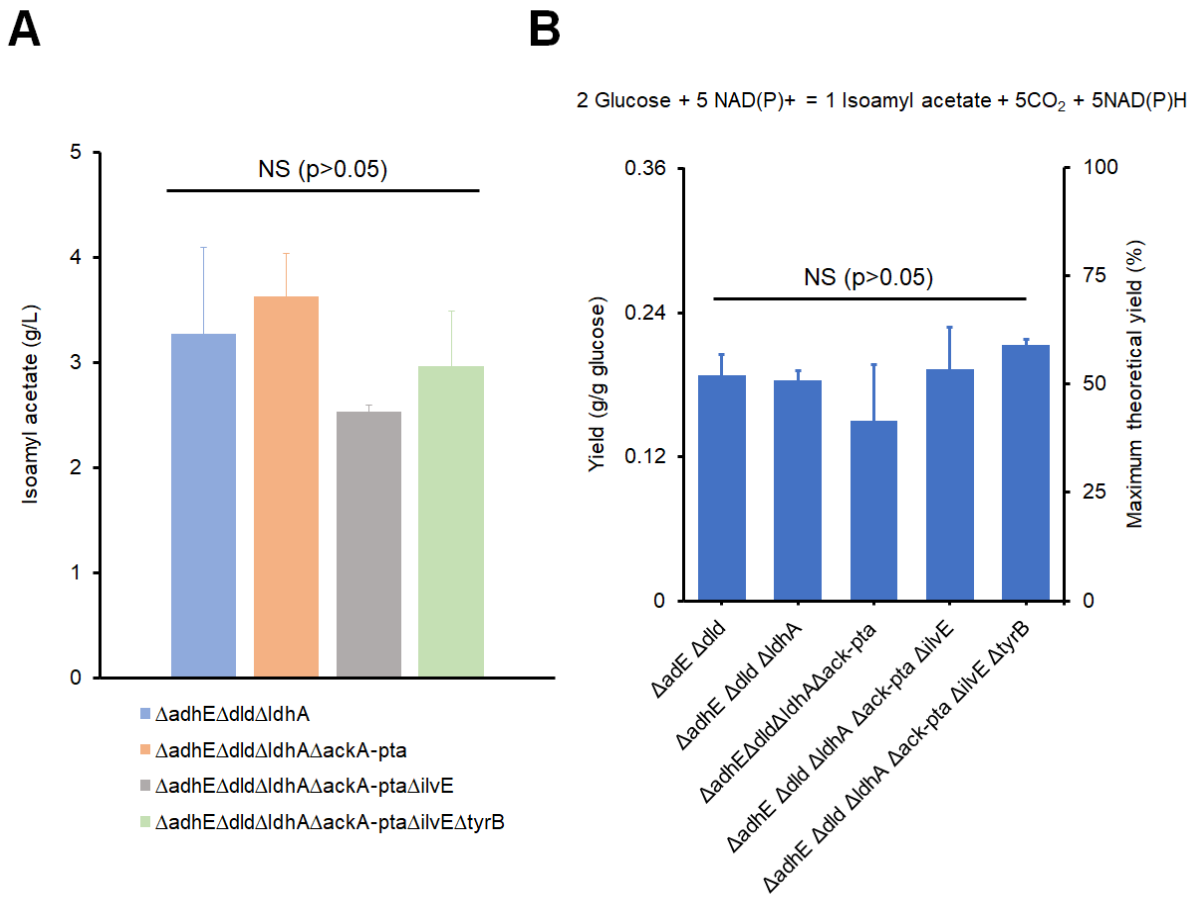


Figure 3-7: Isoamyl acetate production by deletion strains. **(A)** Isoamyl acetate titers after 48 hours. **(B)** Isoamyl acetate production yield and corresponding maximum theoretical yield percentage. The yield was measured at between 24 and 36 hours, the middle of production phase. All the data presents mean ± 1 stdev from three biological replicates.

The pH was adjusted with 5M KOH every 12 hours after 36 hours until 96 hours, maintained at between 6.0 and 7.5 (Figure 3-8). Final isoamyl acetate titers reached to 8.0 g/L corresponding to overall yield of 0.16 (g/g), from 51 g/L of consumed glucose over the 120 hours with volumetric productivity of 2.6 (g/L/day). The intermittent pH adjustment could maintain the fermentation process over the 120 hours, suggesting that pH control is important to prevent the fermentation cessation. Isobutyl acetate, a major byproduct, increased from 36 hours to 72 hours up to final concentrations of 1.0 g/L. A possible explanation of the isobutyl acetate is promiscuous KDC activity of the AlsS. The AlsS enables isobutanol production without KivD when IlvCD are overexpressed in *E. coli*[147].Overall, pH adjusted glucose fed-batch fermentation of the engineered *E. coli* could produce isoamyl acetate up to 8.0 g/L.

3.5 Conclusion

In this chapter, modularity of ester biosynthesis at metabolic network level was engineered in a model mesophilic bacterium *E. coli*, demonstrating high alcohol conversion and ester titer. The metabolic engineering of *E. coli* for isoamyl acetate production articulated strategies to control the product specificity by pathway module engineering. With non-toxic hexadecane overlay, the engineered *E. coli* produced isoamyl acetate above than solubility limit (2 g/L), suggesting that higher chain esters with low solubility can be produced at high titers and yield. Acylation of alcohols generally increases hydrophobicity that allows *in situ* product recovery via hybrid extraction-distillation process which can recover the ester products at high concentration and purity followed by distillation with a lower energy requirement[148]. Because long chain esters are much less toxic than their precursor alcohols[149-151], esterification of the toxic alcohols will be advantageous for overall bioprocessing.

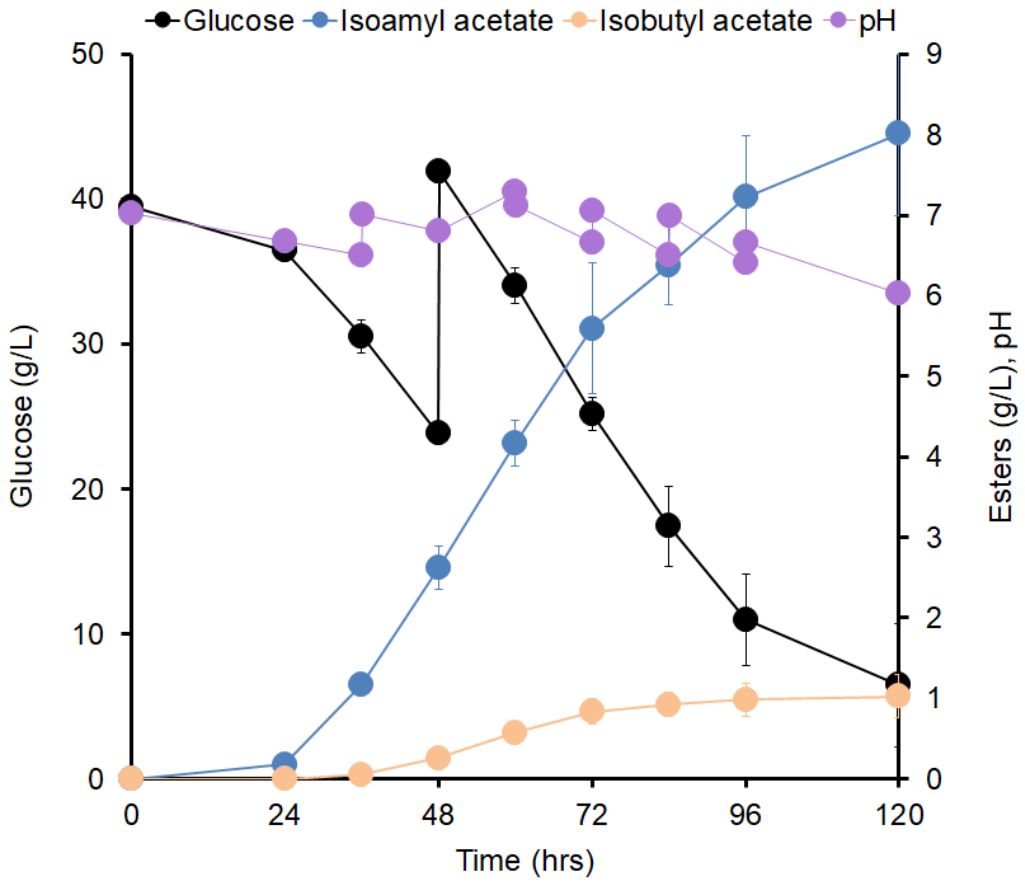


Figure 3-8: Fed-batch isoamyl acetate production by HSEC1614*(ΔadhE Δdld ΔldhA ΔackA-pta ΔilvE ΔtyrB* harboring the isoamyl acetate production modules).

The detoxification by esterification was recently demonstrated for geranyl acetate production in engineered *E. coli*[150]. Acylation of alcohols generally increases hydrophobicity that allows *in situ* product recovery via hybrid extraction-distillation process which can recover the ester products at high concentration and purity followed by distillation with a lower energy requirement[148].

Engineering specificity of LeuA, KivD, and AAT can further exploit the production pathways for a longer chain branched acetate esters such as isohexyl acetate (C_8) and isoheptyl acetate (C_9) with controlled metabolic flux as described here. Protein engineering of LeuA from *E. coli* and KivD from *L. lactis* can increase the binding pocket size for higher chain alcohol production[55], and indolepyruvate decarboxylase (IPDC) having higher specificity towards long chain ketoacids exists[152]. Therefore, microbial production of the long-branched chain esters at high titers will be feasible with a combination of metabolic engineering and process engineering.

Chapter 4. Engineering metabolic network/cellular level modularity of ester biosynthesis in non-model organisms

This chapter focuses on engineering modularity of ester biosynthesis at metabolic network/cellular level in a non-model cellulolytic and thermophilic bacterium, *Clostridium thermocellum*. Parts of this chapter are slightly modified from published papers[58, 59, 153].

4.1 Motivation

More than 80% of metabolic engineering research have employed model-organisms such as *Escherichia coli* and *Saccharomyces cerevisiae* because they are well characterized, and many genetic engineering tools and parts are available. Although the model organisms are useful to develop a novel metabolic pathway, it is often difficult to obtain desirable phenotypes even with extensive engineering. A representative example is to engineer model microbes to efficiently utilize recalcitrant lignocellulose, the most abundant carbon source in the planet, and produce valuable chemicals as well. Degradation and utilization of the polysaccharide usually require multiple enzymes in a concerted action[154, 155], which is challenging to accomplish by engineering a non-cellulolytic microbe. Heterologous engineering of cellulose utilization necessitates effective expression and secretion of multiple cellulases and therefore, building a production pathway on top of an engineered cellulose assimilation system usually results in metabolic burden and imbalance compromising production efficiency and/or the cellulose fermentation capability[156, 157].

Domesticating a non-model cellulolytic bacterium is another approach to develop bioprocesses utilizing lignocellulose for valuable chemical production. Cellulolytic bacteria have been evolved to efficiently degrade and consume cellulose with robust enzyme systems such as

cellulosome[154]. Therefore, cellulolytic bacteria have increasingly become attractive for direct conversion of lignocellulosic biomass into esters in a consolidated bioprocessing (CBP) configuration where biomass degradation and fermentation simultaneously take place in a single step and hence reduce the production cost[158].

A gram-positive and thermophilic anaerobe *Clostridium thermocellum* (*Hungateiclostridium thermocellum*) is an ideal cellulolytic bacterium for CBP because it is the best biomass degrader known to date and has endogenous metabolic pathways for synthesizing various alcohols including ethanol and isobutanol[159-161]. Also, due to high volatility of esters, the bioprocess at elevated temperatures can benefit downstream product separation and hence reduce the process cost. However, most of the AATs known to date are isolated from mesophilic microbes or plants[162-165], and none of them has been reported to be active at elevated temperatures (> 50ºC). The highest temperature reported for ester production is 42℃ in a thermotolerant yeast[68]. Therefore, engineering *C. thermocellum* can attain a milestone for CBP ester production from lignocellulose at elevated temperatures.

In this chapter, we engineered *C. thermocellum* to produce ester from cellulose fermentation by harnessing the engineered robust ester biosynthesis module. We screened esterases responsible for ester hydrolysis and disrupted the two critical esterases that helped alleviate the isobutyl acetate degradation in *C. thermocellum* while not affecting cellulose utilization. By introducing the engineered CATec3 Y20F, the engineered esterase-deficient *C. thermocellum* strain could convert various alcohols into esters by cellulose fermentation. In combination with gene expression and fermentation optimization, we demonstrated the first-generation of engineered *C. thermocellum* capable of producing 0.34 g/L of C4-derived esters by

direct cellulose fermentation and 1 g/L of C4-derived esters with an isobutanol supplemented cellulose fermentation.

4.2 Materials and methods

4.2.1 Bacterial strains and plasmids

Strains and plasmids are listed in Table 4-1. *C. thermocellum* DSM 1313 Δ*hpt* (M1354)[166] was used as the parent strain because the *hpt* deletion allows counter-selection for markerless gene deletion[167]. *Escherichia coli* TOP10 and BL21(DE3) were used for molecular cloning and dam methylation of plasmids for *C. thermocellum* transformation[168], respectively. For characterization of carbohydrate esterases, *E. coli* BL21 CodonPlus (DE3)- RIPL (Agilent, CA, USA) was used as an expression host to avoid potential expression problems caused by codon bias.

4.2.2 Bioinformatics

The homology model of Clo1313_0613 was generated using the Swiss-Model software[169] and visualized using the Molecular Operating Environment software (MOE, version 2019.01). The 3D structure was energy minimized with the Amber10: EHT force, and the binding pocket was searched using the 'Site Finder' tool in MOE.

Identified esterases were sequence aligned with the homolog proteins from different organisms. The representative homologous proteins of Clo1313_0613 and Clo1313_0693 were identified from the Carbohydrate Active Enzymes (CAZy) database (http://www.cazy.org/) along with BLAST search[170, 171] and their sequences were retrieved from the NCBI protein sequence repository. Their domain sequences were aligned by ClustalW[172] in Mega7[81].

Table 4-1: A list of plasmids and strains used in this study.

Name	Descriptions	Source
Plasmids		
pETDuet-1	pBR322 ori, Amp^R, lacI, T7lac promoter	Novagen
pET_1424t	Clo1313_1424t cloned at MCS1 site. Original signal peptides removed and tagged with 6x His at N-terminus	This study
pET_1139	Clo1313_1139 cloned at MCS1 site. Original signal peptides removed and tagged with 6x His at N-terminus	This study
pET_1680	Clo1313_1680 cloned at MCS1 site. Original signal peptides removed and tagged with 6x His at N-terminus	This study
pET_0521	Clo1313_0521 cloned at MCS1 site. Original signal peptides removed and tagged with 6x His at N-terminus	This study
pET_0613	Clo1313_0613 cloned at MCS1 site. 6x His at N-terminus	This study
pET_0693	Clo1313_0693 cloned at MCS1 site. Original signal peptides removed and tagged with 6x His at N-terminus	This study
pHS005	pNW33N derivative, The backbone plasmid for gene deletions	This study
pHS0024_F97W	pNW33N derivative, CATsa_F97W under *C. thermocellum* gapDH promoter	This study
pHS0024_CATec3	pNW33N derivative, CATsa_F97W under *C. thermocellum* gapDH promoter	This study
pHS0024_Y20F	pNW33N derivative, CATsa Y20F under *C. thermocellum* gapDH promoter	This study
pHS0024_Y20F A138T	pNW33N derivative, CATsa_Y20F A138T under *C. thermocellum* gapDH promoter	This study
pHS0059	pHS0024 harboring CATec3 replacing with CATsa	This study
pHS0059_Y20F	CATec3 Y20F under *C. thermocellum* gapDH promoter	This study
pHS0070	pHS0059_Y20F with truncated GapDH promoter	This study
Strains		
E. coli		
Top10	Host for molecular cloning, *mcrA, Δ(mrr-hsdRMS-mcrBC), Phi80lacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(SmR), endA1, nupG*	Invitrogen
BL21 (DE3)	Host for plasmid *in vivo* methylation, *E. coli* B *dcm, ompT, hsdS*(rB-mB-), *gal*	Invitrogen
BL21-CodonPlus (DE3)- RIPL	Host for heterologous esterase expression, *E. coli B F- ompT hsdS(*rB- mB-*) dcm+ gal λ(DE3) endA Hte [argU proL] [argU ileY leuW],* Tet^R , Cm^R, $Strep/Spec^R$	Agilent

Table 4-1 Continued.

Name	Descriptions	Source
C. thermocellum		
M1354	*C. thermocellum* DSM1313 Δ*hpt*	[166]
HSCT1003	M1354 Δ*clo1313_0613*	This study
HSCT1004	M1354 Δ*clo1313_0693*	This study
HSCT2005	M1354 Δ*clo1313_0613* Δ*clo1313_0693*	This study
HSCT3009	M1354 Δ*clo1313_0613* Δ*clo1313_0693* Δ*clo1313_1160 (ldhA)*	This study
HSCT2105	HSCT2005 harboring pHS0024_F97W	This study
HSCT0102	M1354 harboring pHS0024_F97W	This study
HSCT2106	HSCT2005 harboring pHS0024_Y20F	This study
HSCT2107	HSCT2005 harboring pHS0059	This study
HSCT2108	HSCT2005 harboring pHS0059_Y20F	This study
HSCT2113	HSCT2005 harboring pHS0024_Y20F A138T	This study
HSCT3110	HSCT3009 harboring pHS0070	This study

Phylogenetic trees were built based on the aligned sequences using the maximum likelihood method with 1,000 bootstrap replicates. A 40% bootstrap confidence level cutoff was selected.

4.2.3 Genetic engineering of *C. thermocellum*

The primers used in this chapter are listed in Table 4-2. *C. thermocellum* cells were transformed by electroporation[58, 167]. Electroporation was performed outside of the anaerobic chamber followed by cell recovery and plating inside the chamber. A series of two consecutive exponential pulses were applied using the electroporation system (cat # 45-0651, BTX Technologies Inc., MA, USA) set at 1.8 kV, 25 μF, and 350 Ω, which usually resulted in a pulse duration of 7.0-8.0 ms. For markless gene deletions, 500 bp to 750 bp length of upstream (5') and downstream (3') fragments of the target gene were assembled by overlapping PCR and cloned in the MCS1 of pHS005. In the MCS2, 500 bp to 750 bp intermediate fragments (int) were cloned. The *dam* methylated plasmid was introduced to *C. thermocellum* and selected in CTFuD agar supplemented with Thm. A colony from successful transformation was selected and cultured in the liquid CTFuD-NY medium supplemented with Thm until the cell culture reached mid-log growth phase (OD ~ 0.2- 0.6). Cells were then serially diluted 100-fold and plated on a CTFuD-NY agar plate supplemented with Thm and 30 μg/mL 5-fluoro-2′-deoxyuridine (FuDR) followed by 3-5 days of incubation at 55℃. The colonies were PCR screened using the primers, HS252 and HS253 (Table 4-2), to check for plasmid loss. A colony without a PCR product was selected and streaked on the CTFuD-NY agar plate containing Thm and FuDR followed by incubation at 55℃ for 3-5 days. Colonies were PCR screened again to ensure plasmid loss and then cultured in the liquid CTFuD-NY medium without selection. The cultured cells were serial diluted 100-fold and plated on the CTFuD-NY agar supplemented with 500 μg/mL 8-azahypoxanthine (8-AZH).

Table 4-2: A list of primers used in this study. The restriction enzyme sites were underlined.

Primer name	Primer sequence (5' to 3')	Description
HS302	CTCTGGATCCAATGGCCCCGAATGAATACAAG	Clo1313_0693 cloning forward
HS230	CTCTGAGCTCTTAATTGGTTTTCAAATACTTCGC	Clo1313_0693 cloning reverse
HS304	CTCTGGATCCAATGGATGTGGTAATTACGTCAAACC	Clo1313_0521 cloning forward
HS234	CTCTGAGCTCTTACGGTACAGAGTTATAC	Clo1313_0521 cloning reverse
HS306	CTCTGGATCCAATGTCCCGGGAAGACGTTGTAGG	Clo1313_1680 cloning forward
HS238	CTCTGAGCTCTTACTGTTTTTTTATCTGTCTTC	Clo1313_1680 cloning reverse
HS229	CTCTGGATCCAATGGCACAATTATATGATATGC	Clo1313_0613 cloning forward
HS230	CTCTGAGCTCTTACGGATTCAGCTCATCC	Clo1313_0613 cloning reverse
HS303	CTCTGGATCCAATGGCAGCGGTCAATGCG	Clo1313_0500 cloning forward
HS232	CTCTGAGCTCTTATACAGGCAATTGTGTAATTG	Clo1313_0500 cloning reverse
HS305	CTCTGGATCCAATGACAAGAGAGCAGGAAAATCTC	Clo1313_1139 cloning forward
HS236	CTCTGAGCTCTCATTTCTCAAGTATAAGG	Clo1313_1139 cloning reverse
HS307	CTCTGGATCCAATGAAAACGATAAAAATCATGCCTG	Clo1313_1424 cloning forward
HS240	CTCTGAGCTCCTAATTTCTAAGAAGTATTC	Clo1313_1424 cloning reverse
HS186	CTCTGAGCTCAGAATTGCAACAAACATTTC	Clo1313_0613 deletion upstream forward
HS473	CCATATATAGCATTTTCAAAACCCCTTTTTATTC	Clo1313_0613 deletion upstream reverse
HS474	AAAAGGGGTTTTGAAAATG GTATATATGGATGCGGTTCG	Clo1313_0613 deletion downstream forward
HS475	CTCTGGATCCTTACGGATTCAGCTCATCCATAAG	Clo1313_0613 deletion downstream reverse
HS441	TTGTACACGGCCGCATAATCTGCAAATATTGAAGGGTG	Clo1313_0613 deletion intermediate forward
HS442	GCAGCCTAGGTTAATTAAGCTGCGCTTACGGATTCAGCTCATCCATAAGTATC	Clo1313_0613 deletion intermediate reverse

Table 4-2 Continued.

Primer name	Primer sequence (5' to 3')	Description
HS370	GGTTTACTGTACAGGGCTTC	Clo1313_0613 deletion check primer
HS371	GCTCGACTTGCATGTGTTAG	Clo1313_0613 deletion check primer
HS452	CTCTGAGCTCCCGGAAGATATGCCTGC	Clo1313_1160 deletion upstream forward
HS453	GAAAAGGAGAACTT ATTCCTTCCTCCTTTATCAATAAC	Clo1313_1160 deletion upstream reverse
HS454	ATAAAGGAGGAAGGAAT AAGTTCTCCTTTTCTTTTATGA	Clo1313_1160 deletion downstream forward
HS455	CTCTGGATCCATCTTGACAGTATTTCGGATTC	Clo1313_1160 deletion downstream reverse
HS456	TTGTACACGGCCGCATAATC GCCAACCAAAAAGAAGGCG	Clo1313_1160 deletion intermediate reverse
HS457	GCAGCCTAGGTTAATTAAGCTGCGC TCATATATCTAGTGTTTTTATTATTTC	Clo1313_1160 deletion intermediate reverse
HS83	AGCATCGGCATCATCTC	Clo1313_1160 deletion check primer
HS86	CAAGGGGTTTTCGTCGTG	Clo1313_1160 deletion check primer
HS476	CTCTGAGCTCGTGCGGAAAAGAAAATTAGG	Clo1313_0693 deletion upstream forward
HS477	TGATAACGTTATTCCCAAATTGGCCTGATGAA GCCATATTTC	Clo1313_0693 deletion upstream reverse
HS478	CATCAGGCCAATTTGGGAATAACGTTATCAGG GCAACAGC	Clo1313_0693 deletion downstream forward
HS479	CTCTGGATCCTTAATTGGTTTTCAAATACTTCG C	Clo1313_0693 deletion downstream reverse
HS480	TTGTACACGGCCGCATAATCGGAAGCAACGCT TGCATTATATATGCC	Clo1313_0693 deletion intermediate forward
HS481	GCAGCCTAGGTTAATTAAGCTGCGCTTGATTA AGATTTGCCACTATACC	Clo1313_0693 deletion intermediate reverse
HS53	ATTGTGTATCATTTGCATAAACAATG	Clo1313_0693 deletion check primer
HS56	GTTTTTGTCCAGCCACTTAAGCAC	Clo1313_0693 deletion check primer
HS252	TTTTACGGTGTAAACCTTCCTC	Plasmid ori RepB check primer
HS253	AAAAGAAAATCGCTAATGTTGATTAC	Plasmid ori RepB check primer

After three days, colonies were PCR screened using the primers binding at genome-specific upstream and downstream sites of the target knockout gene. A colony was streaked on a new CTFuD-NY plate up to three transfers to isolate a single genotype. Finally, the markerless gene deletion was confirmed by negative cell growth in CTFuD medium supplemented with Thm.

4.2.4 Media and cultivation

E. coli strains were grown in lysogeny broth (LB) medium supplemented with 30 μg/mL chloramphenicol (Cm) and/or 50 μg/mL kanamycin (Kan) when appropriate. All *C. thermocellum* strains were cultivated in an anaerobic chamber (Sheldon manufacturing, OR, USA) with a gas mixture (90% N_2, 5% CO_2, 5% H_2) or rubber stopper sealed Balch tubes outside the chamber. For genetic engineering, *C. thermocellum* strains were grown in rich media, CTFuD or CTFuD-NY[167] supplemented with 10 μg/mL thiamphenicol (Thm) as needed. The CTFuD medium contained 2 g/L yeast extract while CTFuD-NY contained vitamins and trace elements[167] without yeast extract. Except for isobutyl acetate production from cellulose, *C. thermocellum* strains were grown in a defined MTC medium with 5 g/L cellobiose as previously described [173]. To produce isobutyl acetate from cellulose, a modified MTC medium (C-MTC) was used as previously described [58]. The C-MTC medium contained (per liter): 2 g urea, 1.5 g ammonium chloride, 10 g 3-(N-morpholino)propanesulfonic acid (MOPS), 2 g sodium citrate tribasic dihydrate, 1.5 g citric acid monohydrate, 1 g sodium sulfate, 0.5 g potassium phosphate monobasic, 1 g cysteine-HCl, 0.2 g calcium chloride, 1 g magnesium chloride hexahydrate, 0.1 g iron (II) chloride tetrahydrate, 2.5 g sodium bicarbonate, 0.02 g pyridoxamine dihydrochloride, 0.004 g p-aminobenzoic acid, 0.002 g biotin, and 0.002 g vitamin B12. The medium was adjusted to pH of 7.5 and autoclaved for

sterilization. Solid LB and CTFuD media additionally contained 15 g/L and 10 g/L of agar, respectively.

4.2.5 Esterase and thioesterase assay

To screen esterase activity towards isobutyl acetate, seven putative esterases from *C. thermocellum* were heterologously expressed in *E. coli* BL21 CodonPlus-RIPL(DE3) strains under a control of T7lac promoter. The recombinant *E. coli* strains were cultured overnight in 3 mL of LB medium at 37℃, transferred to 25 mL of fresh LB medium with 100-fold dilution, and cultured at 37℃ until OD reached 0.3~0.4. Then, protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) at final concentration of 0.1 mM. The IPTG-induced cells were cultured in a water bath orbital shaker (MaxQ7000, Thermo Fisher Scientific, MA, USA) at 19℃ and 200 rpm up to 16 hours and harvested by centrifugation at 4,700 xg for 5 minutes. The cell pellets were washed with Milli-Q water once and suspended in a B-PER complete bacterial protein extraction reagent (cat# 89821, Thermo Fisher Scientific, MA, USA) to disrupt cell wall. Crude cell extract was collected by 5 minutes of centrifugation at 17,000 xg for further experiments. Deacetylation activity of the seven O-deacetylases was screened by a colorimetric pNPA assay on a 96-well plate. The reaction buffer contained 50 mM Tris-HCl at pH 6.8 and 10 mM pNPA in 200 μL of total reaction volume. The reaction started by adding crude cell extracts at 2X, 4X, and 8X dilutions and then an absorbance at 405nm was monitored in a microplate reader at 50℃ for 10 minutes.

The 6xHis tagged enzymes were purified by Ni-NTA agarose (cat# 25214, Thermo Fisher Scientific, MA, USA) using a mini spin column (cat# 6572, BioVision Inc., CA, USA), based on the manufacturer's recommendation. The eluted enzyme was desalted by a centrifugal filter unit

(cat# C7719, MilliporeSigma, MA, USA) before the enzyme reaction. The reaction assay contained 20 mM Tris-HCl at pH 7.4 and 20 mM of isobutyl acetate in 1 mL total volume. The reaction started by adding 5 μg/mL of the His-tag purified enzymes, followed by 10 hours incubation at 55℃ in an oven. pH change was measured with a pH meter (Hach, CO, USA) before and after the reaction to screen for esterase activity. In addition, HPLC analysis was used to verify the hydrolysis of isobutyl acetate into isobutanol and acetate.

Thioesterase activities of Clo1313_0613 and Clo1313_0693 were measured by DTNB assay using isobutyl-CoA as a substrate[94]. The reaction solution consisted of 50 mM Tris-HCl at pH 6.8, 0.5 mM of isobutyl-CoA, and 1 mg/mL DTNB in 100% DMSO in 50 μL of total reaction volume. 10 μg/mL of the His-tag purified enzymes were added to the reaction solution and immediately incubated at 50℃ in a plate reader followed by kinetic measurement at 412nm for 30 mins. The reaction rate and specific activity were calculated using the extinction coefficient from a standard curve of free coenzyme A (MP Biomedicals, OH, USA) under the same condition.

4.2.6 Proteomics

For proteomics, cell cultures were sampled at 48h in the middle of stationary growth phase and stored in -80℃ before the analysis. Two milliliters of a slurry containing *C. thermocellum* cells growing on Avicel was pelleted by centrifugation (5000 x g for 10 min), supernatant discarded, and pellet resuspended in 350 μl of lysis buffer (4% SDS, 5 mM DTT, 100 mM Tris-HCl, pH 8.0). Cells were lysed by sonication (Branson Sonifier; 20% amplitude, 2 s pulses, 30 s total), precleared by centrifugation (21000 x g for 10 min) and crude protein supernatant with a Nanodrop OneC spectrophotometer (Thermo Scientific). Samples were then adjusted 15 mM iodoacetamide (20 min at room temperature in the dark) and 300 μg cleaned up via protein aggregation capture[174].

Aggregated protein (on 1-micron magnetic Sera-Mag beads; GE Healthcare) was then digested with proteomics-grade trypsin (1:75 w/w; Promega) in 100 mM Tris-HCl, pH 8.0 overnight at 37°C, and again for 3 h at 37°C the following day. Released tryptic peptides were then acidified to 0.5% formic acid, filtered through a 10 kDa MWCO spin filter (Vivaspin 2; Sartorius), and quantified by Nanodrop OneC. Three micrograms of peptides were then analyzed by 1D LC-MS/MS using a Vanquish uHPLC coupled directly to an Orbitrap Q Exactive mass spectrometer (Thermo Scientific) as previously described [175]. Peptides were separated across a 180 min organic gradient using an in-house pulled nanospray emitter packed with 15 cm of 1.7-micron Kinetex reversed-phase resin (Phenomenex). Peptide fragmentation spectra were analyzed/sequenced by Proteome Discoverer software (Thermo Scientific) and peptides quantified by chromatographic area-under-the-curve. Peptide abundances were log2 transformed, distributions normalized by LOESS, and rolled up to their respective proteins via RRollup InfernoRDN[176]**.** Protein abundance distributions were then median centered and statistical analyses performed in Perseus[177]. Manual determination of CATec3 abundance across strains was performed by comparing log2 peptide abundance trends across orthologs as only one AA difference differentiate the two protein forms. All LC-MS/MS raw data have been deposited into the MassIVE and ProteomeXchange repositories with the following accession numbers: MSV000086201 (MassIVE) and PXD021695 (ProteomeXchange). Data can be downloaded directly via ftp://massive.ucsd.edu/MSV000086201/

4.2.7 Conversion of alcohols to esters in *C. thermocellum*

Tube-scale cellulose fermentation was performed in the batch mode. Briefly, 19 g/L of Avicel PH-101 was used as a sole carbon source in a 16 mL culture volume. 0.8 mL of overnight cell culture

was inoculated in 15.2 mL of C-MTC medium, and 4 mL hexadecane was added in the anaerobic chamber. Each tube contained a small magnetic stirrer bar to homogenize cellulose, and the culture was incubated in a water bath connected with a temperature controller and a magnetic stirring system. Alcohols were fed to the culture at 36h time point when cells entered early stationary growth phase. pH was adjusted to between 6.4 and 7.8 with 5 M KOH injection.

4.2.8 Fermentative ester production

Ester production from cellobiose in *C. thermocellum* strains was performed by the two-step bioconversion configuration. Cells were first cultured in MTC minimal medium[167] containing 5 g/L cellobiose in a rubber capped Balch tube until OD reached 0.8~1.0. The cells were cooled down at room temperature for 20 minutes and centrifuged at 4,700 x g and 4℃ for 20 minutes. After removing the supernatant, cells were resuspended in the same volume of fresh MTC minimal media containing 2 g/L isobutanol in an anaerobic chamber. The cell suspension was then divided into 800 μL in a 2.0 mL screw cap microcentrifuge tube with a 200 μL hexadecane overlay. The cells were incubated at 55℃ for 24 hours followed by analysis of gas chromatography coupled with a mass spectrometer (GC/MS) to quantify the amount of isobutyl acetate produced.

For the cellulose fermentation, modified MTC medium (C-MTC medium) was used. 20 g/L of Avicel PH-101 was used as a sole carbon source instead of cellobiose, and 10 g/L of MOPS was added to increase buffer capacity. Initial pH was adjusted to 7.5 by 5M KOH and autoclaved. In an anaerobic chamber, 0.8 mL of overnight cell culture was inoculated in 15.2 mL of C-MTC medium (1:20 inoculation ratio) with 4 mL of overlaid hexadecane. Each tube contained a small magnetic stirrer bar to homogenize cellulose. The rubber capped Balch tube was incubated in a water bath connected with a temperature controller set at 55℃ and a magnetic stirring system.

Following pH adjustment with 70 μL of 5 M KOH injection, 800 μL of cell culture and 200 μL of hexadecane layer were sampled every 12 hours. Culture pH was maintained within a range of 6.4-7.8 during the fermentation.

Cell growth was monitored by measuring pellet protein. The cell-cellulose pellet from 800 μL sampling volumes was washed twice with Milli-Q water and suspended by 200 μL lysis buffer (0.2 M NaOH, 1% SDS) followed by an hour incubation at room temperature. Then, the solution was neutralized with 50 μL 0.8 M HCl and diluted by 550 μL water. The mixture was centrifuged at 17,000 x g for 3 minutes. Protein concentration from the supernatant was analyzed by the detergent-compatible Bradford assay (Thermo Scientific, WA, USA). The residual pellet was boiled in a 98℃ oven for an hour before quantifying residual cellulose.

Residual cellulose was quantified by the phenol-sulfuric acid method[178] with some modifications. The boiled sample was washed twice with Milli-Q water and suspended in 800 μL water to make equivalent volume to the original. The sample was homogenized by pipetting and vortexing for 10 seconds, and 20 μL of the homogenized sample was transferred to a new 2.0 mL microcentrifuge tube or 96-well plate and dried overnight in a 55℃ oven. The dried pellet was suspended in 200 uL of 95% sulfuric acid and incubated for an hour at room temperature. After the pellet was dissolved completely, 20 μL of 5% phenol was added and mixed with the sulfuric acid solution. After 30 min incubation at room temperature, 100 μL of the sample was transferred to a new 96-well plate, and the absorbance at 490 nm was measured. The absorbance was converted to cellulose concentration by the standard curve of Avicel PH-101 treated by the same procedure.

4.2.9 Analytical methods

HPLC system (Shimadzu Inc., MD, USA) was used to quantify extracellular metabolites. 800 μL of samples were centrifuged at 17,000 x g for 3 minutes followed by filtering through 0.2 micron filters. The samples were run with 5 mM H_2SO_4 at 0.6 mL/min on an Aminex HPX-87H (Biorad Inc., CA, USA) column at 50℃. Refractive index detector (RID) and ultra-violet detector (UVD) at 220 nm were used to determine concentrations of sugars, organic acids, and alcohols. GC (HP 6890, Agilent, CA, USA) equipped with a MS (HP 5973, Agilent, CA, USA) was used to quantify esters. A Zebron ZB-5 (Phenomenex, CA, USA) capillary column (30 m x 0.25 mm x 0.25 μm) was used with helium as the carrier gas at a flow rate of 0.5 mL/min. The oven temperature program was set as follows: 50℃ initial temperature, 1℃/min ramp up to 58℃, 25℃/min ramp up to 235℃, 50℃/min ramp up to 300℃, and 2-minutes bake-out at 300℃. 1 μL sample was injected into the column with the splitless mode at an injector temperature of 280°C. For the MS system, selected ion mode (SIM) was used to detect and quantify esters. As an internal standard, 10 mg/L n-decane were added in initial hexadecane layer and detected with m/z 85, 99, and 113 from 12 to 15-minute retention time range.

4.3 Disruption of carbohydrate esterases hydrolyzing isobutyl acetate from *C. thermocellum*

4.3.1 Endogenous esterase activities as a bottleneck for CBP ester production

Ester production in cellulolytic bacteria such as *C. thermocellum* has never studied before and bottlenecks for CBP ester production is widely unknown. One of the potential bottlenecks is ester degradation caused by endogenous esterases. For example, by disrupting an esterase that

hydrolyzes volatile esters from *Saccharomyces cerevisiae*[179, 180], biosynthesis of medium chain esters was increased [181], suggesting that balance of AAT and esterase activity of brewer yeasts is important for the ester accumulation in wine and distillates [181-183]. Cellulolytic bacteria including *C. thermocellum* possess multiple carbohydrate esterases (CEs) whose expression helps deconstruct lignocellulosic biomass in a concerted enzyme reaction [184-186]. However, it is widely unknown which CEs hydrolyze the medium chain esters. Therefore, identification of the CEs degrading the medium chain esters is important to develop an efficient CBP platform for ester biosynthesis from lignocellulosic biomass.

The esterase- and AAT-dependent pathways have different reaction chemistry for ester biosynthesis (Figure 4-1A, B). The pathway driving force depends on the equilibrium constant and species concentrations. For isobutyl acetate biosynthesis by an esterase in an aqueous solution, the equilibrium constant (K_{eq} =1.3x10^{-5} 1/M) at the standard condition (T = 25℃, P = 1 atm, pH 7, [C] = 1M) is relatively low, yielding a positive standard ΔG^o of 27.8 ± 4.2 kJ/mol[187]. This reaction is not thermodynamically favorable for ester biosynthesis but for degradation (or ester hydrolysis)[188], and therefore, relatively high substrate concentrations are required to push the reaction forward (Table 4-3). In contrast, the AAT-dependent pathway is thermodynamically favorable with K_{eq} of 2.42 1/M and a negative standard ΔG^o of -2.4 ± 4.1 kJ/mol[187], hence only relatively low substrate concentrations are needed for ester biosynthesis to occur. In short, esterases play a critical role in ester degradation under aqueous fermentation conditions while AATs are essential for ester biosynthesis[182, 183, 189, 190].

It is challenging to determine which CEs cause ester degradation because *C. thermocellum* possesses a large number of CEs to deacetylate O-acetyl groups of hemicelluloses critical for lignocellulosic biomass deconstruction[184-186] and the functional roles of these CEs responsible for

ester degradation are largely unknown. Here, we aimed to identify which endogenous CEs of *C. thermocellum* could be disrupted to reduce ester degradation while not interfering biomass deconstruction.

4.3.2 Putative carbohydrate esterases degrading isobutyl acetate

To identify endogenous CEs of *C. thermocellum* hydrolyzing isobutyl acetate, we thoroughly searched the CAZy database[170]. We found *C. thermocellum* has 14 CEs that are classified as either N-deacetylase or O-deacetylase, based on the domains and putative substrates of CEs (Table 4-4). Since isobutyl acetate is an O-acetyl ester that can be deacylated by an O-acetyl esterase, we focused on investigating the seven putative O-acetyl esterases, Clo1313_1424, Clo1313_1139, Clo1313_1160, Clo1313_0521, Clo1313_0613, Clo1313_0500, and Clo1313_0693. Based on the sequence motifs, only Clo1313_0613 is a cytosolic enzyme while others are likely secreted and/or anchored to the cell membrane (Figure 4-2A). To characterize these seven CEs for ester hydrolysis, we cloned and heterologously expressed them in *E. coli*. We first removed the signal peptides of these proteins except Clo1313_0613 and tagged all proteins with 6x His residues at their N-termini (Figure 4-2A). For Clo1313_1424 that has two identical DNA sequences encoding a Xyn_B like domain, we built a truncated version of Clo1313_1424 (Clo1313_1424t) with only one Xyn_B like domain and without the N-terminus signal peptide for expression in *E. coli*. Successful expression of the seven modified esterases in *E. coli* was confirmed as shown in Figure 4-2B.

4.3.3 Clo1313_0613 and Clo1313_0693 are potential deletion targets

To evaluate the CE activities for ester hydrolysis, we incubated isobutyl acetate with the purified enzymes at 55℃ for 10 hours and monitored the pH change.

Table 4-3: Gibbs free energy change (ΔG, KJ/mol) for isobutyl acetate biosynthesis by the esterase- and AAT-dependent pathways. Substrate concentrations are varied while the concentration of isobutyl acetate is fixed to be 10^{-6} mM.

Pathways	**Isobutanol (mM)**	**Acetic acid (mM)**	**ΔG (kJ/mol)**
Esterase	0.01	1	22.1 ± 4.2
	0.1	1	16.4 ± 4.2
	1	1	10.7 ± 4.2
	100	1	-0.7 ± 4.2
	100	10	-12.2 ± 4.2
Alcohol acyltransferase (AAT)	**Isobutanol (mM)**	**Acetyl-CoA (mM)**	**ΔG (kJ/mol)**
	0.01	0.01	-47.8 ± 4.1
	0.1	0.01	-53.5 ± 4.1
	1	0.01	-59.2 ± 4.1
	10	0.01	-64.9 ± 4.1
	100	0.01	-70.7 ± 4.1

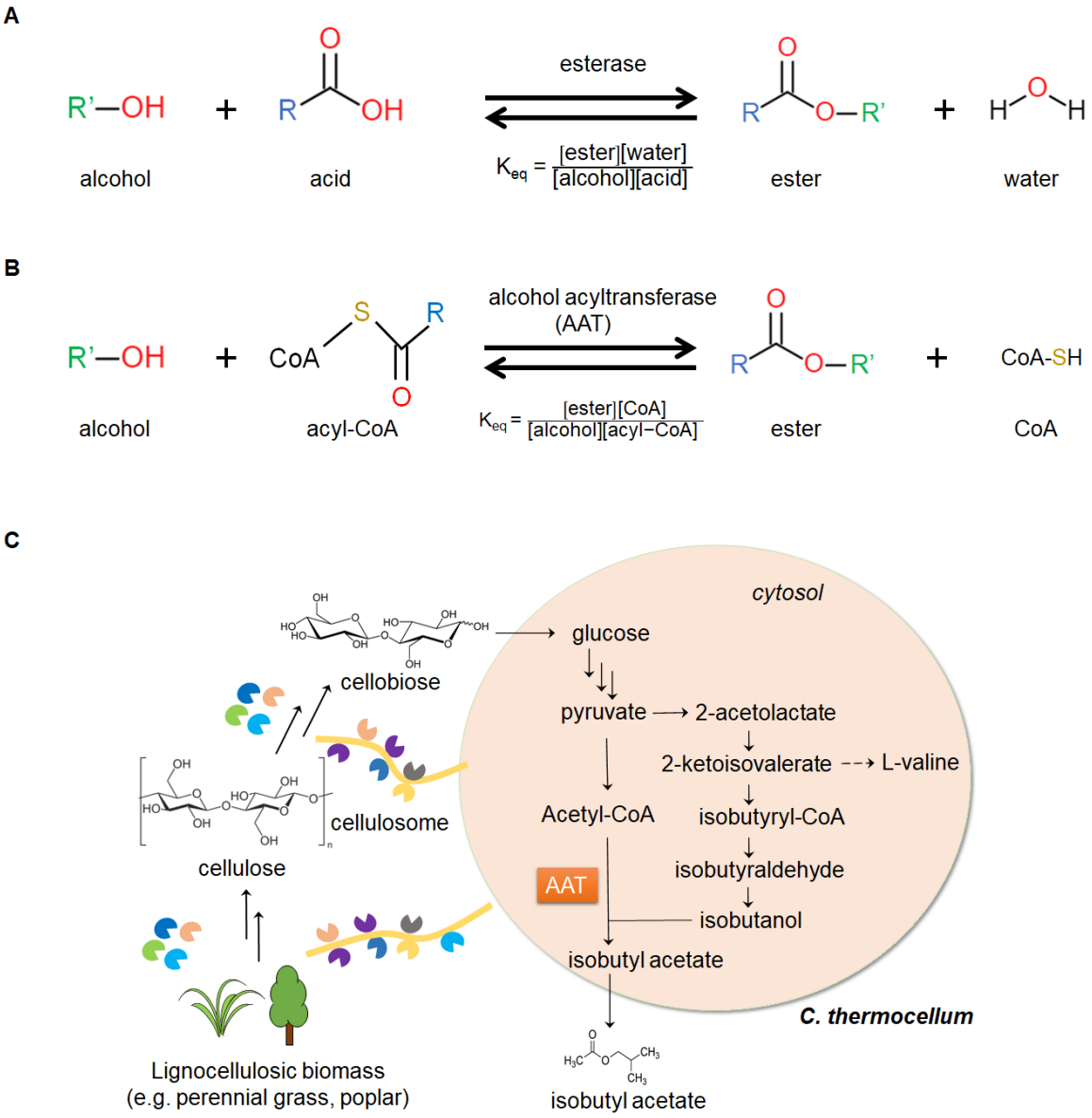


Figure 4-1: Two representative ester biosynthesis pathways and metabolic pathway of *C. thermocellum* for isobutyl acetate synthesis. **(A)** A reaction scheme of esterase/lipase associated ester synthesis pathway and equation of equilibrium constant. Water is necessarily involved in the reaction. **(B)** A reaction scheme of alcohol acyltransferase (AAT) and equation of equilibrium constant. **(C)** A scheme of isobutyl acetate production from lignocellulose via AAT pathway in *C. thermocellum*.

Table 4-4: Carbohydrate esterases (CEs) in *C. thermocellum*.

Class	Locus tag and description
CE-1	Clo1313_2635, glycoside hydrolase family 10, putative N-deacetylase, endoglucanase included Clo1313_1305, glycoside hydrolase family 10, putative N-deacetylase, endoglucanase included Clo1313_2858, carbohydrate binding family 6, putative N-deacetylase, endoglucanase included
CE-2	Clo1313_1425, glycoside hydrolase family 5, endoglucanase included
CE-3	Clo1313_1424, lipolytic protein G-D-S-L family, putative O-deacetylase
CE-4	Clo1313_1139, glycoside hydrolase family 11, putative N-, O-deacetylase Clo1313_1680, delta-lactam-biosynthetic de-N-acetylase, putative N-, O-deacetylase Clo1313_0521, glycoside hydrolase family 11, putative N-, O-deacetylase
CE-7	Clo1313_0613, cephalosporin-C deacetylase, putative O-deacetylase
CE-8	Clo1313_0500, pectinesterase, putative O-deacetylase
CE-9	Clo1313_2853, N-acetylglucosamine-6-phosphate deacetylase, putative N-deacetylase
CE-12	Clo1313_0693, carbohydrate binding family 6, putative O-deacetylase
CE-14	Clo1313_2896, GlcNAc-PI de-N-acetylase, putative N-deacetylase
CENC	Clo1313_0420, dockerin type 1

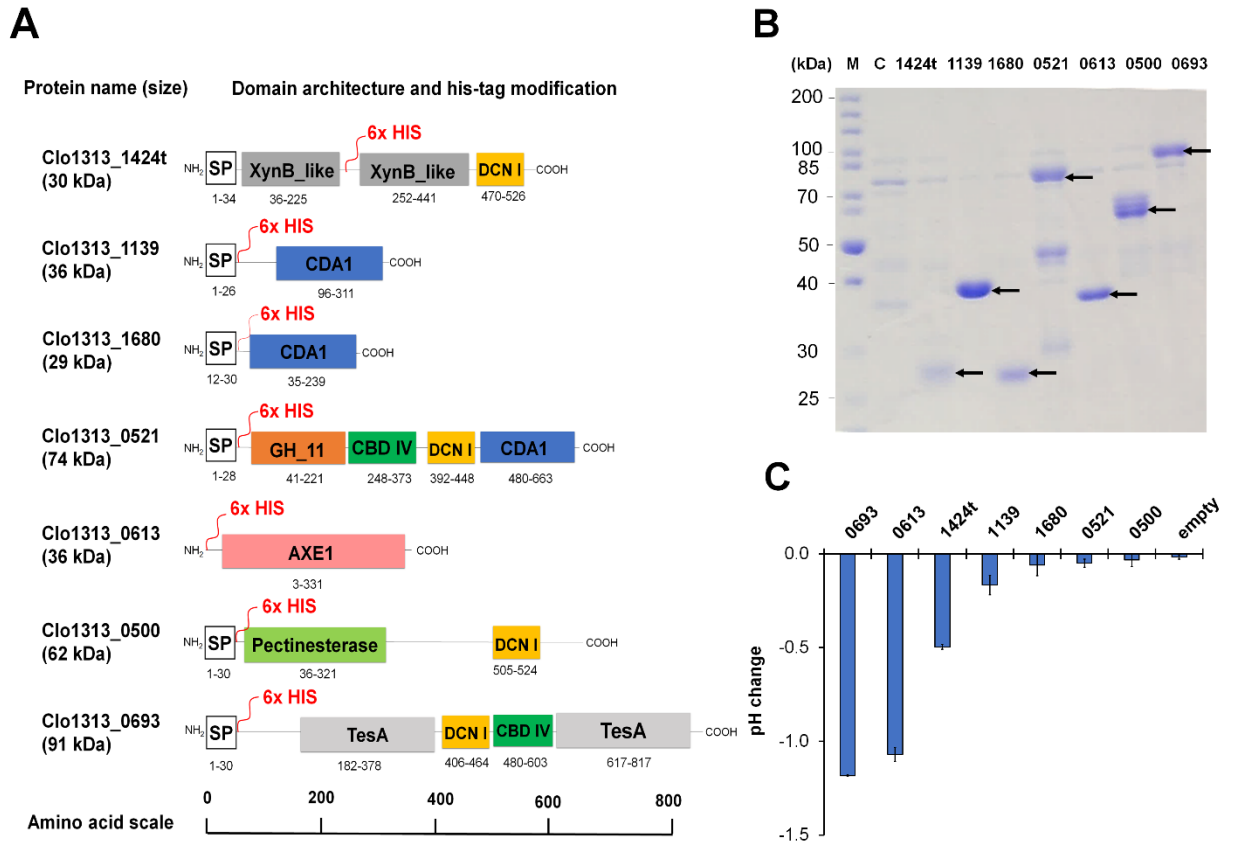


Figure 4-2: Identification of esterases responsible for isobutyl acetate hydrolysis in *C. thermocellum*. **(A)** Seven O-deacetylases in *C. thermocellum* and their domain architectures. The locations tagged with the histidine residues were marked as 6xHis in red. Clo1313_1424t is a truncated version of Clo1313_1424 which is derived from removal of the signal peptide and one XynB_like domain located at 36-225 residues that is identical to the other located at 252-441 residues. Abbreviation − SP: signal peptide, DCN: dockerin, CDA: carbohydrate deacetylase family, GH: glycosyl hydrolases family, CBD: cellulose binding domain, AXE: acetyl xylan esterase, and TesA: thioesterase A (lysophospholipase L1 or related esterase). **(B)** SDS-PAGE of purified His-tagged proteins expressed in the recombinant *E. coli* strains. Abbreviations − M: protein marker ladder and C: *E. coli* harboring an empty backbone plasmid. **(C)** *In vitro* screening of isobutyl acetate hydrolysis activities by endogenous CEs of *C. thermocellum*. The error bars represent standard deviation of three biological replicates.

If the enzymes hydrolyze isobutyl acetate, pH of the reaction solution will decrease because hydrolysis of isobutyl acetate (pKa ~25) by an esterase into isobutanol and acetic acid (pKa ~4.75) lowers pH (Figure 4-2C). The screening discovered three most active CEs, Clo1313_0693, Clo1313_0613, and Clo1313_1424 (Figure 4-2C). The hydrolysis was further confirmed with detection of isobutanol and acetate by HPLC analysis (data not shown). The other CEs showed insignificant pH changes, suggesting that they might have little or no activity towards isobutyl acetate. Taken together, we identified the three esterases, Clo1313_0613 and Clo1313_0693, and Clo1313_1424 that can hydrolyze isobutyl acetate.

Since CEs are important for lignocellulose deconstruction, it is critical to eliminate CEs that alleviate ester hydrolysis but do not affect lignocellulose fermentation in *C. thermocellum*. The key question is which of the three esterases identified above should be chosen for disruption in *C. thermocellum.* Bioinformatic analysis revealed that Clo1313_0613 (cephalosporin C deacetylase) shares 59% protein sequence identity with a well-studied homo-hexamer cephalosporin C deacetylase from *Bacillus subtilis*. Similarly, the homology model of Clo1313_0613 was also presented as a homo-hexamer structure. This deacetylase has a broad substrate range with high specificity towards short substrates including p-nitrophenyl acetate (pNPA) and short-chain or pre-processed xylans[191]. Our *in vitro* assay indeed confirmed that Clo1313_0613 exhibited activity towards pNPA (Figure 4-3). Based on the constructed phylogenetic tree of Clo1313_0613, we found homolog proteins exist in a wide spectrum of bacteria, regardless of their cellulolytic capability (Figure 4-4A). Most of the bacteria identified are gram-positive including *Actinomyces sp.* while a small group of gram-negative bacteria such as *Mesorhizobium sp* are also found. Taken together, Clo1313_0613 was expected to be a non-detrimental deletion target.

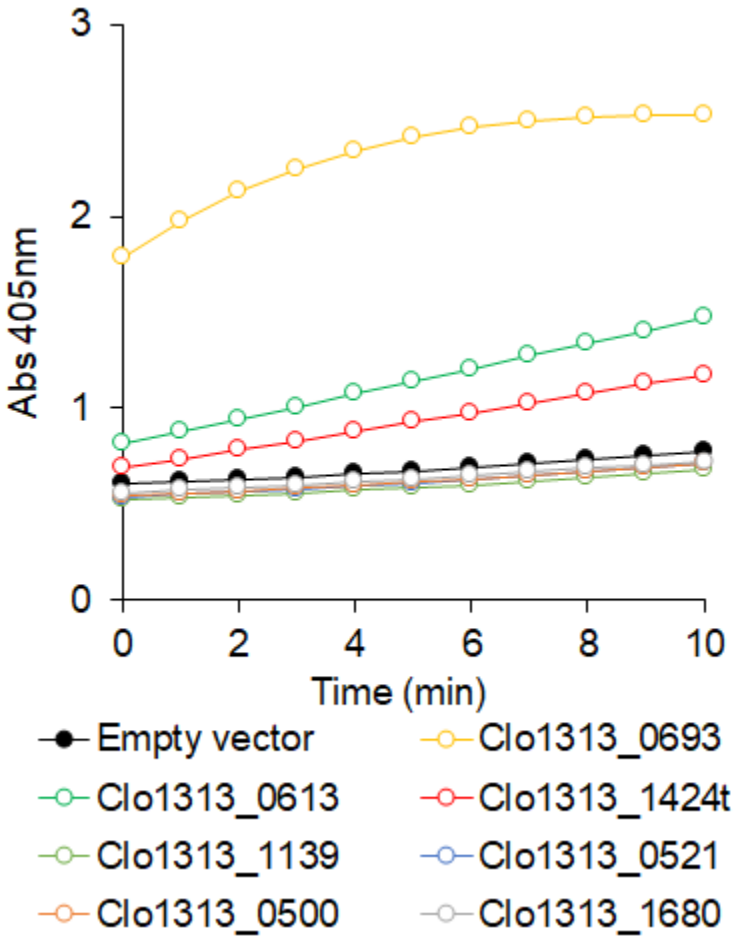


Figure 4-3: p-nitrophenyl (pNPA) assay of the seven O-acetyl esterases.

The catalytic domain of Clo1313_0693, annotated as a carbohydrate binding family 6, shares the highest sequence similarity (68%) with an uncharacterized protein YxiM from *B. subtilis* (locus tag: BSU_39120). YxiM has not yet been characterized experimentally, but its homolog proteins deacetylate broad substrates including apple pectin rhamnogalacturonan[192], pNPA[192, 193], and cephalosporin C[193]. Clo1313_0693 possesses two rhamnogalacturan acetylesterase-like domains that are likely responsible for the hydrolysis activity of isobutyl acetate as experimentally observed (Figure 4-3C). The gene encoding Clo1313_0693 is located downstream of a promoter regulated by the σ^{I3} RNA polymerase[194, 195], which is related to a membrane-associated anti-sigma factor (Rsigl3) that specifically binds to pectin[196]. Another pectinesterase under the control of σ^{I3} regulon is Clo1313_1983, a rhamnogalacturonan lyase[195]. Note that Clo1313_0693 includes a signal peptide and dockerin (Figure 4-2A), which is proposed to be a component of extracellular cellulosome[197]. Interestingly, we found that homolog proteins broadly exist in both cellulolytic and non-cellulolytic bacteria (Figure 4-4B), suggesting that Clo1313_0693 might not be detrimental for lignocellulose fermentation.

Lastly, the catalytic domain of Clo1313_1424 shares 30% amino acid sequence identity with a well-studied XynB domain of xylanase from *Ruminococcus flavefaciecns*[198, 199]. The XynB deacetylates xylan that is important for hemicellulose deconstruction[200]. Even though wild type *C. thermocellum* is not capable of assimilating C5 sugars for growth, it possesses enzymes to deconstruct xylan to enhance accessibility of cellulolytic enzymes to the recalcitrant lignocellulosic biomass[201]. Remarkably, most of the homologs of XynB were observed in cellulolytic microorganisms (Figure 4-4C). Therefore, we reasoned that Clo1313_1424 is not a good target for disruption to alleviate ester hydrolysis in *C. thermocellum* due to its functional role important for lignocellulose deconstruction.

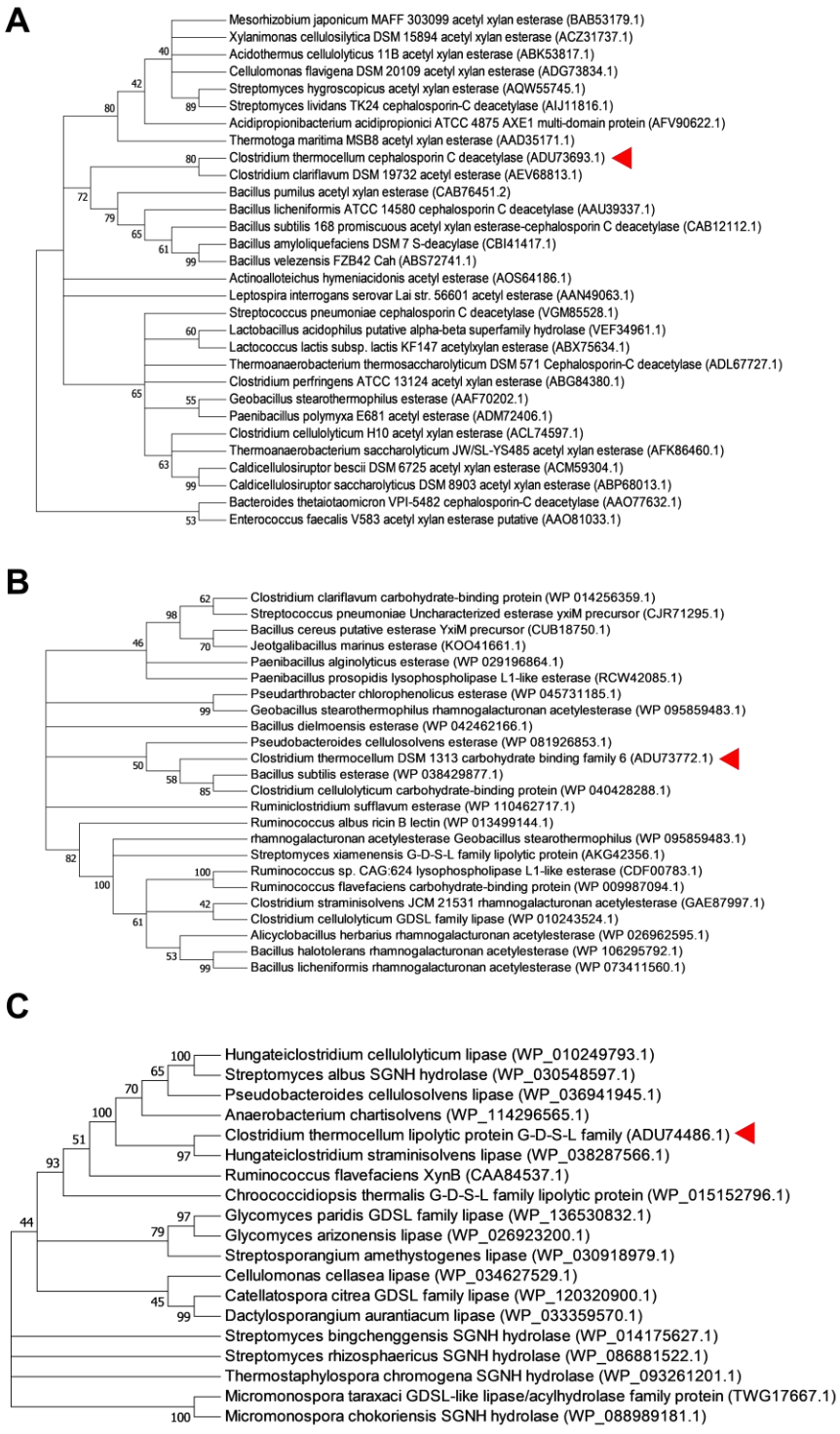


Figure 4-4: Phylogenetic analysis of Clo1313_0613, Clo1313_0693, and Clo1313_1424. **(A)** A phylogenetic tree of AXE1 domain in Clo1313_0613. The red triangle indicates Clo1313_0613. **(B)** A phylogenetic tree of TesA domain in Clo1313_0693. The red triangle indicates Clo1313_0693. **(C)** A phylogenetic tree of the catalytic domain in Clo1313_1424. The red triangle indicates Clo1313_1424.

4.3.4 Disruption of *clo1313_0613* and *clo1313_0693* is not detrimental to cellulose fermentation capability of *C. thermocellum*

To genetically disrupt the targeted esterases, we first constructed a markerless gene deletion platform that utilizes two counter-selection genetic parts[167]. pHS005 (Figure 4-5) was constructed by assembling the required genetic parts using pNW33N as a backbone plasmid[58]. The markless gene deletion platform pHS005 contains two multiple cloning sites (MCS) that allows use of restriction enzyme and ligation for constructing deletion plasmids (Figure 4-5B).

Transcriptomic analysis of *C. thermocellum* shows that both Clo1313_0613 and Clo1313_0693 are expressed in *C. thermocellum*[197, 202]. To test their *in vivo* activities for isobutyl acetate degradation, we disrupted these esterase genes through homologous recombination. We generated three esterase-deficient strains including HSCT1003 with Clo1313_0613 deletion, HSCT1004 with Clo1313_0693 deletion, and HSCT2005 with both Clo1313_0613 and Clo1313_0693 deletion (Figure 4-6A). HSCT1003 (0.271 ± 0.017 1/h) and HSCT2005 (0.270 ± 0.031 1/h) grew about 20% slower than the parent strain (M1354, 0.341 ± 0.013 1/h) while HSCT1004 (0.352 ± 0.018 1/h) did not exhibit significant change in growth rate (Table 4-5). Interestingly, the cellulose degradation profiles were similar among the engineered strains, with the similar growth (Figure 4-6B, Table 4-5). Therefore, we concluded that the esterase disruptions did not inhibit cellulose consumption in *C. thermocellum*.

4.3.5 Alleviated isobutyl acetate degradation by the esterase deficient *C. thermocellum*

We next compared the hydrolysis of isobutyl acetate among the parent and esterase-deficient strains (Figure 4-6C). To estimate the *in vivo* esterase activity, we externally supplemented 1 g/L

Table 4-5: Specific growth rate of *C. thermocellum* strains. Cells were cultured in the defined MTC medium with 5 g/L of cellobiose at 55^{o}C. The values are presented as mean ± 1 standard deviation of six biological replicates.

Strains	Specific growth rate (1/h)
M1354 (Δhpt)	0.341 ± 0.013
HSCT1003 (Δhpt Δ0613)	0.271 ± 0.017
HSCT1004 (Δhpt Δ0693)	0.352 ± 0.018
HSCT2005 (Δhpt Δ0613 Δ0693)	0.270 ± 0.031

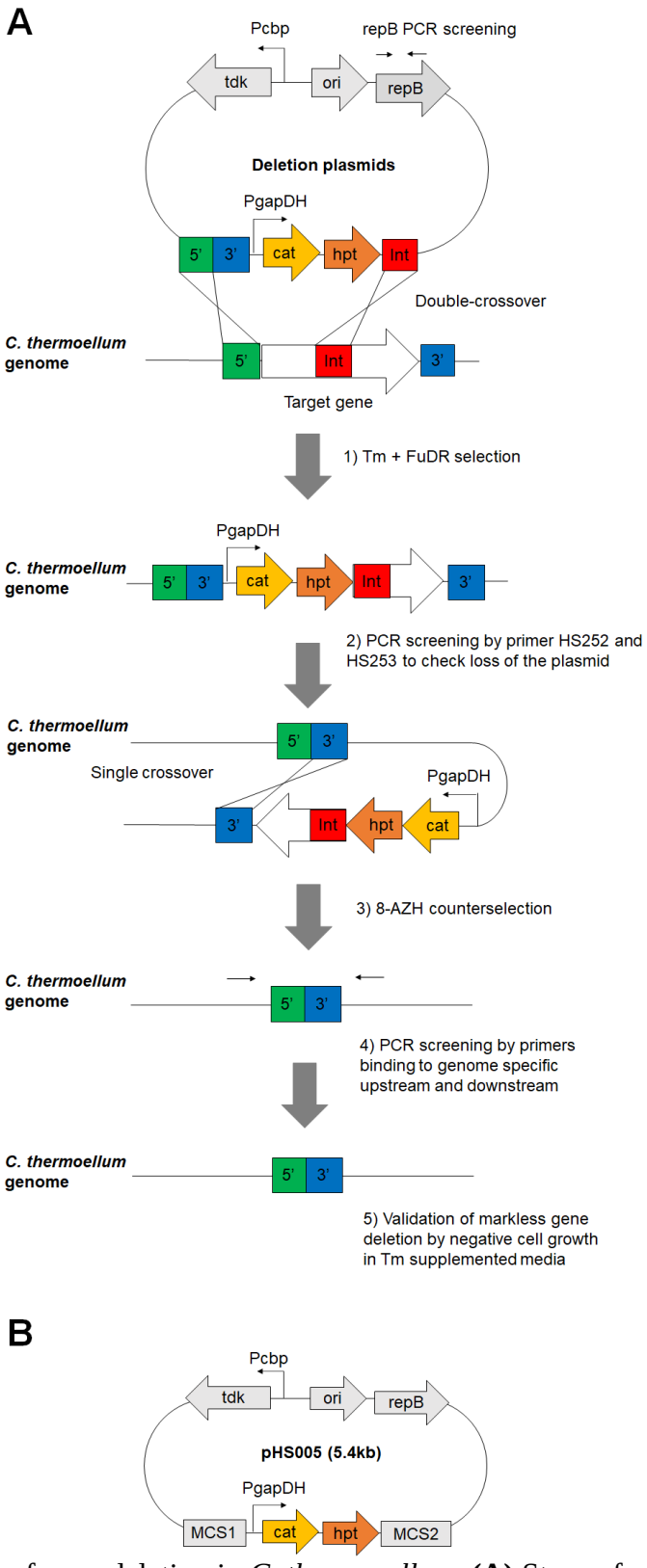


Figure 4-5: Schematic of gene deletion in *C. thermocellum*. **(A)** Steps of markless gene deletion. **(B)** Structure of pHS005 for markerless gene deletion.

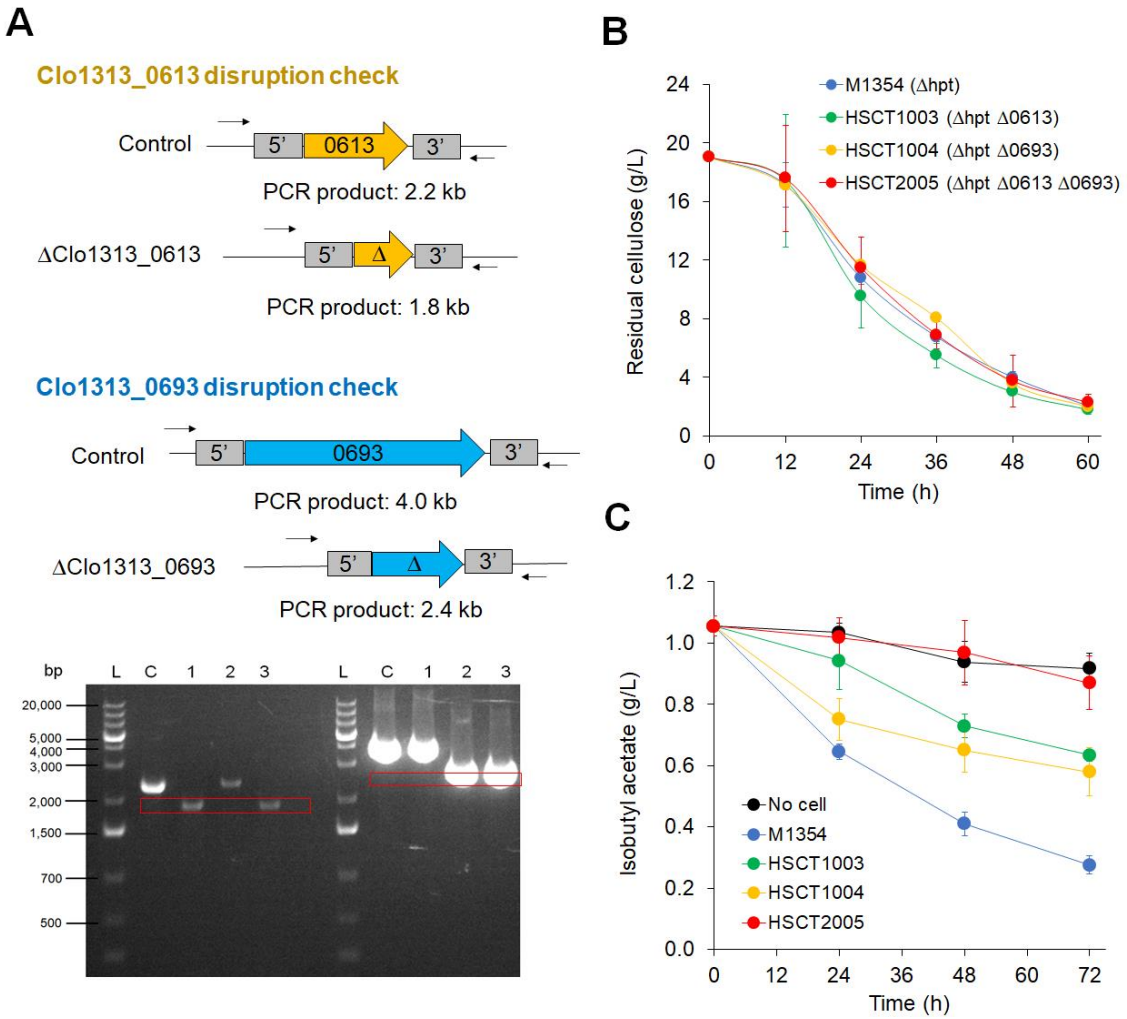


Figure 4-6: Disruption of esterases in *C. thermocellum* and characteristics of the esterase-deficient mutants. **(A)** Confirmation of gene disruption. Abbreviation − L: DNA marker ladder, C: control strain (M1354), 1: HSCT1003 (ΔClo1313_0613), 2: HSCT1004 (ΔClo1313_0693), and 3: HSCT2005 (ΔClo1313_0613, ΔClo1313_0693). **(B)** Comparison of residual cellulose by the engineered *C. thermocellum* strains during 20 g/L cellulose fermentation. The error bar represents standard deviation of three biological replicates. **(C)** Hydrolysis of isobutyl acetate by the engineered strains. The error bars represent standard deviation of three biological replicates.

(8.61 mM) of isobutyl acetate and monitored the changes in ester concentrations by a gas chromatography coupled with mass spectrometry (GC/MS). After 72 h, the parent strain M1354 degraded 74.0 ± 2.4% of isobutyl acetate while HSCT1003 and HSCT1004 exhibited only 40.0 ± 1.5% and 45.1 ± 8.4% of isobutyl acetate degradation, respectively. Remarkably, HSCT2005 with double esterase disruption exhibited the least ester degradation (17.6 ± 3.6%) (Figure 4-6C). Interestingly, we also observed that 13.3 ± 6.8% of isobutyl acetate decreased even without the cells (negative control), likely due to spontaneous hydrolysis and/or experimental loss. The results clearly showed that both Clo1313_0613 and Clo1313_0693 were the major esterases responsible for isobutyl acetate degradation in *C. thermocellum*. Furthermore, we observed that double esterase disruption in HCST2005 helped alleviate degradation of isobutyl isobutyrate but not ethyl esters (i.e., ethyl acetate and ethyl isobutyrate), as compared to the parent strain M1354 (Table 4-6).

Taken together, we identified and disrupted two endogenous CEs, Clo1313_0613 and Clo1313_0693, that could significantly alleviate hydrolysis of isobutyl acetate without potential inhibition on lignocellulose deconstruction in *C. thermocellum*.

4.4 Designer ester biosynthesis in engineered *C. thermocellum*

4.4.1 Conversion of alcohols to esters by CATec3 Y20F expressing *C. thermocellum*

In Chapter 2, we screened and engineered CATs for higher thermostability and catalytic efficiency towards a wide range of alcohols. Here, we investigated performance of the engineered CATs in *C. thermocellum* by comparing isobutanol conversion to the *in vitro* characteristics. By harnessing the high thermostability of CATec3 Y20F, we investigated whether CATec3 Y20F-harboring *C. thermocellum* could catabolize cellulose and convert various alcohols to designer bioesters efficiently at an elevated temperature of 55^{o}C (Figure 4-7A).

Table 4-6: Degradation of various esters by M1354 and HSCT2005 after 72 hour fermentation. The cells were cultured in the MTC medium with external supply of esters. Esters were added at initial concentrations of 1 g/L except for isobutyl isobutyrate at 0.5 g/L due its low solubility. The values represent means ± standard deviation of three biological replicates.

Esters	**M1354**	**HSCT2005**
Ethyl acetate	10.06 ± 2.04%	13.91 ± 3.00%
Ethyl isobutyrate	17.75 ± 2.28%	15.51 ± 9.13%
Isobutyl acetate	81.25 ± 0.70%	17.99 ± 0.88%
Isobutyl isobutyrate	67.35 ± 6.68%	9.34 ± 5.59%

We chose the engineered *C. thermocellum* Δ*clo1313_0613,* Δ*clo1313_0693* as the host because its two carbohydrate esterases was disrupted to alleviate ester degradation. By co-feeding cellulose and each higher alcohol, the recombinant *C. thermocellum* could produce all respective acetate esters (Figure 4-7B, D). Since *C. thermocellum* has the endogenous isobutyl-CoA pathway[160], we also observed the production of isobutyrate esters such as butyl isobutyrate and isobutyl isobutyrate as byproducts (Figure 4-7C). Many of these esters, such as n-butyl, n-pentyl, isoamyl, and geranyl esters, have never been reported to be feasibly synthesized in a thermophile since eukaryotic AATs are not thermostable. Among the esters, isoamyl acetate was produced at the highest conversion yield of > 30% (mol/mol) and titer of 1.2 g/L. Ester production in *C. thermocellum* was not as high as observed in *E. coli* likely due to the metabolic burden required to make cellulolytic enzymes for cellulose degradation along with overexpression of the heterologous gene. An increased titer of isobutyl esters was achieved when feeding a higher concentration of isobutanol (Figure 4-7E) below the lethal concentration (Figure 4-7F), indicating that the enzyme expression and/or alcohol availability were likely unsaturated in *C. thermocellum*. Further strain engineering and/or optimization of medium and operating conditions might help boost the production of designer bioesters by the CBP technology.

4.4.2 Thermostability of CAT is critical for efficient designer ester biosynthesis at elevated temperatures

Functional expression of a heterologous protein in thermophiles requires high thermostability, which is not well understood and often presents a significant bottleneck in metabolic engineering of these organisms.

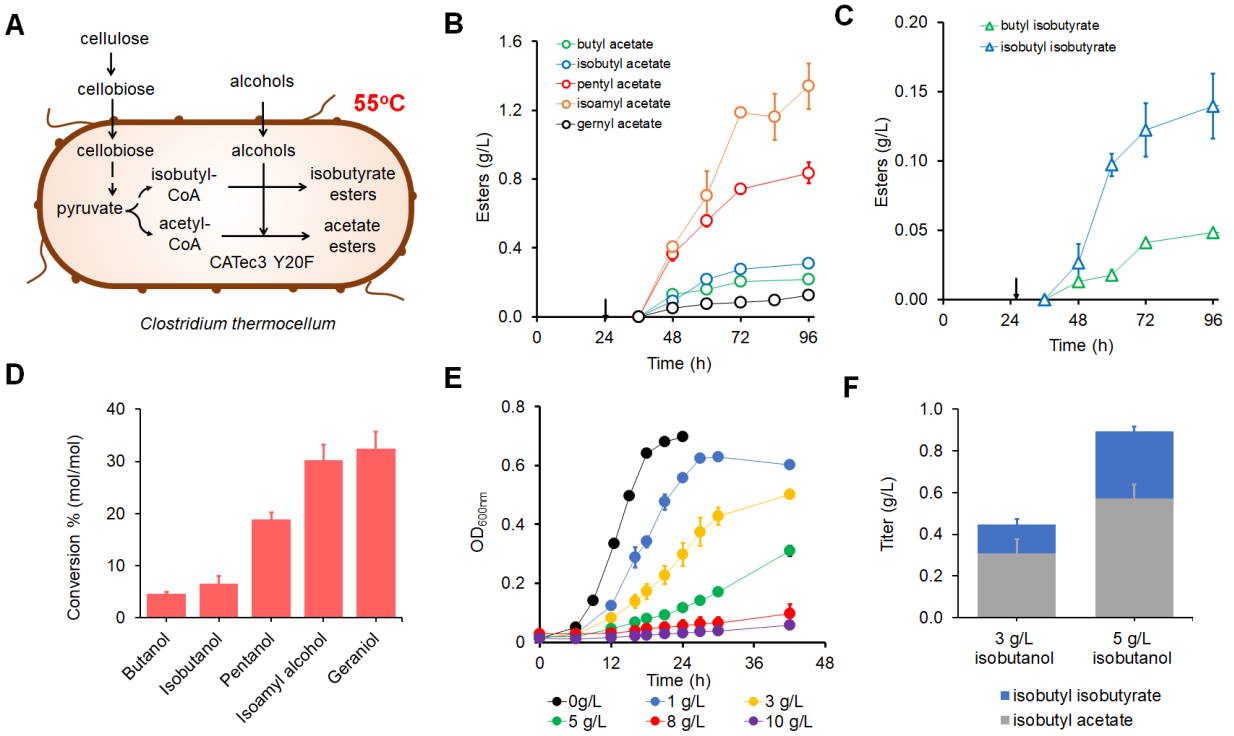


Figure 4-7: Conversion of alcohols by *C. thermocellum*. The error bars represent ±1 stdev from three biological replicates. **(A)** Microbial ester biosynthesis by a CATec3 Y20F-expressing *C. thermocellum* at 55°C. **(B)** Time dependent alcohol conversion to acetate esters by CATec3 Y20F-expressing strain *C. thermocellum* HSCT2108. **(C)** Time dependent alcohol conversion to isobutyrate esters by CATec3 Y20F-expressing strain *C. thermocellum* HSCT2108. **(D)** Alcohol conversion efficiencies. Each value represents mean ± 1 stdev from three biological replicates. **(E)** Toxicity of isobutanol to *C. thermocellum*. M1342 (*C. thermocellum* DSM1313 Δhpt) was cultured in MTC medium containing 5g/L cellobiose and isobutanol. Optical density (OD_{600nm}) was measured over 48 hours. **(F)** Isobutyl ester titers of CATec3 Y20F-expressing strain *C. thermocellum* HSCT2108 upon different isobutanol feeding concentration.

Inspired by the differences in the catalytic efficiency and melting temperatures among the engineered CATs (Table 2-9), we investigated how thermostability of the CATs affected ester production in *C. thermocellum*. We characterized the recombinant *C. thermocellum* Δ*clo1313_0613* Δ*clo1313_0693* (HSCT2005) harboring various CATs with distinctive melting temperatures and catalytic efficiency for *in vivo* isobutyl ester production by co-feeding cellulose and isobutanol at 55℃ (Figure 4-8A). Among the recombinant *C. thermocellum* strains, HSCT2108 harboring CATec3 Y20F, which has the highest catalytic efficiency and melting temperature, produced the highest level of isobutyl esters (892 mg/L), about 14-fold higher than the CATsa F97W-expressing *C. thermocellum* HSCT2105. Even though CATsa Y20F A138T has similar catalytic efficiency but higher melting temperature relative to CATsa Y20F, the CATsa Y20F A138T-expressing strain HSCT2113 produced 46% more esters than the CATsa Y20F-expressing strain HSCT2106. Similarly, we also observed higher ester production in the CATec3-expressing strain HSCT2107 than the CATsa F97W-expressing strain HSCT2105, where CATec3 has similar catalytic efficiency but higher melting temperature. Both HSCT2107 and HSCT2113 produced esters at very similar level, although CATsa Y20F A138T has higher catalytic efficiency but about 10℃ lower melting temperature than CATec3. These results strongly suggested that CAT robustness with enhanced thermostability plays a critical role for efficient ester production in *C. thermocellum* at elevated temperatures.

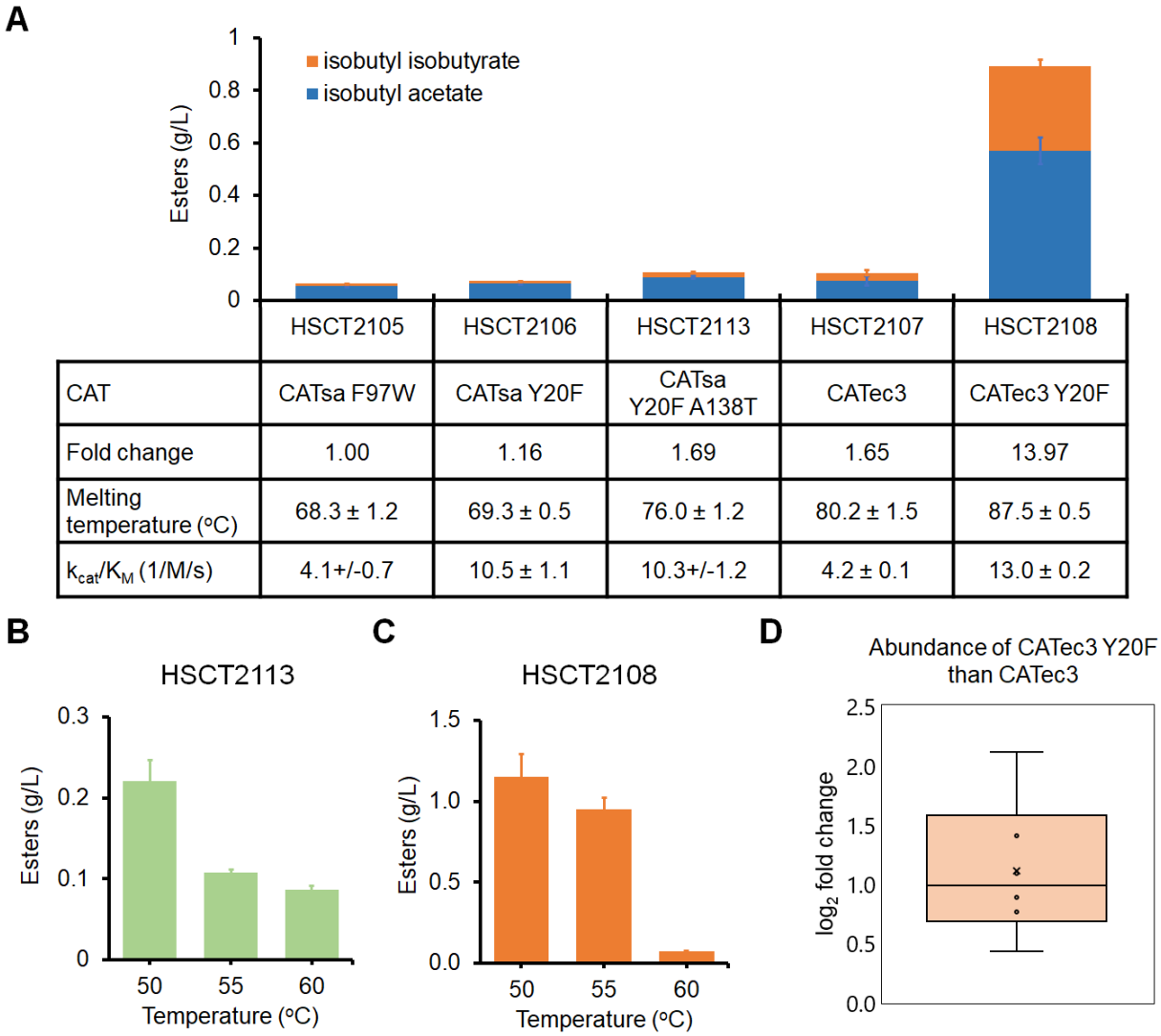
To further elucidate the effect of thermostability of CATs on ester production, we characterized the performance of HSCT2113 and HSCT2108 at various elevated temperatures (Figure 4-8B, C) compatible with *C. thermocellum* growth. Interestingly, HSCT2113 increased the ester production up to 220 mg/L at 50℃, about 2-fold higher than 55℃ (Figure 4-8B). In contrast, HSCT2108 produced esters at relatively similar level of about 1 g/L at 50 and 55℃, while the

production was reduced to 74 mg/L at 60℃ (Figure 4-8C). Proteolysis is a common cellular process to degrade and remove denatured or misfolded proteins[203]. If the cell growth temperature affected integrity of the CATs, their intracellular abundances would be altered. To quantify statistically meaningful changes in the intracellular abundance of CATs, we analyzed and compared proteomes across the two representative strains expressing CATec3, the wildtype (HSCT2107) versus the Y20F mutant (HSCT2108). Since the only difference between these two strains is one amino acid substitution, we could reliably quantify the relative abundance of each CAT by comparing tryptic peptide fragments. The result showed that CATec3 Y20F was 2.2-fold (average difference between peptide abundances) more abundant than CATec3 (Figure 4-8D). Because CATec3 Y20F has a 7℃ higher melting temperature than CATec3 (Table 2-8), the intracellular protein abundance change might imply a lesser degree of denaturation or misfolding due to higher protein thermostability. Taken altogether, CAT thermostability is critical for robust and efficient ester production in thermophiles by maintaining its intracellular protein abundance.

4.5 Metabolic engineering of *C. thermocellum* for C4-derived ester production

4.5.1 Consolidated bioprocessing of ester production from cellulose

The conversion of isobutanol to isobutyl esters at g/L scale by *C. thermocellum* overexpressing the engineered CATs suggested that improving isobutanol flux will guarantee higher ester production from cellulose fermentation (Figure 4-8). Therefore, metabolic engineering strategies to improve alcohol and acyl-CoA flux would enable the ester production from cellulose fermentation without supplementing the alcohols.



CAT	CATsa F97W	CATsa Y20F	CATsa Y20F A138T	CATec3	CATec3 Y20F
Fold change	1.00	1.16	1.69	1.65	13.97
Melting temperature (°C)	68.3 ± 1.2	69.3 ± 0.5	76.0 ± 1.2	80.2 ± 1.5	87.5 ± 0.5
k_{cat}/K_M (1/M/s)	4.1+/-0.7	10.5 ± 1.1	10.3+/-1.2	4.2 ± 0.1	13.0 ± 0.2

Figure 4-8: Effect of CAT thermostability on robust and efficient of microbial biosynthesis esters at elevated temperatures. **(A)** Isobutyl ester production of the recombinant *C. thermocellum* strains harboring various CATs. (B-C) Effect of temperatures on isobutyl ester production by the **(B)** CATec3-expressing strain HSCT2107 and **(C)** CATec3 Y20F-expressing strain HSCT2108. In panels a-c, each data represents mean ± 1 stdev from three biological replicates. An amount of 5 g/L of isobutanol was fed at the early stationary phase of the culture to avoid cell growth inhibition. **(D)** Box-whisker plot of protein level difference between HSCT2108 and HSCT2107. The data represent mean ± 1 stdev from three biological replicates (p-value: 0.0049).

A recent study engineered isobutanol production in *C. thermocellum* by overexpressing heterologous ketoisovalerate decarboxylase (KDC) and endogenous valine pathway[160]. However, the KDC from *Lactococcus lactis* ($KivD_{LL}$) has unsatisfactory thermostability that could potentially reduce the pathway efficiency over prolonged fermentation at elevated temperatures higher than 50°C[204]. Also, the process necessitated high cell density at OD above 16 which is an unfavorable configuration because the usual maximum OD of *C. thermocellum* under common culture conditions is lower than 1.5[173, 205].

Another possible strategy is to harness endogenous ferrodoxin-dependent ketoisovalerate oxidoreductase (KOR) pathway that catalyzes isobutyl-CoA synthesis from 2-ketoisovalerate[160, 206], a key precursor of valine biosynthesis. The isobutyl-CoA serves a key building block for C4-derived esters (e.g., ethyl isobutyrate, isobutyl acetate, isobutyl isobutyrate). The pathway is normally activated under high cellulose loading fermentation associated overflow metabolism[159, 207], resulting from promiscuous PFORs that primarily catalyze conversion of pyruvate to acetyl-CoA[206]. Among the PFORs, two PFORs (Clo1313_1353-1356 and Clo1313_0020-0023) were previously suggested to encode for the main PFOR complexes in *C. thermocellum*[208]. Especially, Clo1313_1353-1356 is responsible for the isobutanol production in *C. thermocellum* due to the detrimental effect of Clo1313_1353-1356 deletions on isobutanol production[206]. However, it has been largely unknown how to harness the endogenous pathway to overproduce C4-derived esters.

Here, we engineered the endogenous metabolism of *C. thermocellum* for ester biosynthesis from cellulose fermentation. We discovered that deletion of two carbohydrate esterases, Clo1313_0613, and Clo1313_0693, improves isobutanol by reducing endogenous thioesterase activity against isobutyl-CoA. Overexpression of a thermostable ester biosynthesis enzyme, CATec3 Y20F[59], synergistically enhanced isobutanol production with the esterase deletion,

improving isobutyl ester production 100-folds compared to a precedent strain. An engineered *C. thermocellum* strain produced 340 mg/L of isobutyl-CoA derived esters (i.e., ethyl isobutyrate, isobutyl acetate, and isobutyl isobutyrate) from one-step cellulose fermentation at 55℃.

4.5.2 Expression of CATsa F97W for endogenous ester production in *C. thermocellum*

As a starting point, we investigated whether *C. thermocellum* overexpressing CATsa F97W could produce esters at elevated temperatures. Previous studies have demonstrated that the wild type *C. thermocellum* has native metabolism capable of endogenously producing precursor metabolites for ester biosynthesis, such as acetyl-CoA, isobutyryl-CoA, as well as ethanol[209] and higher alcohols (e.g., isobutanol) under high cellulose loading fermentation[159, 160, 207]. Therefore, functional expression of the engineered CAT would produce esters using the endogenous pathways.

We first tested whether an engineered CATsa F97W can produce isobutyl acetate, a medium chain ester. We started by generating two isobutyl acetate-producing strains, HSCT0101 and HSCT0102, by introducing the plasmids pHS0024 (harboring the wildtype CATsa) and pHS0024_F97W (harboring the mutant CATsa F97W) into *C. thermocellum* DSM1313 (Table 4-1). Colonies were isolated on antibiotics selective plates at 55℃. Successful transformation clearly indicated that CATsa F97W conferred the thiamphenicol resistance and hence maintained CAT activity.

Next, we evaluated whether the *C. thermocellum* strains could synthesize isobutyl acetate directly from cellobiose. Since the endogenous isobutanol production from a typical cellobiose concentration (5 g/L) is low[159], we supplemented the media with 2 g/L isobutanol. Both HSCT0101 and HSCT0102 could produce isobutyl acetate at 55℃ as expected. HSCT0102

outperformed HSCT0101 with 3.5-fold increase in isobutyl acetate production (Figure 4-9A). Interestingly, we also observed the parent *C. thermocellum* M1354 produced a trace amount of isobutyl acetate (< 0.1 mg/L) even though this strain does not harbor a CAT (Figure 4-9B). This phenomenon was only observed when hexadecane overlay was used during fermentation for ester extraction. One possible explanation is the endogenous activity of esterases in *C. thermocellum* might have been responsible for low isobutyl acetate production while the organic phase overlay helps extract the target ester. It should be noted that the esterase reaction is reversible and more thermodynamically favorable for ester degradation than biosynthesis.

Finally, we tested whether HSCT0102 could endogenously produce isobutyl acetate directly from cellulose at elevated temperature (55℃). After 72 hours, cell mass, containing 550 mg/L of pellet protein, reached 1.04 g/L, and 17 g/L of cellulose were consumed (Figure 4-9C). About 103 mg/L of isobutanol were produced for the first 48 hours, and further increased up to 110 mg/L for additional 24 hours. Besides isobutanol, *C. thermocellum* also produced other fermentative metabolites, including ethanol, formate, acetate, and lactate, as expected (Figure 4-9A, B). For the target isobutyl acetate production, HSCT0102 did not produce isobutyl acetate for the first 24 hours but started accumulating the target product for the next 48 hours. The observed profile of isobutyl acetate production could be attributed to the low substrate affinity of CATsa F97W (Table 2-4). The final titer of isobutyl acetate reached 1.9 mg/L.

Besides the production of the desirable ester isobutyl acetate, we also observed that HSCT0102 produced other detectable esters such as ethyl acetate, ethyl isobutyrate, and isobutyl isobutyrate (Figure 4-10B, C). Endogenous biosynthesis of these esters could be explained from the complex redox and fermentative metabolism of *C. thermocellum*[207, 210]. *C. thermocellum* can endogenously synthesize the precursor metabolites, acetyl-CoA and ethanol via the ethanol

biosynthesis pathway while *C. thermocellum* can endogenously produce the precursor metabolites, isobutyryl-CoA and isobutanol via the valine biosynthesis (Figure 4-10A). With the availability of four precursor metabolites, *C. thermocellum* could produce ethyl acetate, ethyl isobutyrate, isobutyl acetate, and isobutyl isobutyrate as observed experimentally (Figure 4-10C).

Taken altogether, *C. thermocellum* overexpressing CATsa successfully produced the target isobutyl acetate from cellulose at elevated temperature (55^{o}C).

4.5.3 Employing the esterase-deficient *C. thermocellum* for ester production

To evaluate whether the esterase-deficient *C.thermocellum* can improve isobutyl acetate production, we introduced the thermostable ester production module CATsa F97W into HSCT2005 to generate the esterase-deficient, ester-producing strain HSCT2105. Strain characterization showed that HSCT2105 produced 3.1 mg/L of isobutyl acetate (Figure 4-11A) after the 72-hour cellulose fermentation. In comparison to HSCT0102, harboring the identical production module without the esterase disruption, HSCT2105 achieved 1.6-fold higher titer (Figure 4-11B). Since both strains yielded the same maximal pellet protein (~520 mg/L) and consumed cellulose (~17 g/L), it further supports that the esterase disruptions did not interfere with cellulose utilization. HSCT2105 produced isobutyl isobutyrate 10-fold higher than HSCT0102 (Figure 4-11B). This result correlates with decreased lactate and increased isobutanol production (Table 4-7), implying a higher metabolic flux was directed towards isobutyl-CoA, a precursor for isobutanol and isobutyl isobutyrate biosynthesis. Since some esterases exhibit versatile activities such as thioesterases[211-214], double deletion of esterases in HSCT0102 might have alleviated degradation of not only isobutyl isobutyrate but also isobutyl-CoA.

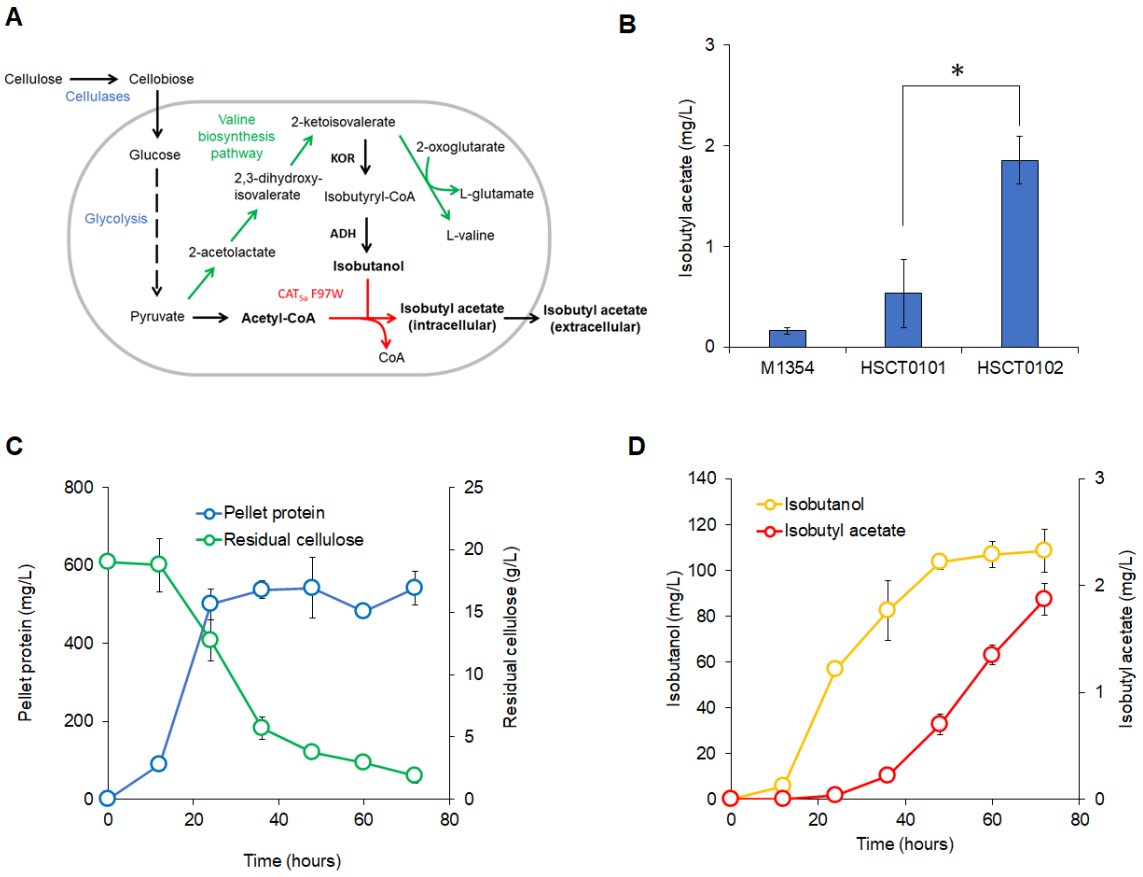


Figure 4-9: Isobutyl acetate production in the engineered *C. thermocellum*. **(A)** A simplified isobutyl acetate production pathway from cellulose in *C. thermocellum*. **(B)** Biosynthesis of isobutyl acetate of the wildtype and engineered *C. thermocellum* strains at 55℃ from MTC medium with 5 g/L cellobiose and external supply of 2 g/L isobutanol. Isobutyl acetate was measured after 24 hours from the hexadecane layer of cell cultures. Initial OD of each cell culture was in a range of 0.8−1.0. The error bars represent standard deviation of five biological replicates. Statistical analysis: t-test, "*" p value < 4 × 10−4, t = -6.475, df = 7. **(C)** Kinetic profiles of cell growth and residual cellulose of HSCT0102. HSCT0102 was cultured in C-MTC medium with 20 g/L Avicel PH-101. The error bars represent standard deviation of three biological replicates. **(D)** Kinetic profiles of isobutanol and isobutyl acetate production by HSCT0102 in C-MTC medium with 20 g/L Avicel PH-101. The error bars represent standard deviation of three biological replicates. Abbreviation: KOR, 2-ketoisovalerate ferredoxin oxidoreductase; ADH, alcohol dehydrogenase.

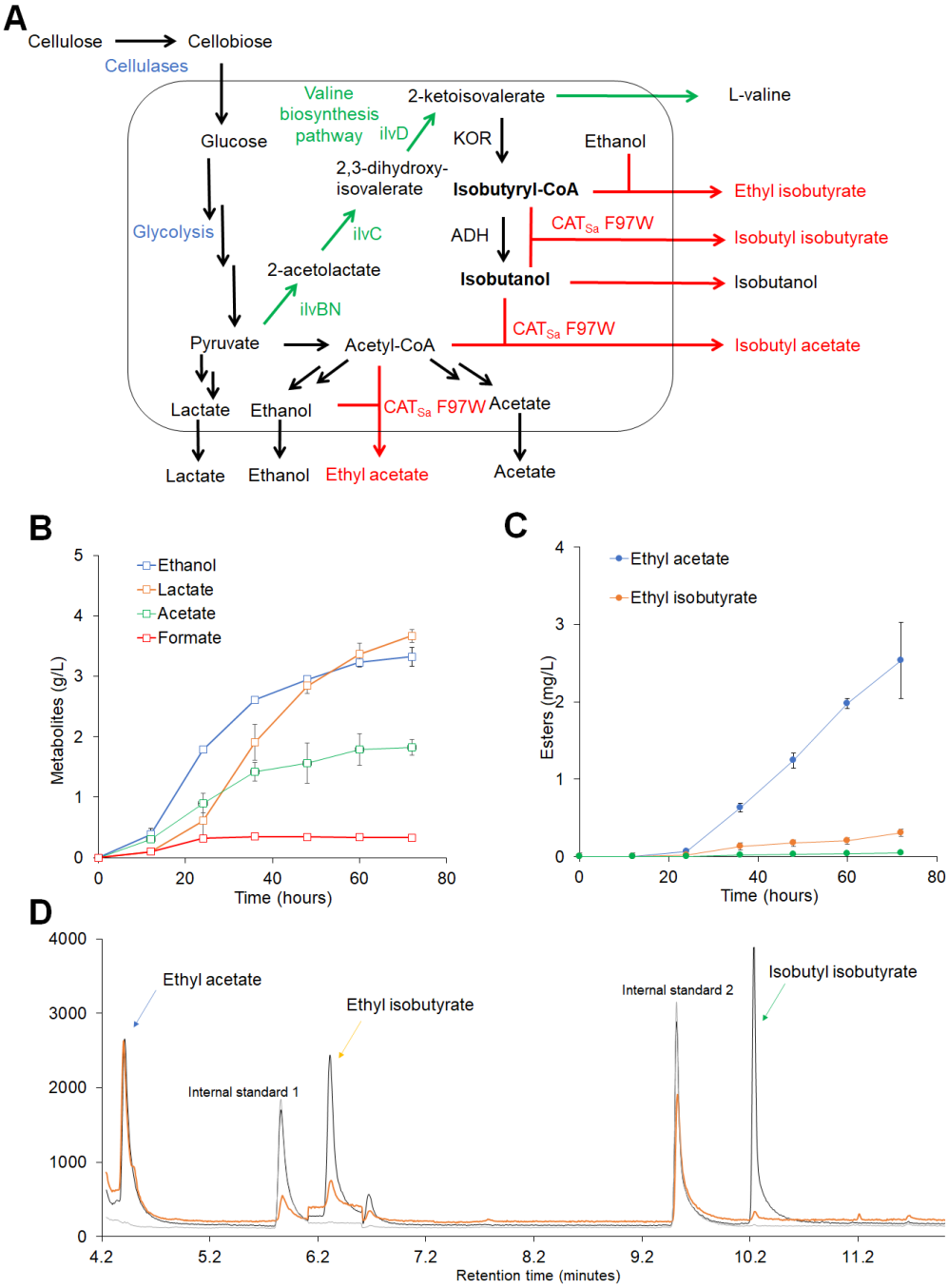


Figure 4-10: Fermentative products from cellulose fermentation of HSCT0102. **(A)** Simplified metabolic pathways of *C. thermocellum* for biosynthesis of native fermentative metabolites and heterologous esters. **(B)** Kinetic profile of biosynthesis of other esters besides isobutyl acetate, including ethyl acetate, ethyl isobutyrate, and isobutyl butyrate. **(C)** An overlaid GC/MS chromatogram demonstrating ester biosynthesis. **(D)** Kinetic profile of other fermentative metabolites. Abbreviation: IS, internal standard.

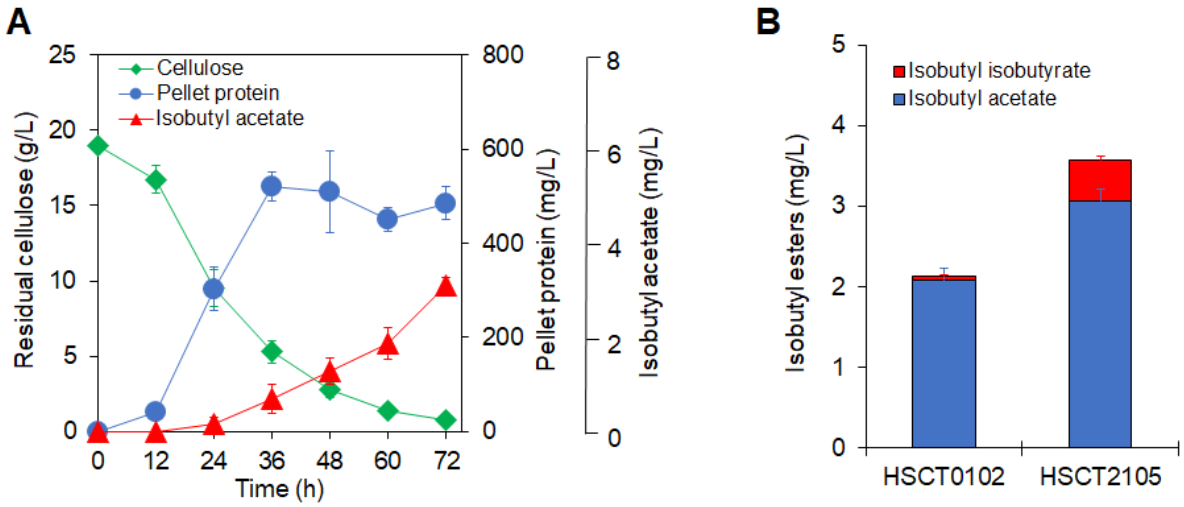


Figure 4-11: Comparison of ester production by HSCT0102 and HSCT2105. **(A)** A time dependent profile of isobutyl acetate production from cellulose by HSCT2105. The error bars indicate standard deviation of three biological replicates. **(B)** Ester concentrations by HSCT0102 and HSCT2105 after fermentation. The esters were quantified after 72 hours of cellulose fermentation. The error bars indicate standard deviation of three biological replicates.

Even though esterase deletion resulted in higher isobutyl ester production in HSCT2005 than HSCT0102, the ester titer was relatively low. Therefore, we hypothesize that other bottlenecks need to be tackled for more efficient ester production. One critical bottleneck is the availability of an efficient thermostable AAT used to construct the ester production module[50, 61, 65, 66, 190]. Based on the thermodynamics (Figure 4-1A), it is anticipated that the beneficial contribution from disrupting esterases to alleviate ester hydrolysis will become more significant for enhanced ester production as observed in Figure 4-8B when the esterase-deficient phenotype is tested in an ester overproducing *C. thermocellum* harboring a more efficient ester production module.

Overall, the esterase-deficient, ester-producing *C. thermocellum* HSCT2105 showed a higher metabolic capacity for isobutyl acetate biosynthesis. This strain can serve as a starting host for metabolic engineering to enhance ester production in future studies.

4.5.4 Overexpression of CATec3 Y20F and the esterase deletion synergistically improves metabolic flux towards isobutanol

Motivated by the fact that *C. thermocellum* DSM1313 overexpressing CATsa F97W could produce isobutyl acetate directly from cellulose, we further hypothesized that harnessing CATec3 Y20F would increase the isobutyl ester production titer because CATec3 Y20F outperformed at thermostability and catalytic efficiency towards isobutanol than CATsa F97W (Figure 4-8A). To compare the *in vivo* performance, we transformed *C. thermocellum* M1354 with a plasmid harboring CATec3 Y20F (HSCT0103). HSCT0103 produced 13.1 mg/L of isobutyl esters from 20 g/L cellulose fermentation (Figure 4-12A), 6.5 times high titers than the previously reported strain (HSCT0102) overexpressing CATsa F97W that produced 2.0 mg/L of isobutyl esters[58].

Table 4-7: Endpoint fermentative metabolites of cellulose fermentation by HSCT0102 and HSCT2105. The values represent mean ± 1 standard deviation from three biological replicates.

Strains	Acetate (g/L)	Lactate (g/L)	Formate (g/L)	Ethanol (g/L)	Isobutanol (g/L)
HSCT0102	1.82 ± 0.13	3.66 ± 0.10	0.33 ± 0.01	3.32 ± 0.15	0.11 ± 0.01
HSCT2105	1.36 ± 0.03	1.96 ± 0.09	0.46 ± 0.02	3.12 ± 0.07	0.30 ± 0.02

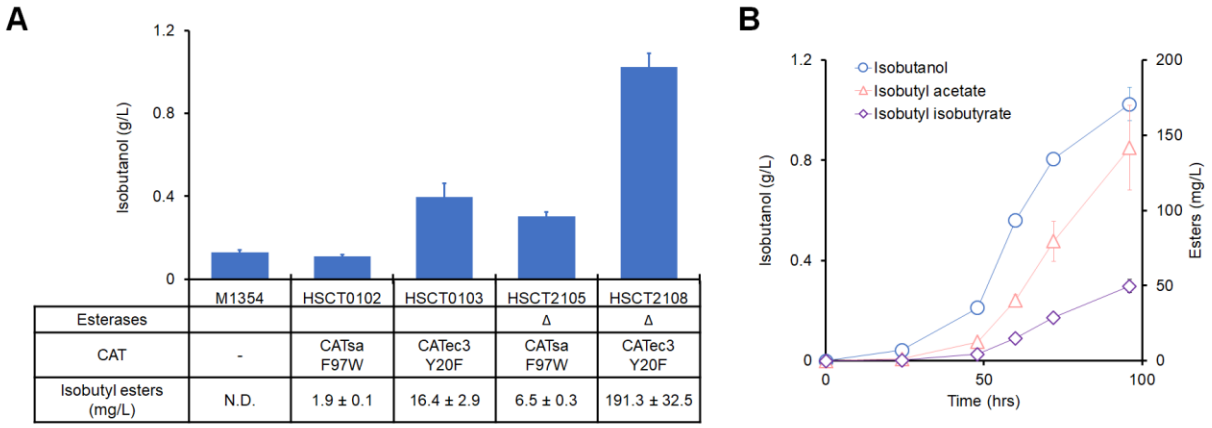
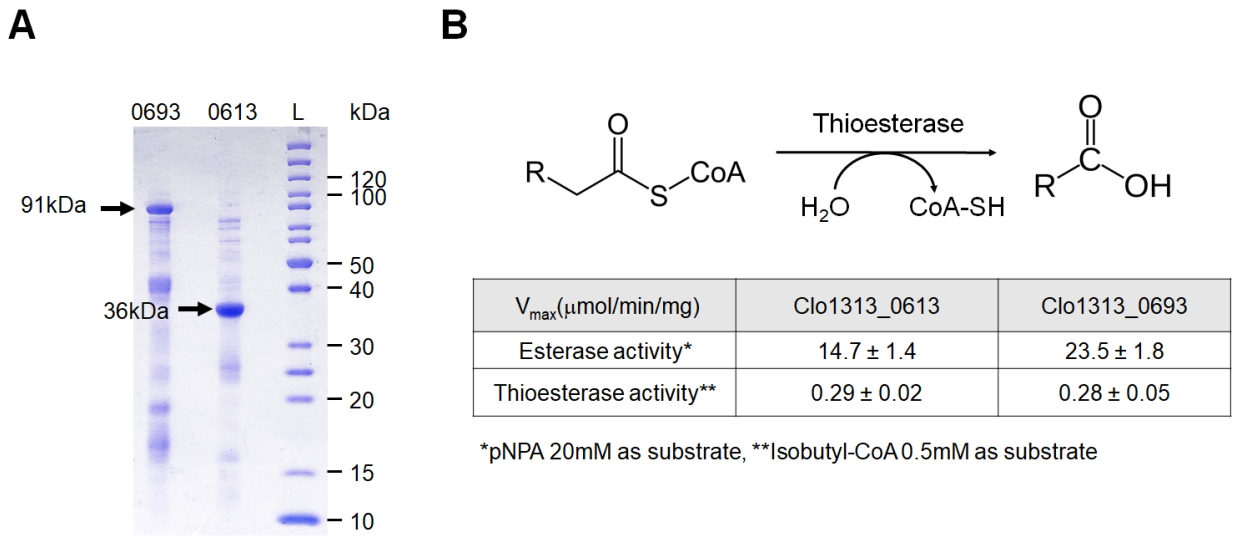


Figure 4-12: Effect of the esterase deletion and CAT expression on the CBP isobutyl ester production. **(A)** Isobutanol titers from cellulose fermentation by engineered *C. thermocellum* strains. **(C)** A profile of isobutanol, and isobutyl ester production by HSCT2108.

Surprisingly, overexpression of CATec3 Y20F also improved isobutanol titer compared with the parental strain (M1354) and HSCT0102 (Figure 4-12B). We also observed that disruption of two carbohydrate esterases and CATec3 Y20F overexpression (i.e., HSCT2108) synergistically improved isobutanol production achieving titers of 1.0 g/L (Figure 4-12B). HSCT2108 produced isobutyl esters (i.e., isobutyl acetate and isobutyl isobutyrate) up to 191.3 mg/L that was 16.6 times higher titer compared with HSCT0103.

Motivated by the increased isobutanol by *clo1313_0613* and *clo1313_0693* deletions, we further questioned whether the esterases have catalytic activity related to the isobutyl ester production pathway. We hypothesized that the esterases may possess thioesterase activities degrading isobutyl-CoA. To test the hypothesis, we compared esterase and thioesterase activities of Clo1313_0613 and Clo1313_0693 (Figure 4-13). The His-tag semi-purified enzymes were reacted with pNPA and isobutyl-CoA at the saturated concentration (Figure 4-13A). Interestingly, Clo1313_0613 and Clo1313_0693 exhibited a relatively weak thioesterase activity towards isobutyl-CoA, at maximum reaction rate (V_{max}) of 0.29 and 0.28 (μmol/min/mg), respectively. The promiscuous thioesterase activity of Clo1313_0613 and Clo1313_0693 was 50-fold and 84-fold lower compared to the primary esterase activity, respectively. Therefore, deletion of Clo1313_0613 and Clo1313_0693 might reduce intracellular isobutyl-CoA degradation and increase the isobutyl-CoA pool, leveraging the isobutanol and isobutyl ester production as observed (Figure 4-12). Clo1313_0613 might affect the intracellular isobutyl-CoA pool more than Clo1313_0693 because Clo1313_0693 is likely an extracellular enzyme (Figure 4-2A) and isobutyl-CoA is an intracellular metabolite.



V_{max}(μmol/min/mg)	Clo1313_0613	Clo1313_0693
Esterase activity*	14.7 ± 1.4	23.5 ± 1.8
Thioesterase activity**	0.29 ± 0.02	0.28 ± 0.05

*pNPA 20mM as substrate, **Isobutyl-CoA 0.5mM as substrate

Figure 4-13: Promiscuous thioesterase activity of Clo1313_0613 and Clo1313_0693. **(A)** SDS-PAGE of His-tag purified enzymes. 0693; Clo1313_0693, 0613; Clo1313_0613, L; protein ladder. The black arrow indicates the expected band sizes for each enzyme. **(B)** Maximum reaction rate (V_{max}) of Clo1313_0613 and Clo1313_0693 towards pNPA and isobutyl-CoA. The values represent mean ± 1 stdev from three biological replicates.

4.5.5 Deletion of lactate dehydrogenase and promoter modification

To further enhance the ester production, we next aimed to increase isobutanol flux and CATec3 Y20F expression level. First, we hypothesized that deletion of the lactate pathway would increase not only pyruvate pool, a key precursor of the valine synthesis pathway, but also NADH availability that can provide a driving force for isobutanol and isobutyl ester production. However, the *ldhA* (*clo1313_1160*) deletion did not significantly improve the isobutanol and isobutyl ester production but rather increased ethyl esters (Figure 4-14A).

We next hypothesized that higher CATec3 Y20F expression can further improve the isobutyl ester production. Previous studies reported that plasmid-based gene expression in *C. thermocellum* can cause undesired integration of plasmid into the genome by homologous recombination[215, 216]. The crossover event particularly occurs at the promoter region where is usually between 200 to 400 bp length[217]. We also confirmed that the plasmid-based CAT expression under the 310 bp length of gapDH promoter triggered such plasmid integration at genomic gapDH promoter region after 60 hours. The plasmid integration can result in loss of plasmid in the culture population, decreasing expression level of the target protein[216]. Because frequency of the homologous recombination depends on homology length[218], we truncated the gapDH promoter by eliminating 130 bp of the initial sequences to improve integrity of CATec3 Y20F expression in *C. thermocellum*. The truncated gapDH promoter included a putative SigA/RpoD motif to ensure integrity of transcription initiation[215]. Interestingly, CATec3 Y20F expression under the truncated promoter (HSCT3111) improved ester production by 48% compared to the original promoter. The total ester titers reached up to 370 mg/L from 17.2 g/L cellulose (Figure 4-14B). Especially, titers of ethyl acetate and ethyl isobutyrate were increased by 4-fold and 2-fold, respectively. The results suggested that ethanol pathway is a major

competitive pathway against the C4-derived ester production. The final ethanol concentration from the HSCT3111 culture was 4.1 g/L. Although CATec3 Y20F has poor affinity to ethanol (Table 2-9), the high concentration of ethanol can compete with isobutanol for the esterification.

4.5.6 Deletion of pyruvate-formate lyase

Because *C. thermocellum* employs two metabolic bifurcations, pyruvate-formate lyase (PFL) and the PFOR, at pyruvate node and the PFOR is closely related to the isobutyl-CoA and isobutanol production[206], we further tested whether deletion of PFL can activate the KOR pathway and reduce the ethanol formation. To eliminate the PFL activity, we added 6 mM of a PFL inhibitor sodium hypophosphite (HPP) to the culture medium[173, 219]. Unfortunately, HSEC3111 was not able to grow up in the HPP containing MTC-C medium, indicating that the *pfl* deletion is toxic to the cell. Lower concentration of HPP could restore the cell growth but the ester production was decreased by 50-60% compared to the culture without HPP. One possible explanation is the PFL deletion might reduce acetyl-CoA availability that is essential for detoxification of thiamphenicol selection marker and acetate ester production. Also, PFL is vital for redox balance when reduced ferredoxin is accumulated during fermentation[173].

Therefore, further engineering should focus on improving pathway specificity towards isobutanol over ethanol to reduce ethyl esters and improve isobutyl esters. Since deleting the ethanol producing bifunctional alcohol dehydrogenase (*adhE*) from *C. thermocellum* is detrimental to cell growth[220] and often problematic[221], one possible strategy is to engineer the AdhE for higher preference towards isobutyl-CoA than acetyl-CoA.

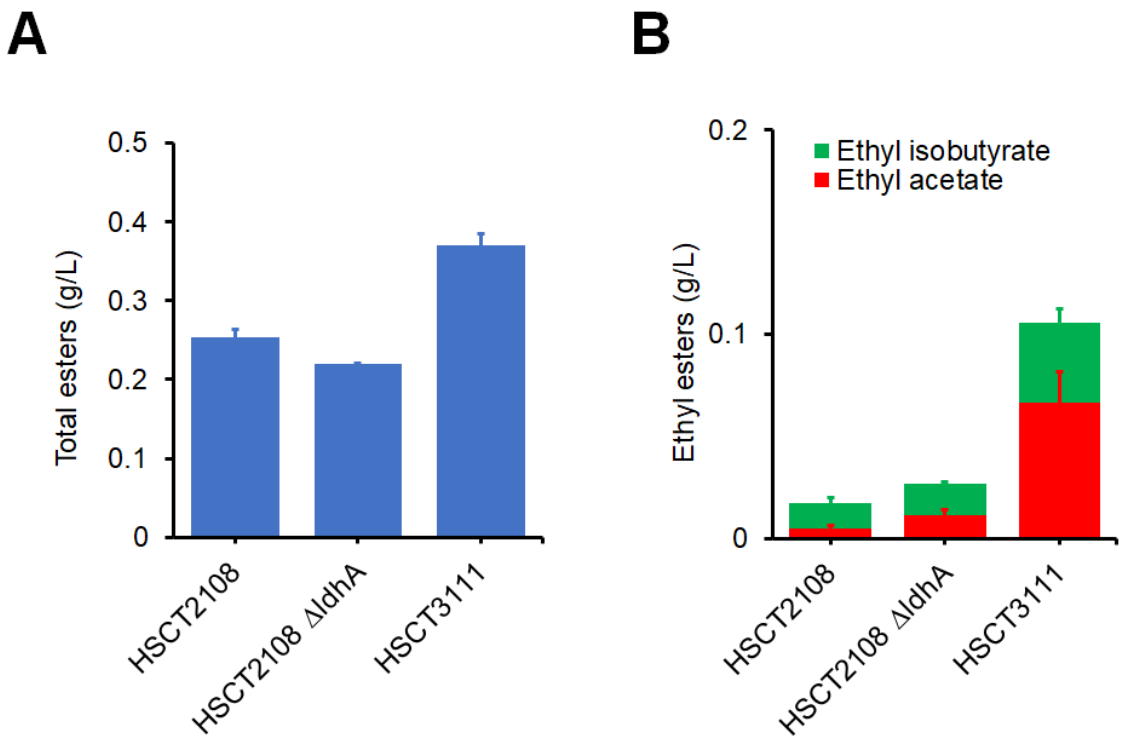


Figure 4-14: Effect of lactate dehydrogenase (*ldhA*) deletion and CAT expression under truncated version of GapDH promoter. **(A)** Final titers of total esters (i.e., ethyl acetate, ethyl isobutyrate, isobutyl acetate, and isobutyl isobutyrate). The values present mean ± 1 stdev from three biological replicates. **(B)** Final titers of ethyl esters (i.e., ethyl acetate and ethyl isobutyrate). The values present mean ± 1 stdev from three biological replicates.

4.6 Conclusion

In this chapter, *C. thermocellum* as a chassis was rendered more compatible to produce esters by disrupting endogenous esterases responsible for ester product degradation. The deletion of carbohyrdrate esterases to improve ester production can be broadly applicable to development of non-model microbial ester producer because the two identified carbohydrate esterases broadly exist in both non-cellulolytic bacteria and cellulolytic bacteria. The engineered esterase-deficient *C. thermocellum* was deployed to synthesize a wide range of esters at elevated temperatures with expression of a robustness, efficiency, and compatibility of engineered ester biosynthesis module.

The comparative proteomics analysis of CAT-expressing *C. thermocellum* strains unveiled that the CAT robustness with enhanced thermostability is critical for efficient ester production in thermophiles by maintaining high level of intracellular CAT abundance. The results highlight the significant role of enzyme thermostability on engineering heterologous pathways in a thermophilic whole-cell biocatalyst. While the enzyme has melting temperatures higher than the maximal growth temperatures of the host, the enzyme's functionality *in vivo* can be significantly hampered by the enzyme integrity, offering a design principle in engineering a heterologous pathway in thermophiles (e.g., *C. thermocellum)* for direct conversion of cellulosic biomass to chemicals and fuels.

In combination with gene expression and fermentation optimization, we demonstrated the first-generation of engineered *C. thermocellum* capable of producing 0.34 g/L of C4-derived esters by direct cellulose fermentation and 1 g/L of C4-derived esters with an isobutanol supplemented cellulose fermentation.

Chapter 5. Engineering community level modularity of ester biosynthesis

In plants, ester formulation has been considered as a defensive mechanism against pathogens[222, 223], and herbiovores[224, 225], or as an attractant for pollinators[226, 227]. Therefore, ester biosynthesis at the community level has been largely studied from an ecological or evolutionary point of view. However, from an engineering point of view, microbial community engineering can be an innovative approach to efficiently synthesize a wide range of esters. For example, it is challenging to design a strain capable of both efficiently utilizing the feedstocks and producing a target product. The balance between resource assimilation and product formulation can be more effectively accomplished by co-culture engineering by distributing the complicated tasks to multiple strains. The division of labor concept for microbial community engineering can be beneficial for specificity control and production of esters. This chapter focuses on engineering microbial syntrophic co-culture systems to produce esters by co-utilizing mixed sugars.

5.1 Motivation

Microbes naturally exist in populations and communities where the cells communicate each other by exchanging metabolites and molecular signals. The communication often exhibits cooperative behaviors that improve robustness of the system against environmental fluctuation[228, 229]. The cooperative behaviors also enable performing functions that are difficult or even impossible for individual strains[230, 231].

In metabolic engineering and synthetic biology, a microbial co-culture comprising of multiple strains has gained increasing interests to overcome many technical difficulties of

engineering monoculture, consisting of one microbial strain[232]. Despite of successful expansion of bioproducts, optimizing the biosynthetic pathways remains a major challenge for industrialization of the bioprocess[14]. The pathway optimization in mono-culture has pursued engineering biological parts such as promoter strength[233], ribosomal binding site (RBS)[234], and gene copy number[235] that are largely relied on availability of genetic tools and parts. Moreover, engineering of complicated biosynthetic pathways involving multiple heterologous gene expression often leads to undesirable outputs due to multiple reasons such as metabolic imbalance[236], metabolic burden[237], inefficient utilization of resources[238], which are not easily predictable. In contrast, co-culture engineering can separate multiple functions into individually optimized strains by compartmentalizing the molecular components of each pathway. By modularizing biosynthetic pathways, the co-culture strategy can reduce metabolic burden on individual strains and tune cellular environments optimal for specific reactions[239]. Also, pathway can be optimized simply adjusting population of the strain without reliance on availability of the genetic tools and parts. The division of labor concept has been experimentally demonstrated in many biochemical productions requiring a long linear heterologous pathways such as n-butanol[240], oxygenated taxanes[241], muconic acid[242], and resveratrol[243].

Co-culture engineering can be particularly advantageous when the biosynthetic pathway has highly interconnected precursor pathways undergoing diverging and converging steps. Microbial ester production pathway is an important example where precursor alcohol and acyl-CoA pathways are diverged from pyruvate or acetyl-CoA and then converged to esters by an alcohol acyltransferase (AAT). The alcohol and acyl-CoA pathways can be engineered for different chain length and isomeric structures which can be condensed together into combinatorial numbers of different ester classes. Critical to the ester biosynthesis is achieving a desirable balance

of the alcohol and acyl-CoA pathways for optimal production. However, engineering heterologous long chain alcohol and acyl-CoA pathways in a single strain often imposes metabolic burden impairing protein expression and cell viability[244]. As a result, few studies could engineer a single strain to efficiently produce esters derived from long chain alcohol and acyl-CoA modules[50, 65].

The innate modularity of ester biosynthesis pathway suggests that microbial co-culture engineering can be a promising approach for diverse ester production. Particularly, orthogonal pathway engineering of alcohol and acyl-CoA in individual strains can develop a flexible and programmable ester biosynthesis platform by modular co-culture design. For example, a co-culture with a specialized butanol producer and a valeryl-CoA producer can be designed to produce butyl valerate, a fatty wax ester with fragrance, flavor and fuel applications[62], from fermentation. The CATec3 Y20F is a robust AAT compatible with a wide range of alcohols and acyl-CoAs as well as microbial hosts[59] that can help the modular co-culture design for combinatorial numbers of ester production.

In this study, we demonstrated engineering of non-linear ester biosynthesis pathway by designing a syntrophic *E. coli*-*E. coli* co-culture. As a proof-of-concept, we first developed an *E. coli* modular chassis cell for glucose independent secondary sugar utilization. Using the platform, we constructed an *E. coli* strain preferring xylose over glucose by antibiotics-free genome engineering. The engineered strain avoided glucose induced carbon catabolite repression (CCR) allowing design of synthetic *E. coli*- *E.coli* co-cultures for ester production. An engineered *E. coli*-*E.coli* co-culture compartmenting isobutanol and butyryl-CoA pathways produced 392 mg/L of isobutyl butyrate from glucose-xylose co-fermentation, reporting the highest production titers up to date. This work presents a modular co-culture design platform for engineering various non-linear ester biosynthesis pathways.

5.2 Materials and methods

5.2.1 Bacterial strains and plasmids

The *E. coli* strains and plasmids used in this chapter were listed in Table 5-1.

5.2.2 Media and cultivation

E. coli strains were grown in lysogeny broth (LB) medium or M9 hybrid medium[80] containing glucose or xylose as carbon sources. For the ester production by co-culture, 25 g/L glucose and 10 g/L xylose, and 5 g/L yeast extract were used as major carbon and nitrogen sources. To increase buffer capacity, 10 g/L MOPS were added as an additional buffering reagent. Antibiotics of 100 μg/mL ampicillin and/or 100 μg/mL spectinomycin and/or 50 μg/mL kanamycin were supplemented to media when appropriate.

5.2.3 Genetic engineering of *E. coli*

E. coli genome engineering was carried out using the λ red mediated recombineering method [245]. The plasmid pSIM6 was used to induce homologous recombination of double strand DNA into the genome[122] and pCP20 was used to induce the FLP-mediated recombineering for subsequent gene deletions[125]. For the genome substitution of XylR mutant, cells were selected through multiple passages of culture transfers in the M9 medium containing 5 g/L glucose and 5 g/L xylose mixture and 1 g/L yeast extract. The primers used for the genetic engineering were listed in Table 5-2.

Table 5-1: A list of plasmids and *E. coli* strains used in this study.

Name	Descriptions	Source
Plasmids		
pET29a	pBR322 ori, Kan^R, lacI, T7lac promoter	Novagen
pET::GFP	pET29a::*iLOV* (GFP)	This study
pET::RFP	pET29a::*mCHERRY* (RFP)	This study
pETite*	pBR322 ori, Amp^R, lacI, T7lac promoter	Novagen
pSIM6	pSC101 $repA^{ts}$, Gam, Beta, Exo under the control of a temperature sensitive promoter	[122]
pCP20	$repA101^{ts}$, Cm^R, Amp^R, FLP recombinase under the control of a temperature senstivie protmoer.	[123]
pHM47	pET23a::*ilvC, ilvD, yqhD*	[124]
pHS121	pRSFDuet-1::*alsS*::*kivD*	This study
pDL3	pETite*P_{T7}::*atoB, hbd, crt, ter*::P_{T7}::*saat*	
pHS116	pRSFDuet-1::*catec3 Y20F*	This study
E. coli		
Top10	Host for molecular cloning, mcrA, Δ(mrr-hsdRMS-mcrBC), Phi80lacZ(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, rpsL(SmR), endA1, nupG	Invitrogen
BL21(DE3)	F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)	Novagen
HSEC0201	BL21(DE3) Δ*adhE,* Δ*dld*	Chapter 3
HSEC0415	BL21(DE3) Δ*adhE,* Δ*dld,* Δ*ptsI,* Δ*glk*	This study
HSEC0415xyl	HSEC0315 *xylR*::*xylR* R121C, P363S	This study
HSEC0403	BL21(DE3) Δ*adhE,* Δ*dld,* Δ*ldhA,* Δ*ackA-pta*	Chapter 3
HSEC0916	HSEC0403 harboring pHM47 and pHS121	This study
HSEC1017	HSEC0315 harboring pDL3 and pHS116	This study

5.2.4 Analytical methods

To quantify sugars and alcohols, HPLC system (Shimadzu Inc., MD, USA) was used. 600-750 μL of samples were centrifuged at 17,000 x g for 3 minutes followed by filtering through 0.2 micron filters. The HPLC samples were run with 5 mM H_2SO_4 at 0.6 mL/min on an Aminex HPX-87H (Biorad Inc., CA, USA) column at 50℃. Refractive index detector (RID) and ultra-violet detector (UVD) at 220 nm were used to determine metabolite concentrations.

Isobutyl acetate and isobutyl butyrate were measured by GC (HP 6890, Agilent, CA, USA) equipped with a MS (HP 5973, Agilent, CA, USA). For the GC system, the Zebron ZB-5 (Phenomenex, CA, USA) capillary column (30 m x 0.25 mm x 0.25 μm) was used to separate analytes, and helium was used as the carrier with a flow rate of 0.5 mL/min. The oven temperature program was set as follows: 50℃ initial temperature, 1℃/min ramp up to 58℃, 25℃/min ramp up to 235℃, 50℃/min ramp up to 300℃, and 2-minutes bake-out at 300℃. 1 μL of sampled hexadecane layer was injected into the column in the splitless mode with an injector temperature of 280℃. For the MS system, selected ion mode (SIM) was used to detect and quantify esters.

5.3 Modular design of *E.coli-E.coli* syntrophic co-culture for isobutyl butyrate production

5.3.1 Design of *E. coli* chassis for glucose-independent secondary sugar utilization

Capability of utilizing renewable organic wastes or agricultural materials as feedstocks is one of the striking benefits of bioprocessing in a mode of biochemical production. The natural materials are comprised of mixed nutrients which often need additional process to isolate each compound.

Table 5-2: A list of primers used in this study. The bold and underlined letters indicate restriction and site-directed mutation sites, respectively.

Primer name	Primer sequence (5' to 3')	Description
HS847	GGTTCTTTTAAAAATCAGTCACAAGTAAGGTA GGGTTATGGTGTAGGCTGGAGCTGCTTC	*ptsI* deletion forward
HS848	TGATCTTCTCCTAAGCAGTAAATTGGGCCGCA TCTCGTGGACATATGAATATCCTCCTTA	*ptsI* deletion reverse
HS849	AAGACGAGCAGAAAGCGG	*ptsI* check upstream
HS850	AGTTCCGGTATCCTTCTTG	*ptsI* check downstream
HS843	GAAAGAATTATTTTGACTTTAGCGGAGCAGTT GAAGAATGGTGTAGGCTGGAGCTGCTTC	*glk* deletion forward
HS844	GATTATCGGGGAGAGTTACCTCCCGATATAAA AGGAAGGATCATATGAATATCCTCCTTA	*glk* deletion reverse
HS845	GATATTTACAGGGAGCCTGCCTTTC	*glk* check upstream
HS846	ACGCAGGTCGGCCTTGTG	*glk* check downstream
HS740	TAACCTAGGCTGCTGCCACCGCTGAGCAAT	pRSFDuet-1 backbone for pHS116 forward
HS739	GTGGTGATGATGGTGATGGCTGCTGCCCATGG	pRSFDuet-1 backbone for pHS116 reverse
HS741	CCATGGGCAGCAGCCATCACCATCATCACCAC ATGAATTATACTAAATTCGATG	CATec3 Y20F for pHS116 forward
HS742	ATTGCTCAGCGGTGGCAGCAGCCTAGGTTA TTACTTCAATTTCGAATTGCAG	CATec3 Y20F for pHS116 reverse
HS902	CTCT**GAGCTC**AGGAAAAGAACCATGTTTACT AAACGTCAC	XylR cloning in a plasmid forward
HS903	AACCA**GCGGCCGC**CTACAACATGACCTCGCT ATTTACATC	XylR cloning in a plasmid reverse
HS904	CGTTAACTGTTTTGCTTTTTATGGTCTTCCG	XylR R121C mutagenesis forward
HS905	AAGCAAAACAGTTAACGCCTTTCTCTTTTAAA TGC	XylR R121C mutagenesis reverse
HS906	CGGTTATTCCTCGCTGCAATATTTCTACTCTG	XylR P363S mutagenesis forward
HS907	GCAGCGAGGAATAACCGCACATTTGCGATAT C	XylR P363S mutagenesis reverse
HS908	AGGAAAAGAACCATGTTTACTAAACGTCAC	XylR PCR for recombineering forward
HS909	CTACAACATGACCTCGCTATTTACATC	XylR PCR for recombineering reverse

A possible strategy to reduce the additional separation steps is to engineer microbial catalysts consuming the mixed nutrients better, allowing use of the mixed nutrients as feedstocks.

Lignocellulose is a representative renewable and abundant biomass comprising a mixture of fermentable hexose (e.g., glucose) and pentose (e.g., xylose, arabinose) sugars[246]. However, most microbes have hierarchical sugar preferences inhibiting efficient utilization of the mixed sugar. *E. coli*, for instance, consumes glucose first over the other sugars (e.g., xylose) due to the cAMP-CRP associated carbon catabolite repression (CCR) (Figure 5-1). When glucose is present, intracellular cAMP level decreases and the cAMP-CRP mediated transcription of catabolite genes is inactivated, inhibiting the secondary sugar catabolism. In contrast, when glucose is absent but other secondary sugars are present, cAMP-CRP mediated transcription is activated, expressing the secondary sugar catabolic systems. The non-glucose sugars are ranked in a hierarchy that the presence of higher sugar partially inhibits the expression of the lower sugar systems[247]. The intrinsic bias results in uneven sugar utilization that accumulates or wastes the less preferred sugars during multiple industrially relevant fermentation processes (e.g., fed-batch and continuous fermentation)[248]. Therefore, efficient fermentation of mixed sugars has been a major metabolic engineering challenge[249, 250].

Co-culture engineering is a promising approach for sugar co-utilization because each strain can be designed to orthogonally consume each sugar, avoiding competition between the strains. To establish a model system, various *E. coli* co-cultures have been demonstrated through extensive disruption of sugar assimilation machineries of each strain[242, 251]. However, such strategy has caveats such as remained CCR delaying the sugar consumption[242] and compromised cellular robustness caused by the extensive genetic manipulations[252, 253]. We hypothesized that engineering the CCR mechanism will be more practical to develop co-culture strain modules for various

biochemical production. Therefore, we aimed to design a platform *E. coli* chassis that can select for regulatory protein mutants relieving the CCR.

We chose HSEC0201 as a starting strain that lacks *adhE* and *dld* and also produces longer chain alcohol and/or acyl-CoA more effectively than the wild type BL21 strain (Figure 3-5B). We disrupted glucose assimilation pathway by deleting *ptsI* and *glk* genes, generating HSEC0415 (Δ*adhE* Δ*dld*, Δ*ptsI*, Δ*glk*) (Table 5-1). As expected, HSEC0415 was unable to grow in a glucose medium (Figure 5-2A). The slight growth is likely due to yeast extract in the medium because glucose was not consumed for the 50 hours (Figure 5-2B). When grown in a xylose medium, HSEC0415 showed 4% less growth rate (0.62 ± 0.02, 1/h) than the growth rate of HSEC0201 (0.64 ± 0.07, 1/h), suggesting that the effect of *ptsI* and *glk* deletion on xylose metabolism was minimal (Figure 5-2C, D).

5.3.2 Genomic XylR substitution by growth selection in glucose-xylose medium

The *ptsI* deletion disrupts the phosphotransferase system (PTS) that is involved in cAMP synthesis (Figure 5-3A). The decreased cAMP level significantly inhibits cAMP-CRP mediated activation of secondary sugar operons. Because the cAMP-CRP complex is an initial trigger for the catabolic gene expression, the *ptsI* deletion consequences inhibition of secondary sugar consumption delaying cell growth in a glucose-secondary sugar mixed medium. We hypothesized that mutant regulatory proteins increasing secondary sugar assimilation capability can be selected by HSEC0415 in a medium containing glucose and a corresponding secondary sugar. To test the hypothesis, we PCR amplified a XylR mutant (XylR R121C, P363S) that has higher binding affinity to operator sequences of *xyl* operon[254]. Then, we introduced the PCR fragment to HSEC0415 via recombineering followed by two subsequent overnight culture in a glucose-xylose

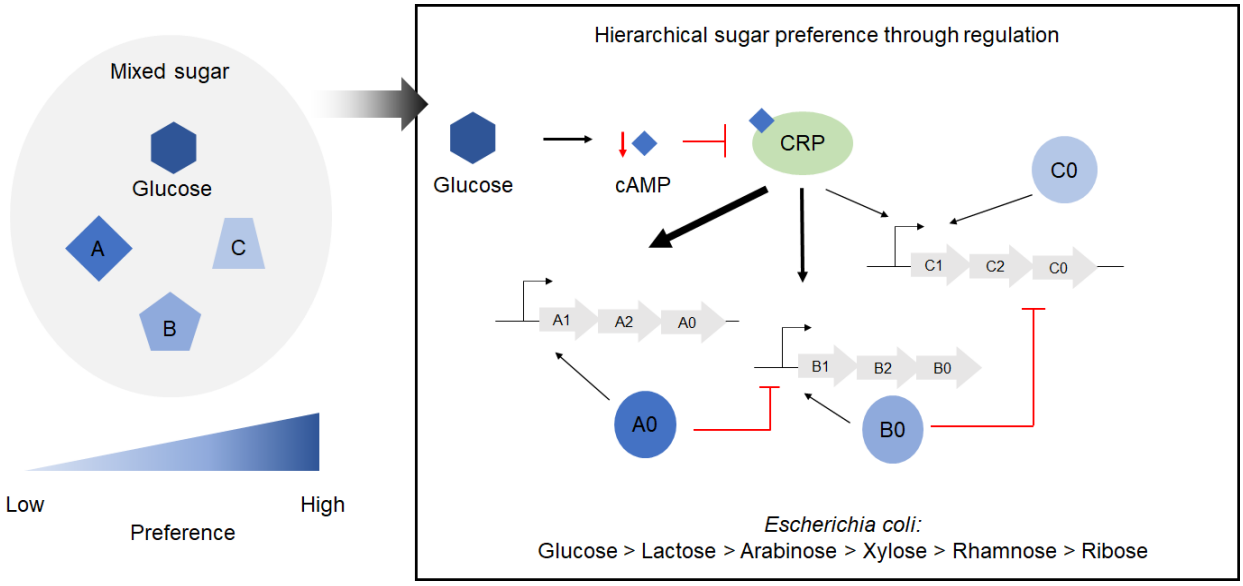


Figure 5-1: A scheme of hierarchical sugar preference dictated by carbon catabolite repression. The A0, B0, C0 represent regulatory proteins.

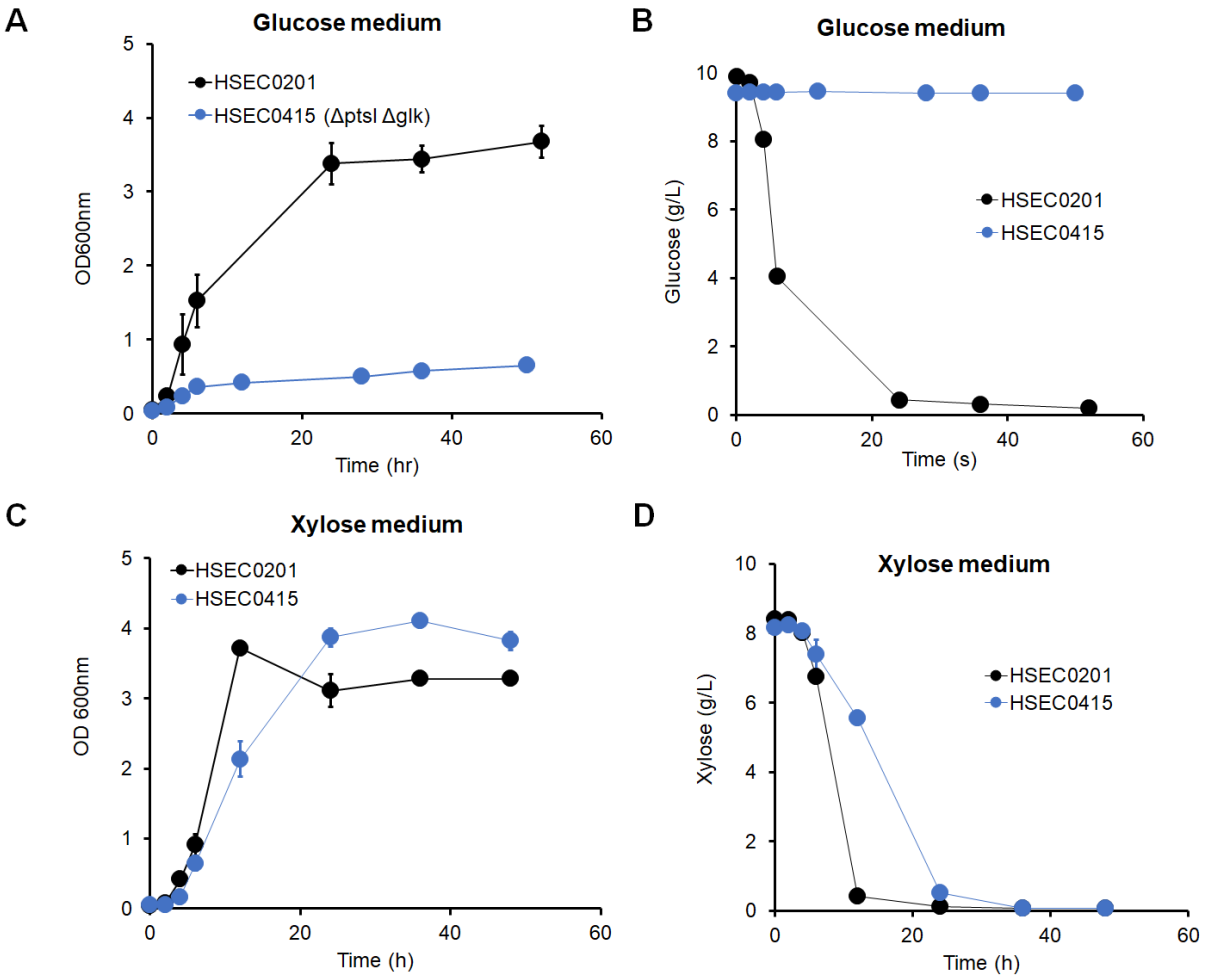


Figure 5-2: Disruption of glucose assimilation pathway. **(A)** Cell growth in glucose medium. **(B)** Residual glucose concentration in the medium during the culture. **(C)** Cell growth in xylose medium. **(D)** Residual xylose concentration in the medium during the culture. **(A-D)** The data represent mean ± 1stdev from three biological replicates.

mixture M9 medium without antibiotics (Figure 5-3B). Sequencing results of the XylR region from five isolated colonies revealed that two colonies obtained the XylR substitution in the genome, suggesting that HSEC0415 platform successfully selected for the XylR mutant. To verify the relieved CCR, we characterized cell growth and sugar consumption of HSEC0415 and HSEC0415xyl in a glucose-xylose M9 medium (Figure 5-3C, D). As expected, HSEC0415 preferred xylose over glucose but the xylose consumption was significantly inhibited by the presence of glucose due to the CCR (Figure 5-3C). In contrast, HSEC0415xyl bypassed the glucose induced CCR, consuming 5 g/L of xylose in 12 hours (Figure 5-3D). Although HSEC0415 and HSEC0415xyl did not grow in a glucose only M9 medium for 72 hours, both strains consumed glucose after depletion of xylose, suggesting that the *ptsI* and *glk* deletions did not eradicate the glucose assimilation ability.

Overall, a glucose-independent, xylose-preferring *E. coli* was successfully constructed by the genomic XylR mutation. The results suggest that HSEC0415 chassis can serve as a selection platform for regulatory protein mutants relieving the glucose induced CCR.

5.3.3 Compartmentation of isobutyl butyrate pathway

The independent xylose consumption of HSEC0415xyl can minimize sugar competition in a consortium, facilitating orthogonal sugar utilization. We next hypothesized that a synthetic *E. coli-E. coli* syntrophic co-culture can compartment a non-linear ester biosynthesis pathway and improve the production. To test the strategy, we designed a co-culture for isobutyl butyrate production from glucose and xylose (Figure 5-4A). The co-culture compartments the isobutyl butyrate pathway into two strains that prefer different sugars.

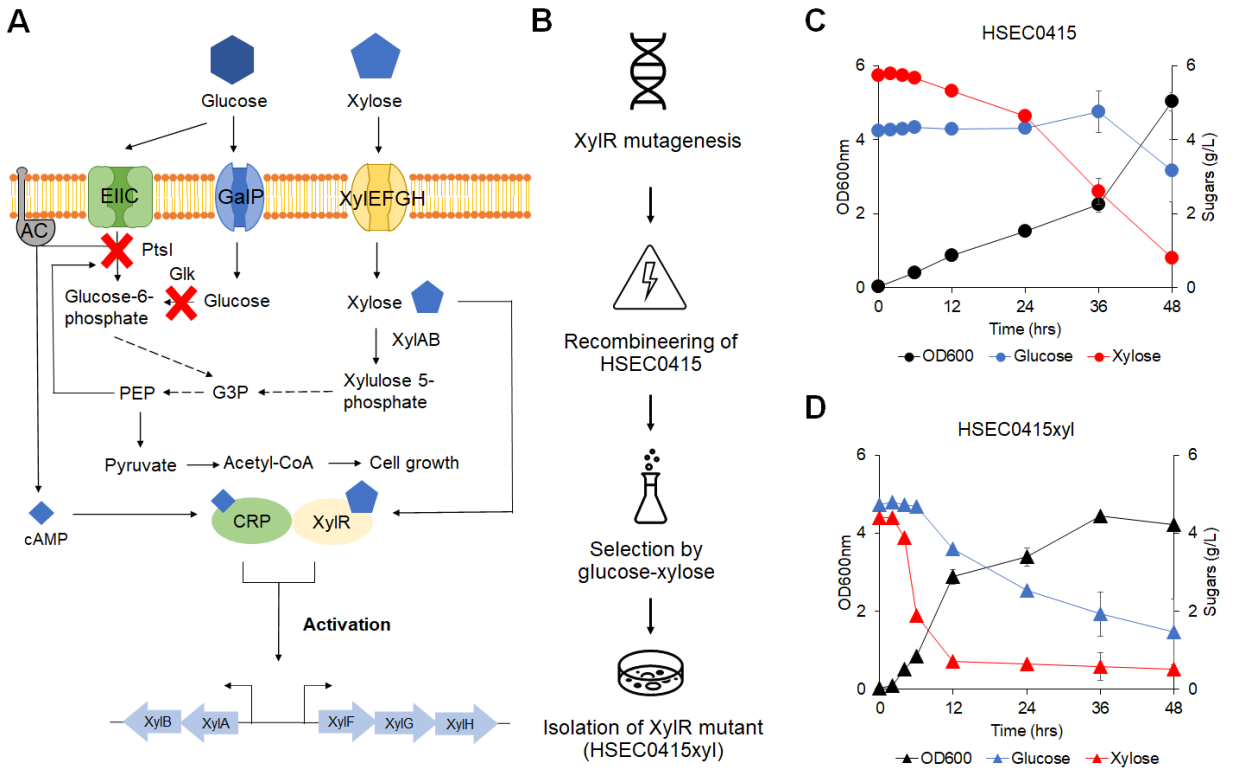


Figure 5-3: Selection of a genomic XylR substitution avoiding glucose induced CCR using HSEC0415. **(A)** A metabolic pathway map of glucose and xylose assimilation. The red cross marks indicate deletion of *ptsI* and *glk*. When glucose is present, xylose consumption is inhibited by the CCR. AC, adenylate cyclase; G3P, Glyceraldehyde-3-phoshpate; PEP, phosphoenolpyruvate. **(B)** A workflow of genomic XylR mutagenesis in HSEC0415. **(C)** A profile of HSEC0415 cell growth, glucose, and xylose concentration in a glucose-xylose medium. **(D)** A profile of HSEC0415xyl cell growth, glucose, and xylose concentration in a glucose-xylose medium. **(C-D)** The data represent mean ± 1stdev from three biological replicates.

A strain is specialized to produce isobutanol (HSEC0916), while the other strain is specialized to produce butyryl-CoA (HSEC1017) (Figure 5-4B). Because isobutanol is a more reduced product requiring more carbon source than butyryl-CoA, we designed the glucose consuming strain to produce isobutanol. To synthesize isobutyl butyrate, the butyryl-CoA producing strain should express the AAT because butyryl-CoA cannot transfer to the other strain but isobutanol can. For a high AAT activity, a SAAT from strawberry and an engineered CATec3 Y20F were co-expressed because they are capable of synthesizing butyrate esters[59, 65, 165].

5.3.4 Isobutyl butyrate production by *E. coli-E.coli* co-culture co-fermenting glucose and xylose

The co-culture can produce isobutyl acetate as a byproduct because the AATs can condense intracellular acetyl-CoA with isobutanol. Therefore, we tested whether the product specificity can be engineered by controlling the population. We altered inoculation ratio and measured isobutyl butyrate and isobutyl acetate production upon the different population (Figure 5-4C). Interestingly, the higher HSEC0916 inoculation produced the more isobutyl acetate, while isobutyl butyrate titers were not significantly changed. This suggests that the isobutanol pathway is not likely the bottleneck. Isobutanol was accumulated up to 0.5 g/L from the 1:1 inoculum ratio, further supporting that butyryl-CoA or AAT activity were likely the limiting step. The highest isobutyl butyrate titer of 133 mg/L was achieved at 1:4 inoculum ratio, 11-fold higher titer than isobutyl butyrate production from an engineered monoculture[65]. The dynamic isobutyl ester profile with different inoculum ratio demonstrates that isobutyl ester production can be engineered by controlling the population.

Because the co-culture at 1:4 ratio produced ~133 mg/L of isobutyl butyrate as the main ester product, we investigated the system to further improve isobutyl butyrate production. Although the M9 medium additionally contained 10 g/L of MOPS, the pH rapidly decreased below 5 after 48 hours, and the production did not proceed further. Therefore, we intermittently adjusted the pH with 5M KOH every 12 hours, maintaining the pH between 6.0 to 7.0 (Figure 5-4D, E). The pH adjustment avoided the fermentation cessation, producing 392 mg/L of isobutyl butyrate for 96 hours (Figure 5-4D). The co-culture maintained steady production and sugar consumption rates over the 96 hours (Figure 5-4E), suggesting that the population was well maintained over the fermentation. During the culture, 9.2 g/L glucose and 6.4 g/L xylose were consumed at consumption ratio of 0.69 (g xylose/g glucose).

Overall, design of a synthetic *E. coli-E.coli* co-culture produced isobutyl butyrate by co-fermenting glucose and xylose with robust population over 96 hours. Because the individual strains can be further engineered to produce other esters with non-linear biosynthesis pathways, the co-culture strategy can be widely applicable to various ester production.

5.4 Conclusion

In this study, we engineered a syntrophic *E.coli-E.coli* co-culture for isobutyl butyrate production. The co-culture simultaneously utilized glucose and xylose and produced isobutyl butyrate up to 392 mg/L which was 32-fold higher titer than a previously reported mono-culture[65].

For a robust *E.coli-E.coli* co-culture, we developed an *E. coli* chassis cell capable of selecting regulatory protein mutants allowing orthogonal sugar utilization in a consortium. Using the platform, we could edit XylR from the *E. coli* genome without an antibiotic marker.

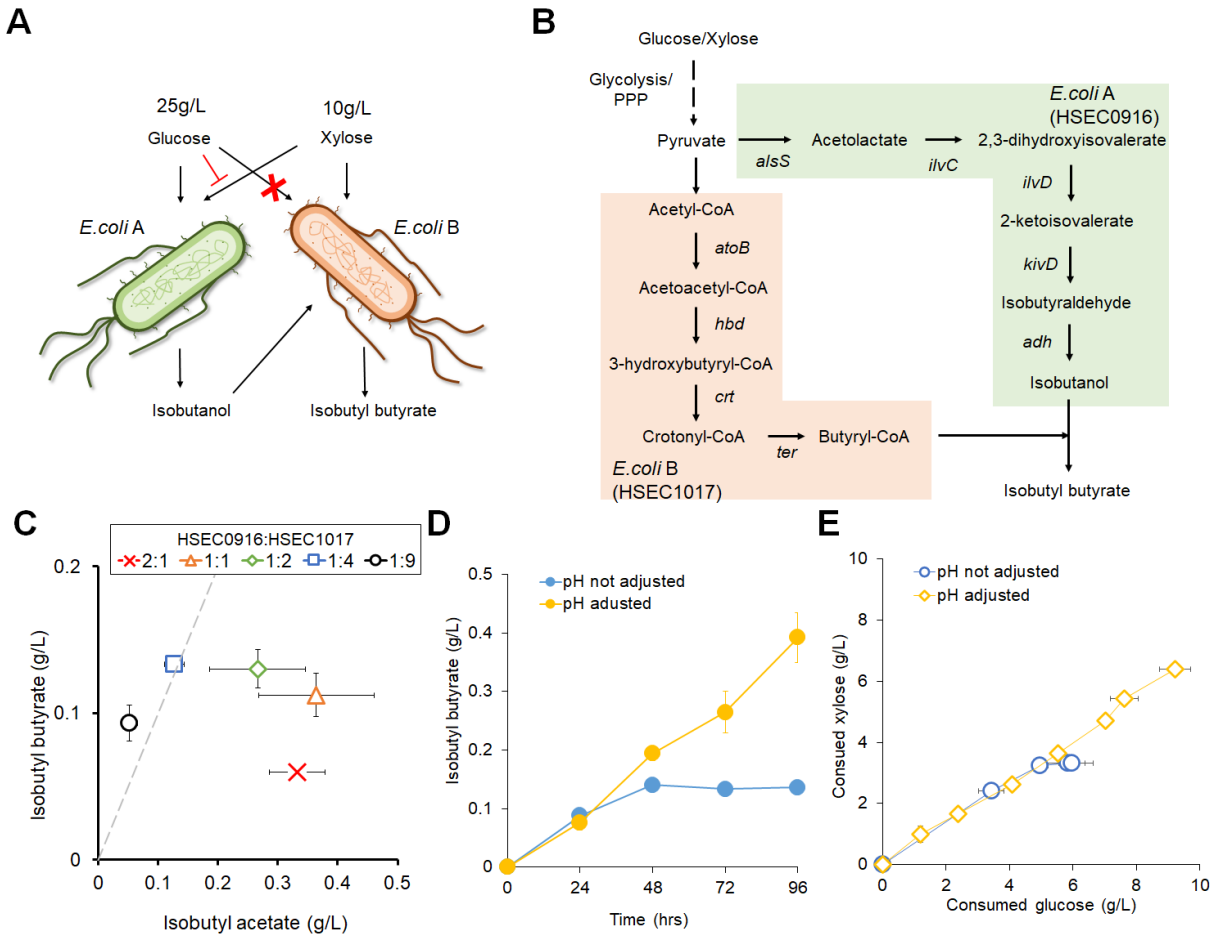


Figure 5-4: Designing a synthetic *E.coli-E.coli* co-culture for isobutyl butyrate production. **(A)** A scheme of co-culture design for isobutyl butyrate production by co-utilization of glucose and xylose **(B)** Compartmentation of non-linear isobutyl butyrate biosynthesis pathway to two strains. **(C)** Titers of isobutyl acetate and isobutyl butyrate at different inoculum ratio after 48 hours. The grey dotted line represents 50% (w/w) product specificity. **(D)** Titers of isobutyl butyrate upon pH adjustment. The cell inoculum ratio was at 1:4 (strain A:B) **(E)** A profile of consumed glucose and xylose over the culture. **(C-E)** The data represent mean ± 1stdev from three biological replicates.

The engineered strain efficiently consumed xylose in the presence of glucose, formulating a robust population steadily co-utilizing glucose and xylose over 96 hours. Because the glucose induced CCR is a complicated mechanism, the *E. coli* chassis platform will be useful to engineer sugar preference of individual *E. coli* strains employable to modular microbial co-culture designs.

The co-culture could produce isobutyl butyrate as a major ester product through the relatively simple inoculation optimization. The results demonstrate viability of the coculture strategy for optimizing pathway balance of non-linear ester pathway. Similarly, engineering of non-linear rosmarinic acid pathway using microbial co-culture was recently demonstrated[255], suggesting viability of the co-culture strategy for pathway balance. Because many esters have complicated non-linear precursor pathways, the co-culture strategy will be generally applicable for production of esters difficult in monoculture systems (Figure 5-5).

The *E.coli-E.coli* co-culture can exemplify multiple synthetic and/or semi-synthetic co-culture designs harnessing diverse microbes specialized to produce long chain alcohols and acyl-CoAs. For example, some non-model Clostridia species overproduce long chain alcohols and acyl-CoA from various resources including sugars, organic acids, and ethanol[256-259] and therefore, hijacking their metabolism can potentially expand biological routes for ester production. Because of limited genetic tools and parts for the non-model microbes, engineering metabolism of individual strains for optimal pathway balance is challenging. In contrast, co-culture engineering can simplify the pathway balance and optimization by manipulating population via inoculation engineering and environmental control. Taken altogether, modular co-culture design can tackle many technical challenges of ester biosynthesis.

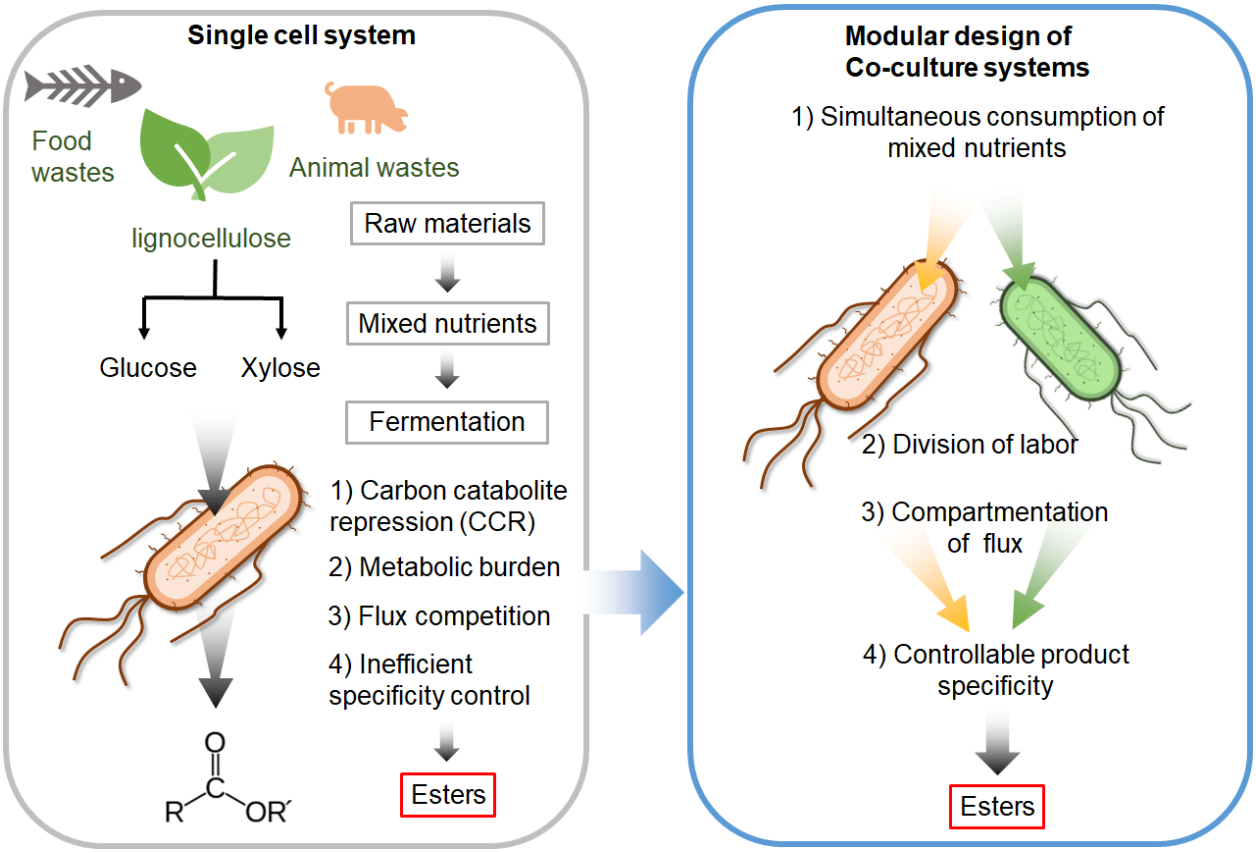


Figure 5-5: Modular design of microbial co-cultures for ester production.

Chapter 6. Conclusion and future directions

Engineering biological modules with high modularity offers distinct advantages for scalable design, construction, operation, and maintenance of complex biological systems. As a testbed, this dissertation engineered modularity of ester biosynthesis across biological scale for novel biocatalysis.

At the molecular level, an AAT with high robustness, compatibility, and efficiency was developed by engineering promiscuity of robust CAT enzymes. The AAT module exhibited relatively good catalytic efficiency towards a broad range of alcohols and acyl-CoAs, enabling combinatorial ester biosynthesis. The approach can lead to a breakthrough in microbial ester production because the conventional eukaryotic AATs lack sufficient robustness, efficiency, and compatibility for an interchangeable catalytic module as they commonly exhibit poor enzyme expression[66], solubility[67], and thermostability[68] in microbial hosts. Also, high stability and promiscuity of CATs offer a better enzyme evolvability[76-78] than eukaryotic AATs, enabling it to tolerate more functionally beneficial but destabilizing mutations[76]. Such characteristics provoke many interesting scientific questions including the possibility of refining CAT's functions for higher specificity against small alcohols (e.g., methanol, ethanol) and/or long chain acyl-CoAs. Due to relative short amino acid length and well identified protein structures of CATs, additional mutagenesis via rational design and/or directed evolution can further tune CATec3 Y20F for specific ester production.

At the cellular level, mesophilic and thermophilic whole-cell biocatalysts were developed for ester production. In a whole-cell biocatalyst, enzyme performance does not solely rely on catalytic efficiency because of complicated reaction environments (e.g., intracellular enzyme abundance, precursor availability, intracellular pH, and temperature). Particularly, the precursor

alcohol and acyl-CoA should be sufficiently supplied with amplified metabolic flux for optimal ester production. By using metabolic engineering and synthetic biology strategies, cellular metabolism of *E. coli* and *C. thermocellum* was successfully rewired to overproduce medium chain esters (C_6-C_8) from glucose and cellulose fermentation, respectively. The whole-cell biocatalysts further demonstrated robustness, compatibility, and efficiency of the engineered ester biosynthesis module by efficiently producing various esters at the cellular level.

Although biological roles of the volatile esters in microbes are highly debated and elusive [183], ester biosynthesis can improve fitness of the host cells by detoxifying toxic metabolites. It has recently been shown that for the same total carbon chain length, short-chain esters are less toxic to microbial health than alcohols[127]. Especially, esters are more hydrophobic than their precursor alcohols, enabling easy extraction during fermentation process. Moreover, ester biosynthesis can be beneficial for cellular robustness by controlling cellular metabolism. As shown in Figure 1-2, alcohol and acyl-CoA biosynthesis pathways exhibit modularity that can be managed by an ester biosynthesis module having a certain substrate specificity. For example, the reversal beta-oxidation cycle requires a metabolite sink to prevent uncontrolled carbon elongation iterations[52]. Esterification reaction of the alcohol and acyl-CoA can serve as a metabolic controller by generating esters as the metabolite sinks.

Such evolutionary incentives can be harnessed to develop novel ester producing strains using modular cell (MODCELL) engineering. Modular cell harnesses the compatibility between the production module and chassis cell for pathway optimization. Computational metabolic pathway modeling helps design of auxotrophic chassis cell requires the production module for healthy cell growth[27]. Understanding evolutionary incentive of ester biosynthesis at the cellular level will provide many insights on implementing the modular cell engineering for ester production.

At the microbial community level, a synthetic microbial co-culture was engineered for isobutyl butyrate production, which involves a complicated non-linear biosynthetic pathway. For robust co-culture designs, an *E. coli* chassis platform was developed by disrupting glucose assimilation pathways. The platform can select regulatory protein mutants that avoid glucose induced CCR and improve assimilation of secondary sugars. The platform will be helpful to construct *E. coli* modules that orthogonally utilize multiple sugars in a consortium.

Although the *E. coli-E.coli* co-culture produced about 30-fold higher titers of isobutyl butyrate than an engineered monoculture, it has been not well understood why and how the division of labor can relieve metabolic burden. For instance, a quantitative definition of metabolic burden has not been established and therefore, analyzing the performance of co-cultures is somewhat superficial. Therefore, it is an important future direction to quantify the metabolic burden in heterologous pathway engineering. Systems biology techniques such as transcriptomics, proteomics, and metabolomics can offer promising ways for the quantitative analysis of metabolic burden. Especially, identifying biomarkers responding the metabolic burden will be a critical step.

While the study focused on finding a laboratory set up to understand microbial community system better, the results provoke many fundamental questions. One of the important questions is evolutionary incentives of ester production by microbes in a community. Because many esters have pleasant odors, evolution of AATs can be promoted through selection by higher organisms' sensation. A study successfully demonstrated that volatile ester formulation by a yeast promoted cell dispersion mediated by insect vectors[260], suggesting a hypothesis that the microbial ester production might attract animal vectors and promote the cell dispersion. Selection of ester forming beverage fermenting yeasts is a representative example where humans' preferences on beer and wine can deliver a significant evolutionary driving force[261]. It will be an interesting engineering

problem to develop an AAT selection platform at a laboratory scale through the interkingdom interactions.

In conclusion, engineering modularity of ester biosynthesis facilitated scalable engineering of biological systems. Because many biochemical synthesis pathways naturally exhibit a certain level of modularity, engineering strategies presented in this dissertation will be useful for developing other biochemical production platforms.

Bibliography

1. Dowd, J.B. et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci U S A* **117**, 9696-9698 (2020).
2. Dillon, M.E., Wang, G. & Huey, R.B. Global metabolic impacts of recent climate warming. *Nature* **467**, 704-706 (2010).
3. Landrigan, P.J. et al. The Lancet Commission on pollution and health. *Lancet* **391**, 462-512 (2018).
4. Hmiel, B. et al. Preindustrial (14)CH4 indicates greater anthropogenic fossil CH4 emissions. *Nature* **578**, 409-412 (2020).
5. Vohra, K. et al. Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem. *Environ Res* **195**, 110754 (2021).
6. Frazzetto, G. White biotechnology. *EMBO Rep* **4**, 835-837 (2003).
7. Dick, G.J. et al. The microbiology of deep-sea hydrothermal vent plumes: ecological and biogeographic linkages to seafloor and water column habitats. *Front Microbiol* **4**, 124 (2013).
8. Macfarlane, G. Alexander Fleming, the man and the myth. (Oxford University Press, Oxford Oxfordshire ; New York; 1985).
9. Pham, J.V. et al. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front Microbiol* **10**, 1404 (2019).
10. Lee, S.Y. Metabolic engineering and synthetic biology in strain development. *ACS Synth Biol* **1**, 491-492 (2012).
11. Lee, S.Y. et al. A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis* **2**, 18-33 (2019).

12. Choi, K.R. et al. Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends Biotechnol* **37**, 817-837 (2019).

13. Ko, Y.S. et al. Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chem Soc Rev* **49**, 4615-4636 (2020).

14. Nielsen, J. & Keasling, J.D. Engineering Cellular Metabolism. *Cell* **164**, 1185-1197 (2016).

15. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).

16. Warner, J.R., Reeder, P.J., Karimpour-Fard, A., Woodruff, L.B. & Gill, R.T. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat Biotechnol* **28**, 856-862 (2010).

17. Zeitoun, R.I. et al. Multiplexed tracking of combinatorial genomic mutations in engineered cell populations. *Nat Biotechnol* **33**, 631-637 (2015).

18. McCarty, N.S., Graham, A.E., Studena, L. & Ledesma-Amaro, R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat Commun* **11**, 1281 (2020).

19. Williams, T.C., Pretorius, I.S. & Paulsen, I.T. Synthetic Evolution of Metabolic Productivity Using Biosensors. *Trends Biotechnol* **34**, 371-381 (2016).

20. Siedler, S. et al. SoxR as a single-cell biosensor for NADPH-consuming enzymes in Escherichia coli. *ACS Synth Biol* **3**, 41-47 (2014).

21. Dietrich, J.A., Shis, D.L., Alikhani, A. & Keasling, J.D. Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth Biol* **2**, 47-58 (2013).

22. Yang, J. et al. Synthetic RNA devices to expedite the evolution of metabolite-producing microbes. *Nat Commun* **4**, 1413 (2013).

23. Eckdahl, T.T. et al. Programmed evolution for optimization of orthogonal metabolic output in bacteria. *PLoS One* **10**, e0118322 (2015).

24. Trinh, C.T., Unrean, P. & Srienc, F. Minimal Escherichia coli cell for the most efficient production of ethanol from hexoses and pentoses. *Appl Environ Microbiol* **74**, 3634-3643 (2008).

25. Fong, S.S. et al. In silico design and adaptive evolution of Escherichia coli for production of lactic acid. *Biotechnol Bioeng* **91**, 643-648 (2005).

26. Raman, S., Rogers, J.K., Taylor, N.D. & Church, G.M. Evolution-guided optimization of biosynthetic pathways. *Proc Natl Acad Sci U S A* **111**, 17803-17808 (2014).

27. Trinh, C.T. & Mendoza, B. Modular cell design for rapid, efficient strain engineering toward industrialization of biology. *Curr Opin Chem Eng* **14**, 18-25 (2016).

28. Garcia, S. & Trinh, C.T. Harnessing Natural Modularity of Metabolism with Goal Attainment Optimization to Design a Modular Chassis Cell for Production of Diverse Chemicals. *ACS Synth Biol* **9**, 1665-1681 (2020).

29. Garcia, S. & Trinh, C.T. Multiobjective strain design: A framework for modular cell engineering. *Metab Eng* **51**, 110-120 (2019).

30. Wilbanks, B., Layton, D.S., Garcia, S. & Trinh, C.T. A Prototype for Modular Cell Engineering. *ACS Synth Biol* **7**, 187-199 (2018).

31. Salvador, F., Forza, C. & Rungtusanatham, M. Modularity, product variety, production volume, and component sourcing: theorizing beyond generic prescriptions. *Journal of Operations Management* **20**, 549-575 (2002).

32. Uetz, P. et al. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**, 623-627 (2000).

33. Kvasnicka, V. & Pospichal, J. Emergence of modularity in genotype-phenotype mappings. *Artif Life* **8**, 295-310 (2002).

34. Lorenz, D.M., Jeng, A. & Deem, M.W. The emergence of modularity in biological systems. *Phys Life Rev* **8**, 129-160 (2011).

35. Wagner, G.P., Pavlicev, M. & Cheverud, J.M. The road to modularity. *Nat Rev Genet* **8**, 921-931 (2007).

36. Melo, D., Porto, A., Cheverud, J.M. & Marroig, G. Modularity: genes, development and evolution. *Annu Rev Ecol Evol Syst* **47**, 463-486 (2016).

37. Garcia, S. & Trinh, C.T. Modular design: Implementing proven engineering principles in biotechnology. *Biotechnol Adv* **37**, 107403 (2019).

38. Lees, J.A., Galardini, M., Bentley, S.D., Weiser, J.N. & Corander, J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics* **34**, 4310-4312 (2018).

39. Gronlund, A., Bhalerao, R.P. & Karlsson, J. Modular gene expression in Poplar: a multilayer network approach. *New Phytol* **181**, 315-322 (2009).

40. Schroder, J. & Tauch, A. Transcriptional regulation of gene expression in Corynebacterium glutamicum: the role of global, master and local regulators in the modular and hierarchical gene regulatory network. *FEMS Microbiol Rev* **34**, 685-737 (2010).

41. Johnson, C.H., Ivanisevic, J. & Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat Rev Mol Cell Biol* **17**, 451-459 (2016).

42. Grilli, J., Rogers, T. & Allesina, S. Modularity and stability in ecological communities. *Nat Commun* **7**, 12031 (2016).

43. Huccetogullari, D., Luo, Z.W. & Lee, S.Y. Metabolic engineering of microorganisms for production of aromatic compounds. *Microb Cell Fact* **18**, 41 (2019).

44. Park, J.H., Lee, K.H., Kim, T.Y. & Lee, S.Y. Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc Natl Acad Sci U S A* **104**, 7797-7802 (2007).

45. Alper, H., Moxley, J., Nevoigt, E., Fink, G.R. & Stephanopoulos, G. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**, 1565-1568 (2006).

46. Tomazou, M., Barahona, M., Polizzi, K.M. & Stan, G.B. Computational Re-design of Synthetic Genetic Oscillators for Independent Amplitude and Frequency Modulation. *Cell Syst* **6**, 508-520 e505 (2018).

47. Hutchinson, C.R. Polyketide and non-ribosomal peptide synthases: falling together by coming apart. *Proc Natl Acad Sci U S A* **100**, 3010-3012 (2003).

48. Yuzawa, S. et al. Short-chain ketone production by engineered polyketide synthases in Streptomyces albus. *Nat Commun* **9**, 4569 (2018).

49. Jakociunas, T. et al. Programmable polyketide biosynthesis platform for production of aromatic compounds in yeast. *Synth Syst Biotechnol* **5**, 11-18 (2020).

50. Rodriguez, G.M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in Escherichia coli. *Nat Chem Biol* **10**, 259-265 (2014).

51. Layton, D.S. & Trinh, C.T. Microbial synthesis of a branched-chain ester platform from organic waste carboxylates. *Metab Eng Commun* **3**, 245-251 (2016).

52. Dellomonaco, C., Clomburg, J.M., Miller, E.N. & Gonzalez, R. Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**, 355-359 (2011).

53. Dekishima, Y., Lan, E.I., Shen, C.R., Cho, K.M. & Liao, J.C. Extending carbon chain length of 1-butanol pathway for 1-hexanol synthesis from glucose by engineered Escherichia coli. *J Am Chem Soc* **133**, 11399-11401 (2011).

54. Marcheschi, R.J. et al. A synthetic recursive "+1" pathway for carbon chain elongation. *ACS Chem Biol* **7**, 689-697 (2012).

55. Zhang, K., Sawaya, M.R., Eisenberg, D.S. & Liao, J.C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc Natl Acad Sci U S A* **105**, 20653-20658 (2008).

56. Houten, S.M. & Wanders, R.J. A general introduction to the biochemistry of mitochondrial fatty acid beta-oxidation. *J Inherit Metab Dis* **33**, 469-477 (2010).

57. Choi, S.Y. et al. Metabolic engineering for the synthesis of polyesters: A 100-year journey from polyhydroxyalkanoates to non-natural microbial polyesters. *Metab Eng* **58**, 47-81 (2020).

58. Seo, H., Lee, J.-W., Garcia, S. & Trinh, C.T. Single mutation at a highly conserved region of chloramphenicol acetyltransferase enables isobutyl acetate production directly from cellulose by Clostridium thermocellum at elevated temperatures. *Biotechnology for Biofuels* **12**, 245 (2019).

59. Seo, H., Lee, J.-W., Giannone, R.J., Dunlap, N.J. & Trinh, C.T. Repurposing chloramphenicol acetyltransferase for a robust and efficient designer ester biosynthesis platform. *bioRxiv*, 2020.2011.2004.368696 (2020).

60. Swiegers, J.H. & Pretorius, I.S. Yeast modulation of wine flavor. *Adv Appl Microbiol* **57**, 131-175 (2005).

61. Lee, J.W. & Trinh, C.T. Microbial biosynthesis of lactate esters. *Biotechnol Biofuels* **12**, 226 (2019).

62. Lange, J.P. et al. Valeric biofuels: a platform of cellulosic transportation fuels. *Angew Chem Int Ed Engl* **49**, 4479-4483 (2010).

63. Steen, E.J. et al. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-562 (2010).

64. Kantsa, A. et al. Community-wide integration of floral colour and scent in a Mediterranean scrubland. *Nat Ecol Evol* **1**, 1502-1510 (2017).

65. Layton, D.S. & Trinh, C.T. Engineering modular ester fermentative pathways in Escherichia coli. *Metab Eng* **26**, 77-88 (2014).

66. Tai, Y.S., Xiong, M. & Zhang, K. Engineered biosynthesis of medium-chain esters in Escherichia coli. *Metab Eng* **27**, 20-28 (2015).

67. Zhu, J., Lin, J.L., Palomec, L. & Wheeldon, I. Microbial host selection affects intracellular localization and activity of alcohol-O-acetyltransferase. *Microb Cell Fact* **14**, 35 (2015).

68. Urit, T., Li, M., Bley, T. & Loser, C. Growth of Kluyveromyces marxianus and formation of ethyl acetate depending on temperature. *Appl Microbiol Biotechnol* **97**, 10359-10371 (2013).

69. Kruis, A.J. et al. From Eat to trEat: engineering the mitochondrial Eat1 enzyme for enhanced ethyl acetate production in Escherichia coli. *Biotechnol Biofuels* **13**, 76 (2020).

70. Kleanthous, C. & Shaw, W.V. Analysis of the mechanism of chloramphenicol acetyltransferase by steady-state kinetics. Evidence for a ternary-complex mechanism. *Biochem J* **223**, 211-220 (1984).

71. Schwarz, S., Kehrenberg, C., Doublet, B. & Cloeckaert, A. Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS Microbiol Rev* **28**, 519-542 (2004).

72. Li, W.M., Ruf, S. & Bock, R. Chloramphenicol acetyltransferase as selectable marker for plastid transformation. *Plant Molecular Biology* **76**, 443-451 (2011).

73. Kamrud, K.I. et al. Detection of expressed chloramphenicol acetyltransferase in the saliva of Culex pipiens mosquitoes. *Insect Biochem Mol Biol* **27**, 423-429 (1997).

74. Shaw, W.V. Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. *Methods Enzymol* **43**, 737-755 (1975).

75. Alonso-Gutierrez, J. et al. Metabolic engineering of Escherichia coli for limonene and perillyl alcohol production. *Metab Eng* **19**, 33-41 (2013).

76. Bloom, J.D., Labthavikul, S.T., Otey, C.R. & Arnold, F.H. Protein stability promotes evolvability. *Proc Natl Acad Sci U S A* **103**, 5869-5874 (2006).

77. Aharoni, A. et al. The 'evolvability' of promiscuous protein functions. *Nat Genet* **37**, 73-76 (2005).

78. Agozzino, L. & Dill, K.A. Protein evolution speed depends on its stability and abundance and on chaperone concentrations. *Proc Natl Acad Sci U S A* **115**, 9092-9097 (2018).

79. Copley, S.D. Shining a light on enzyme promiscuity. *Curr Opin Struct Biol* **47**, 167-175 (2017).

80. M9 minimal medium (standard). *Cold Spring Harbor Protocols* **2010**, pdb.rec12295 (2010).

81. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870-1874 (2016).

82. Chenna, R. et al. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**, 3497-3500 (2003).

83. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324 (2014).

84. Biswas, T., Houghton, J.L., Garneau-Tsodikova, S. & Tsodikov, O.V. The structural basis for substrate versatility of chloramphenicol acetyltransferase CATI. *Protein Sci* **21**, 520-530 (2012).

85. Leslie, A.G. Refined crystal structure of type III chloramphenicol acetyltransferase at 1.75 A resolution. *J Mol Biol* **213**, 167-186 (1990).

86. Beaman, T.W., Sugantino, M. & Roderick, S.L. Structure of the hexapeptide xenobiotic acetyltransferase from Pseudomonas aeruginosa. *Biochemistry* **37**, 6689-6696 (1998).

87. Bordoli, L. et al. Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* **4**, 1-13 (2009).

88. Day, P.J., Shaw, W.V., Gibbs, M.R. & Leslie, A.G. Acetyl coenzyme A binding by chloramphenicol acetyltransferase: long-range electrostatic determinants of coenzyme A recognition. *Biochemistry* **31**, 4198-4205 (1992).

89. Gibson, D.G. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* **498**, 349-361 (2011).

90. Zheng, L., Baumann, U. & Reymond, J.L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res* **32**, e115 (2004).

91. Lee, J.-W. & Trinh, C.T. Microbial biosynthesis of lactate esters. *Biotechnology for Biofuels* **12**, 226 (2019).

92. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254 (1976).

93. Lo, M.C. et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* **332**, 153-159 (2004).

94. Ellman, G.L. Tissue sulfhydryl groups. *Arch Biochem Biophys* **82**, 70-77 (1959).

95. Gross, M.L. & University of Nebraska--Lincoln. High performance mass spectrometry : chemical applications : a symposium. (American Chemical Society, Washington; 1978).

96. Eric Jones, T.O., Pearu Peterson, and Others SciPy: Open source scientific tools for Python. **https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html** (2019).

97. Mayorov, N. Robust non linear regression in SciPy. **https://scipy-cookbook.readthedocs.io/items/robust_regression.html** (2019).

98. Marasovic, M., Marasovic, T. & Milos, M. Robust Nonlinear Regression in Enzyme Kinetic Parameters Estimation. *J Chem-Ny* (2017).

99. Taylor, M.P., Esteban, C.D. & Leak, D.J. Development of a versatile shuttle vector for gene expression in Geobacillus spp. *Plasmid* **60**, 45-52 (2008).

100. Groom, J. et al. Promiscuous plasmid replication in thermophiles: Use of a novel hyperthermophilic replicon for genetic manipulation of Clostridium thermocellum at its optimum growth temperature. *Metab Eng Commun* **3**, 30-38 (2016).

101. Mohr, G. et al. A Targetron System for Gene Targeting in Thermophiles and Its Application in Clostridium thermocellum. *Plos One* **8** (2013).

102. Kannuchamy, S., Mukund, N. & Saleena, L.M. Genetic engineering of Clostridium thermocellum DSM1313 for enhanced ethanol production. *BMC Biotechnol* **16 Suppl 1**, 34 (2016).

103. De Rossi, E., Brigidi, P., Welker, N.E., Riccardi, G. & Matteuzzi, D. New shuttle vector for cloning in Bacillus stearothermophilus. *Res Microbiol* **145**, 579-583 (1994).

104. Rhee, M.S., Kim, J.W., Qian, Y., Ingram, L.O. & Shanmugam, K.T. Development of plasmid vector and electroporation condition for gene transfer in sporogenic lactic acid bacterium, Bacillus coagulans. *Plasmid* **58**, 13-22 (2007).

105. Pillai, K.M.S. (ARIZONA STATE UNIVERSITY, 2016).

106. Zada, B. et al. Metabolic engineering of Escherichia coli for production of mixed isoprenoid alcohols and their derivatives. *Biotechnol Biofuels* **11**, 210 (2018).

107. Jang, H.J. et al. Selective retinol production by modulating the composition of retinoids from metabolically engineered E. coli. *Biotechnol Bioeng* **112**, 1604-1612 (2015).

108. Leslie, A.G., Moody, P.C. & Shaw, W.V. Structure of chloramphenicol acetyltransferase at 1.75-A resolution. *Proc Natl Acad Sci U S A* **85**, 4133-4137 (1988).

109. Kobayashi, J., Furukawa, M., Ohshiro, T. & Suzuki, H. Thermoadaptation-directed evolution of chloramphenicol acetyltransferase in an error-prone thermophile using improved procedures. *Appl Microbiol Biotechnol* **99**, 5563-5572 (2015).

110. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89 (2008).

111. Layton, D.S. & Trinh, C.T. Expanding the modular ester fermentative pathways for combinatorial biosynthesis of esters from volatile organic acids. *Biotechnol Bioeng* **113**, 1764-1776 (2016).

112. Murray, I.A., Lewendon, A. & Shaw, W.V. Stabilization of the imidazole ring of His-195 at the active site of chloramphenicol acetyltransferase. *J Biol Chem* **266**, 11695-11698 (1991).

113. Van der Schueren, J., Robben, J. & Volckaert, G. Misfolding of chloramphenicol acetyltransferase due to carboxy-terminal truncation can be corrected by second-site mutations. *Protein Eng* **11**, 1211-1217 (1998).

114. Taguchi, S. et al. A microbial factory for lactate-based polyesters using a lactate-polymerizing enzyme. *Proc Natl Acad Sci U S A* **105**, 17323-17327 (2008).

115. Yang, T.H. et al. Biosynthesis of polylactic acid and its copolymers using evolved propionate CoA transferase and PHA synthase. *Biotechnol Bioeng* **105**, 150-160 (2010).

116. Day, P.J. & Shaw, W.V. Acetyl coenzyme A binding by chloramphenicol acetyltransferase. Hydrophobic determinants of recognition and catalysis. *J Biol Chem* **267**, 5122-5127 (1992).

117. Guo, D. et al. Metabolic Engineering of Escherichia coli for Production of 2-Phenylethanol and 2-Phenylethyl Acetate from Glucose. *J Agric Food Chem* **66**, 5886-5891 (2018).

118. Rodriguez, G.M., Tashiro, Y. & Atsumi, S. Expanding ester biosynthesis in *Escherichia coli*. *Nature chemical biology* **10**, 259-265 (2014).

119. Tashiro, Y., Desai, S.H. & Atsumi, S. Two-dimensional isobutyl acetate production pathways to improve carbon yield. *Nat Commun* **6**, 7488 (2015).

120. Horton, C.E., Huang, K.-X., Bennett, G.N. & Rudolph, F.B. Heterologous expression of the Saccharomyces cerevisiae alcohol acetyltransferase genes in Clostridium acetobutylicum and Escherichia coli for the production of isoamyl acetate. *Journal of Industrial Microbiology and Biotechnology* **30**, 427-432 (2003).

121. Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. & Court, D.L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* **4**, 206-223 (2009).

122. Datta, S., Costantino, N. & Court, D.L. A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**, 109-115 (2006).

123. Cherepanov, P.P. & Wackernagel, W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14 (1995).

124. Seo, H.M. et al. Combinatorial application of two aldehyde oxidoreductases on isobutanol production in the presence of furfural. *J Ind Microbiol Biotechnol* **43**, 37-44 (2016).

125. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645 (2000).

126. Gibson, D.G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345 (2009).

127. Wilbanks, B. & Trinh, C.T. Comprehensive characterization of toxicity of fermentative metabolites on microbial growth. *Biotechnology for Biofuels* **10**, 262 (2017).

128. Liu, W. et al. Engineering Escherichia coli for high-yield geraniol production with biotransformation of geranyl acetate to geraniol under fed-batch culture. *Biotechnol Biofuels* **9**, 58 (2016).

129. Liu, Y. et al. High production of fatty alcohols in Escherichia coli with fatty acid starvation. *Microb Cell Fact* **15**, 129 (2016).

130. Shalit, M. et al. Volatile ester formation in roses. Identification of an acetyl-coenzyme A. Geraniol/Citronellol acetyltransferase in developing rose petals. *Plant Physiol* **131**, 1868-1876 (2003).

131. IndustryResearch (2021).

132. IndustryARC (2019).

133. Fang, D. et al. Metabolic and Process Engineering of Clostridium beijerinckii for Butyl Acetate Production in One Step. *J Agric Food Chem* **68**, 9475-9487 (2020).

134. Cumplido-Laso, G. et al. The fruit ripening-related gene FaAAT2 encodes an acyl transferase involved in strawberry aroma biogenesis. *J Exp Bot* **63**, 4275-4290 (2012).

135. Horton, C.E., Huang, K.X., Bennett, G.N. & Rudolph, F.B. Heterologous expression of the Saccharomyces cerevisiae alcohol acetyltransferase genes in Clostridium acetobutylicum and Escherichia coli for the production of isoamyl acetate. *J Ind Microbiol Biotechnol* **30**, 427-432 (2003).

136. Abe, F. & Horikoshi, K. Enhanced production of isoamyl alcohol and isoamyl acetate by ubiquitination-deficient Saccharomyces cerevisiae mutants. *Cell Mol Biol Lett* **10**, 383-388 (2005).

137. Vadali, R.V., Horton, C.E., Rudolph, F.B., Bennett, G.N. & San, K.Y. Production of isoamyl acetate in ackA-pta and/or ldh mutants of Escherichia coli with overexpression of yeast ATF2. *Appl Microbiol Biotechnol* **63**, 698-704 (2004).

138. Vadali, R.V., Bennett, G.N. & San, K.Y. Applicability of CoA/acetyl-CoA manipulation system to enhance isoamyl acetate production in Escherichia coli. *Metab Eng* **6**, 294-299 (2004).

139. Connor, M.R. & Liao, J.C. Engineering of an Escherichia coli strain for the production of 3-methyl-1-butanol. *Appl Environ Microbiol* **74**, 5769-5775 (2008).

140. Connor, M.R., Cann, A.F. & Liao, J.C. 3-Methyl-1-butanol production in Escherichia coli: random mutagenesis and two-phase fermentation. *Appl Microbiol Biotechnol* **86**, 1155-1164 (2010).

141. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89 (2008).

142. Tai, M. & Stephanopoulos, G. Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metab Eng* **15**, 1-9 (2013).

143. Mikhail Markovich Gusyatiner, M.G.L., Yuly Ivanovich Kozlov, Lirina Valerievna Ivanovskaya, Elvira Borisovna Voroshilova (Ajinomoto Co Inc, United States; 1999).

144. Wilbanks, B. & Trinh, C.T. Comprehensive characterization of toxicity of fermentative metabolites on microbial growth. *Biotechnol Biofuels* **10**, 262 (2017).

145. Lim, H.N., Lee, Y. & Hussein, R. Fundamental relationship between operon organization and gene expression. *Proceedings of the National Academy of Sciences* **108**, 10626-10631 (2011).

146. Schutze, A., Benndorf, D., Puttker, S., Kohrs, F. & Bettenbrock, K. The Impact of ackA, pta, and ackA-pta Mutations on Growth, Gene Expression and Protein Acetylation in Escherichia coli K-12. *Front Microbiol* **11**, 233 (2020).

147. Atsumi, S., Li, Z. & Liao, J.C. Acetolactate synthase from Bacillus subtilis serves as a 2-ketoisovalerate decarboxylase for isobutanol biosynthesis in Escherichia coli. *Appl Environ Microbiol* **75**, 6306-6311 (2009).

148. Saboe, P.O. et al. In situ product recovery of bio-based ethyl esters via hybrid extraction-distillation. *Green Chemistry* **21**, 5306-5315 (2019).

149. Dunlop, M.J. et al. Engineering microbial biofuel tolerance and export using efflux pumps. *Mol Syst Biol* **7**, 487 (2011).

150. Chacon, M.G., Marriott, A., Kendrick, E.G., Styles, M.Q. & Leak, D.J. Esterification of geraniol as a strategy for increasing product titre and specificity in engineered Escherichia coli. *Microb Cell Fact* **18**, 105 (2019).

151. Sattayawat, P., Yunus, I.S. & Jones, P.R. Bioderivatization as a concept for renewable production of chemicals that are toxic or poorly soluble in the liquid phase. *Proceedings of the National Academy of Sciences* **117**, 1404-1413 (2020).

152. Xiong, M. et al. A bio-catalytic approach to aliphatic ketones. *Sci Rep* **2**, 311 (2012).

153. Seo, H., Nicely, P.N. & Trinh, C.T. Endogenous carbohydrate esterases of Clostridium thermocellum are identified and disrupted for enhanced isobutyl acetate production from cellulose. *Biotechnol Bioeng* **117**, 2223-2236 (2020).

154. Doi, R.H. & Kosugi, A. Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Microbiol* **2**, 541-551 (2004).

155. Bayer, E.A., Belaich, J.P., Shoham, Y. & Lamed, R. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol* **58**, 521-554 (2004).

156. Gao, D., Luan, Y., Wang, Q., Liang, Q. & Qi, Q. Construction of cellulose-utilizing Escherichia coli based on a secretable cellulase. *Microb Cell Fact* **14**, 159 (2015).

157. Oh, E.J. & Jin, Y.S. Engineering of Saccharomyces cerevisiae for efficient fermentation of cellulose. *FEMS Yeast Res* **20** (2020).

158. Lynd, L.R., van Zyl, W.H., McBride, J.E. & Laser, M. Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol* **16**, 577-583 (2005).

159. Holwerda, E.K. et al. The exometabolome of Clostridium thermocellum reveals overflow metabolism at high cellulose loading. *Biotechnol Biofuels* **7**, 155 (2014).

160. Lin, P.P. et al. Consolidated bioprocessing of cellulose to isobutanol using *Clostridium thermocellum*. *Metab Eng* **31**, 44-52 (2015).

161. Akinosho, H., Yee, K., Close, D. & Ragauskas, A. The emergence of Clostridium thermocellum as a high utility candidate for consolidated bioprocessing applications. *Front Chem* **2**, 66 (2014).

162. Aharoni, A. et al. Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *Plant Cell* **12**, 647-662 (2000).

163. Stribny, J., Querol, A. & Perez-Torrado, R. Differences in Enzymatic Properties of the Saccharomyces kudriavzevii and Saccharomyces uvarum Alcohol Acetyltransferases and Their Impact on Aroma-Active Compounds Production. *Front Microbiol* **7**, 897 (2016).

164. Shalit, M. et al. Acetyl-coa: alcohol acetyltransferase activity and aroma formation in ripening melon fruits. *J Agric Food Chem* **49**, 794-799 (2001).

165. Beekwilder, J. et al. Functional characterization of enzymes forming volatile esters from strawberry and banana. *Plant Physiol* **135**, 1865-1878 (2004).

166. Argyros, D.A. et al. High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. *Appl Environ Microbiol* **77**, 8288-8294 (2011).

167. Olson, D.G. & Lynd, L.R. Transformation of Clostridium thermocellum by electroporation. *Methods Enzymol* **510**, 317-330 (2012).

168. Guss, A.M., Olson, D.G., Caiazza, N.C. & Lynd, L.R. Dcm methylation is detrimental to plasmid transformation in Clostridium thermocellum. *Biotechnol Biofuels* **5**, 30 (2012).

169. Bordoli, L. et al. Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* **4**, 1-13 (2009).

170. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**, D490-495 (2014).

171. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).

172. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680 (1994).

173. Thompson, R.A. et al. Elucidating central metabolic redox obstacles hindering ethanol production in Clostridium thermocellum. *Metab Eng* **32**, 207-219 (2015).

174. Batth, T.S. et al. Protein Aggregation Capture on Microparticles Enables Multipurpose Proteomics Sample Preparation. *Molecular & Cellular Proteomics* **18**, 1027-1035 (2019).

175. Walker, C., Ryu, S., Giannone, R.J., Garcia, S. & Trinh, C.T. Understanding and Eliminating the Detrimental Effect of Thiamine Deficiency on the Oleaginous Yeast Yarrowia lipolytica. *Appl Environ Microbiol* **86** (2020).

176. Taverner, T. et al. DanteR: an extensible R-based tool for quantitative analysis of -omics data. *Bioinformatics* **28**, 2404-2406 (2012).

177. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016).

178. Dubois, M., Gilles, K., Hamilton, J.K., Rebers, P.A. & Smith, F. A colorimetric method for the determination of sugars. *Nature* **168**, 167 (1951).

179. Fukuda, K. et al. Molecular cloning and nucleotide sequence of the isoamyl acetate-hydrolyzing esterase gene (EST2) from Saccharomyces cerevisiae. *J Ferment Bioeng* **82**, 8-15 (1996).

180. Fukuda, K. et al. Purification and characterization of isoamyl acetate-hydrolyzing esterase encoded by the IAH1 gene of Saccharomyces cerevisiae from a recombinant Escherichia coli. *Appl Microbiol Biotechnol* **53**, 596-600 (2000).

181. Fukuda, K. et al. Brewing properties of sake yeast whose EST2 gene encoding isoamyl acetate-hydrolyzing esterase was disrupted. *J Ferment Bioeng* **85**, 101-106 (1998).

182. Lilly, M. et al. The effect of increased yeast alcohol acetyltransferase and esterase activity on the flavour profiles of wine and distillates. *Yeast* **23**, 641-659 (2006).

183. Saerens, S.M., Delvaux, F.R., Verstrepen, K.J. & Thevelein, J.M. Production and biological function of volatile esters in Saccharomyces cerevisiae. *Microb Biotechnol* **3**, 165-177 (2010).

184. Bomble, Y.J. et al. Lignocellulose deconstruction in the biosphere. *Curr Opin Chem Biol* **41**, 61-70 (2017).

185. Himmel, M.E. et al. Microbial enzyme systems for biomass conversion: emerging paradigms. *Biofuels* **1**, 323-341 (2010).

186. Pawar, P.M., Koutaniemi, S., Tenkanen, M. & Mellerowicz, E.J. Acetylation of woody lignocellulose: significance and regulation. *Front Plant Sci* **4**, 118 (2013).

187. Flamholz, A., Noor, E., Bar-Even, A. & Milo, R. eQuilibrator--the biochemical thermodynamics calculator. *Nucleic Acids Res* **40**, D770-775 (2012).

188. Faber, K. Biotransformations in organic chemistry : a textbook, Edn. 5th rev. & corr. (Springer-Verlag, Berlin ; New York; 2004).

189. Kruis, A.J. et al. Ethyl acetate production by the elusive alcohol acetyltransferase from yeast. *Metab Eng* **41**, 92-101 (2017).

190. Kruis, A.J. et al. Microbial production of short and medium chain esters: Enzymes, pathways, and applications. *Biotechnol Adv* (2019).

191. Vincent, F. et al. Multifunctional xylooligosaccharide/cephalosporin C deacetylase revealed by the hexameric structure of the Bacillus subtilis enzyme at 1.9A resolution. *J Mol Biol* **330**, 593-606 (2003).

192. Kauppinen, S. et al. Molecular cloning and characterization of a rhamnogalacturonan acetylesterase from Aspergillus aculeatus. Synergism between rhamnogalacturonan degrading enzymes. *J Biol Chem* **270**, 27172-27178 (1995).

193. Navarro-Fernandez, J. et al. Characterization of a new rhamnogalacturonan acetyl esterase from Bacillus halodurans C-125 with a new putative carbohydrate binding domain. *J Bacteriol* **190**, 1375-1382 (2008).

194. Nataf, Y. et al. Clostridium thermocellum cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. *Proc Natl Acad Sci U S A* **107**, 18646-18651 (2010).

195. Munoz-Gutierrez, I. et al. Decoding Biomass-Sensing Regulons of Clostridium thermocellum Alternative Sigma-I Factors in a Heterologous Bacillus subtilis Host System. *PLoS One* **11**, e0146316 (2016).

196. Kahel-Raifer, H. et al. The unique set of putative membrane-associated anti-sigma factors in Clostridium thermocellum suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. *FEMS Microbiol Lett* **308**, 84-93 (2010).

197. Xu, Q. et al. Dramatic performance of Clostridium thermocellum explained by its wide range of cellulase modalities. *Sci Adv* **2**, e1501254 (2016).

198. Aurilia, V. et al. Three multidomain esterases from the cellulolytic rumen anaerobe Ruminococcus flavefaciens 17 that carry divergent dockerin sequences. *Microbiology* **146 (Pt 6)**, 1391-1397 (2000).

199. Flint, H.J., McPherson, C.A. & Martin, J. Expression of two xylanase genes from the rumen cellulolytic bacterium Ruminococcus flavefaciens 17 cloned in pUC13. *J Gen Microbiol* **137**, 123-129 (1991).

200. Biely, P., Mackenzie, C.R., Puls, J. & Schneider, H. Cooperativity of Esterases and Xylanases in the Enzymatic Degradation of Acetyl Xylan. *Bio-Technol* **4**, 731-733 (1986).

201. Poudel, S. et al. Integrated omics analyses reveal the details of metabolic adaptation of Clostridium thermocellum to lignocellulose-derived growth inhibitors released during the deconstruction of switchgrass. *Biotechnol Biofuels* **10**, 14 (2017).

202. Wei, H. et al. Comparison of transcriptional profiles of Clostridium thermocellum grown on cellobiose and pretreated yellow poplar using RNA-Seq. *Front Microbiol* **5**, 142 (2014).

203. Ciechanover, A. Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nat Rev Mol Cell Biol* **6**, 79-87 (2005).

204. Soh, L.M.J. et al. Engineering a Thermostable Keto Acid Decarboxylase Using Directed Evolution and Computationally Directed Protein Design. *ACS Synth Biol* **6**, 610-618 (2017).

205. Biswas, R., Prabhu, S., Lynd, L.R. & Guss, A.M. Increase in ethanol yield via elimination of lactate production in an ethanol-tolerant mutant of Clostridium thermocellum. *PLoS One* **9**, e86389 (2014).

206. Hon, S. et al. Expressing the Thermoanaerobacterium saccharolyticum pforA in engineered Clostridium thermocellum improves ethanol production. *Biotechnol Biofuels* **11**, 242 (2018).

207. Thompson, R.A. & Trinh, C.T. Overflow metabolism and growth cessation in Clostridium thermocellum DSM1313 during high cellulose loading fermentations. *Biotechnol Bioeng* **114**, 2592-2604 (2017).

208. Xiong, W. et al. CO2-fixing one-carbon metabolism in a cellulose-degrading bacterium Clostridium thermocellum. *Proc Natl Acad Sci U S A* **113**, 13180-13185 (2016).

209. Tian, L. et al. Simultaneous achievement of high ethanol yield and titer in Clostridium thermocellum. *Biotechnol Biofuels* **9**, 116 (2016).

210. Thompson, R.A. et al. Elucidating central metabolic redox obstacles hindering ethanol production in *Clostridium thermocellum*. *Metabolic Engineering* **32**, 207-219 (2015).

211. Lescic Asler, I. et al. Probing enzyme promiscuity of SGNH hydrolases. *Chembiochem* **11**, 2158-2167 (2010).

212. Ganguly, J. Studies on the mechanism of fatty acid synthesis. VII. Biosynthesis of fatty acids from malonyl CoA. *Biochim Biophys Acta* **40**, 110-118 (1960).

213. Lo, Y.C., Lin, S.C., Shaw, J.F. & Liaw, Y.C. Substrate specificities of Escherichia coli thioesterase I/protease I/lysophospholipase L1 are governed by its switch loop movement. *Biochemistry* **44**, 1971-1979 (2005).

214. Bonner, W.M. & Bloch, K. Purification and properties of fatty acyl thioesterase I from Escherichia coli. *J Biol Chem* **247**, 3123-3133 (1972).

215. Olson, D.G. et al. Identifying promoters for gene expression in Clostridium thermocellum. *Metabolic Engineering Communications* **2**, 23-29 (2015).

216. Tian, L. et al. Enhanced ethanol formation by Clostridium thermocellum via pyruvate decarboxylase. *Microb Cell Fact* **16**, 171 (2017).

217. Kim, S.K., Groom, J., Chung, D., Elkins, J. & Westpheling, J. Expression of a heat-stable NADPH-dependent alcohol dehydrogenase from Thermoanaerobacter pseudethanolicus 39E in Clostridium thermocellum 1313 results in increased hydroxymethylfurfural resistance. *Biotechnol Biofuels* **10**, 66 (2017).

218. Fujitani, Y., Yamamoto, K. & Kobayashi, I. Dependence of frequency of homologous recombination on the homology length. *Genetics* **140**, 797-809 (1995).

219. Unkrig, V., Neugebauer, F.A. & Knappe, J. The free radical of pyruvate formate-lyase. Characterization by EPR spectroscopy and involvement in catalysis as studied with the substrate-analogue hypophosphite. *Eur J Biochem* **184**, 723-728 (1989).

220. Marcano-Velazquez, J.G., Lo, J., Nag, A., Maness, P.C. & Chou, K.J. Developing Riboswitch-Mediated Gene Regulatory Controls in Thermophilic Bacteria. *ACS Synth Biol* **8**, 633-640 (2019).

221. Tian, L. et al. Metabolic engineering of Clostridium thermocellum for n-butanol production from cellulose. *Biotechnol Biofuels* **12**, 186 (2019).

222. Lopez-Gresa, M.P. et al. A New Role For Green Leaf Volatile Esters in Tomato Stomatal Defense Against Pseudomonas syringe pv. tomato. *Front Plant Sci* **9**, 1855 (2018).

223. Ameye, M. et al. Green leaf volatile production by plants: a meta-analysis. *New Phytol* **220**, 666-683 (2018).

224. Kessler, A. & Baldwin, I.T. Defensive function of herbivore-induced plant volatile emissions in nature. *Science* **291**, 2141-2144 (2001).

225. Aharoni, A. et al. Terpenoid metabolism in wild-type and transgenic Arabidopsis plants. *Plant Cell* **15**, 2866-2884 (2003).

226. Klatt, B.K., Burmeister, C., Westphal, C., Tscharntke, T. & von Fragstein, M. Flower volatiles, crop varieties and bee responses. *PLoS One* **8**, e72724 (2013).

227. Pichersky, E. & Gershenzon, J. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr Opin Plant Biol* **5**, 237-243 (2002).

228. Keller, L. & Surette, M.G. Communication in bacteria: an ecological and evolutionary perspective. *Nat Rev Microbiol* **4**, 249-258 (2006).

229. Hays, S.G., Yan, L.L.W., Silver, P.A. & Ducat, D.C. Synthetic photosynthetic consortia define interactions leading to robustness and photoproduction. *J Biol Eng* **11**, 4 (2017).

230. Dwidar, M. et al. Co-culturing a novel Bacillus strain with Clostridium tyrobutyricum ATCC 25755 to produce butyric acid from sucrose. *Biotechnol Biofuels* **6**, 35 (2013).

231. Beri, D. et al. Coculture with hemicellulose-fermenting microbes reverses inhibition of corn fiber solubilization by Clostridium thermocellum at elevated solids loadings. *Biotechnology for Biofuels* **14**, 24 (2021).

232. Brenner, K., You, L. & Arnold, F.H. Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol* **26**, 483-489 (2008).

233. Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *P Natl Acad Sci USA* **102**, 12678-12683 (2005).

234. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946-950 (2009).

235. Tyo, K.E., Ajikumar, P.K. & Stephanopoulos, G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nat Biotechnol* **27**, 760-765 (2009).

236. Lechner, A., Brunk, E. & Keasling, J.D. The Need for Integrated Approaches in Metabolic Engineering. *Cold Spring Harb Perspect Biol* **8** (2016).

237. Ajikumar, P.K. et al. Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* **330**, 70-74 (2010).

238. Zaldivar, J., Nielsen, J. & Olsson, L. Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Appl Microbiol Biotechnol* **56**, 17-34 (2001).

239. Zhang, H. & Wang, X. Modular co-culture engineering, a new approach for metabolic engineering. *Metab Eng* **37**, 114-121 (2016).

240. Saini, M., Hong Chen, M., Chiang, C.J. & Chao, Y.P. Potential production platform of n-butanol in Escherichia coli. *Metab Eng* **27**, 76-82 (2015).

241. Zhou, K., Qiao, K., Edgar, S. & Stephanopoulos, G. Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nature Biotechnology* **33**, 377-383 (2015).

242. Zhang, H., Pereira, B., Li, Z. & Stephanopoulos, G. Engineering Escherichia coli coculture systems for the production of biochemical products. *Proc Natl Acad Sci U S A* **112**, 8266-8271 (2015).

243. Hong, J., Im, D.K. & Oh, M.K. Investigating E. coli Coculture for Resveratrol Production with (13)C Metabolic Flux Analysis. *J Agric Food Chem* **68**, 3466-3473 (2020).

244. Wu, G. et al. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol* **34**, 652-664 (2016).

245. Yu, D. et al. An efficient recombination system for chromosome engineering in Escherichia coli. *Proc Natl Acad Sci U S A* **97**, 5978-5983 (2000).

246. Carroll, A. & Somerville, C. Cellulosic biofuels. *Annu Rev Plant Biol* **60**, 165-182 (2009).

247. Aidelberg, G. et al. Hierarchy of non-glucose sugars in Escherichia coli. *BMC Syst Biol* **8**, 133 (2014).

248. Kim, J.H., Block, D.E. & Mills, D.A. Simultaneous consumption of pentose and hexose sugars: an optimal microbial phenotype for efficient fermentation of lignocellulosic biomass. *Appl Microbiol Biotechnol* **88**, 1077-1085 (2010).

249. Wu, Y., Shen, X., Yuan, Q. & Yan, Y. Metabolic Engineering Strategies for Co-Utilization of Carbon Sources in Microbes. *Bioengineering (Basel)* **3** (2016).

250. Fox, K.J. & 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* **30**, 9-16 (2020).

251. Xia, T., Eiteman, M.A. & Altman, E. Simultaneous utilization of glucose, xylose and arabinose in the presence of acetate by a consortium of Escherichia coli strains. *Microb Cell Fact* **11**, 77 (2012).

252. Klamt, S. Generalized concept of minimal cut sets in biochemical networks. *Biosystems* **83**, 233-247 (2006).

253. Klamt, S. & Gilles, E.D. Minimal cut sets in biochemical reaction networks. *Bioinformatics* **20**, 226-234 (2004).

254. Sievert, C. et al. Experimental evolution reveals an effective avenue to release catabolite repression via mutations in XylR. *Proc Natl Acad Sci U S A* **114**, 7349-7354 (2017).

255. Li, Z., Wang, X. & Zhang, H. Balancing the non-linear rosmarinic acid biosynthetic pathway by modular co-culture engineering. *Metab Eng* **54**, 1-11 (2019).

256. Qureshi, N. & Blaschek, H.P. Butanol production using Clostridium beijerinckii BA101 hyper-butanol producing mutant strain and recovery by pervaporation. *Appl Biochem Biotechnol* **84-86**, 225-235 (2000).

257. Kim, H., Jeon, B.S. & Sang, B.I. An Efficient New Process for the Selective Production of Odd-Chain Carboxylic Acids by Simple Carbon Elongation Using Megasphaera hexanoica. *Sci Rep* **9**, 11999 (2019).

258. Jeon, B.S., Choi, O., Um, Y. & Sang, B.I. Production of medium-chain carboxylic acids by Megasphaera sp. MH with supplemental electron acceptors. *Biotechnol Biofuels* **9**, 129 (2016).

259. Wu, Q. et al. Medium chain carboxylic acids production from waste biomass: Current advances and perspectives. *Biotechnol Adv* **37**, 599-615 (2019).

260. Christiaens, J.F. et al. The fungal aroma gene ATF1 promotes dispersal of yeast cells through insect vectors. *Cell Rep* **9**, 425-432 (2014).

261. van Rijswijck, I.M.H., Wolkers-Rooijackers, J.C.M., Abee, T. & Smid, E.J. Performance of non-conventional yeasts in co-culture with brewers' yeast for steering ethanol and aroma production. *Microb Biotechnol* **10**, 1591-1602 (2017).

Vita

Hyeongmin Seo (Hyung-Min Seo) was born to Yoo-ok Seo and Young-Sook Cho on September 23, 1989 in Anyang-si, Gyeonggi-do, South Korea. Hyeongmin graduated from Dong-An High School in 2008. He attended Konkuk University, earning Bachelor of Science degree of Biological Engineering with microbial engineering concentration in 2014. During Sophomore, he joined enzyme engineering laboratory and assisted Dr. Sang-Hyun Lee's projects. After two years of military service from 2011 to 2013 as Korean Augmentation to the United States Army (KATUSA), he worked with Dr. Yung-Hun Yang for three years as undergraduate research assistant and post-bachelor research assistant. During this time, Hyeongmin was fascinated by metabolic engineering and synthetic biology and decided to study abroad in the United States. Thereafter, he joined Dr. Cong T. Trinh's group in 2016. He plans to continue his career in the field of metabolic engineering as an academic researcher and teacher.