Strain engineering and bioprocess development for bio-based production of porphyrins

by

Bahareh Arab

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2025

© Bahareh Arab 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Guneet Kaur Assistant professor, School of Engineering, University of Guelph
Supervisor(s):	Perry Chou Professor, Dept. of Chemical Engineering, University of Waterloo
	Murray Moo Young Professor Emeritus, Dept. of Chemical Engineering, University of Waterloo
Internal Member:	Yilan Liu Assistant professor, Dept. of Chemical Engineering, University of Waterloo
Internal Member:	Valerie Ward Associate professor, Dept. of Chemical Engineering, University of Waterloo
Internal-External Member:	Subha Kalyaanamoorthy Assistant professor, Dept. of Chemistry, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

I was the sole author for Chapters 1, 2, and 8 which were written under the supervision of Dr. Perry Chou and Dr. Murray Moo-Young.

This thesis consists in part of five manuscripts written for publication. As the lead author of all of these chapters, I contributed to conceptualizing the study design, conducting data collection and analysis, and drafting the manuscripts. My coauthors provided guidance throughout the research process and provided feedback on manuscript drafts.

Citations:

Chapter 3 **Arab, B.**, Westbrook, A., Moo-Young, M., & Chou, C. H. P. (2022). A toolkit for effective and successive genome engineering of *Escherichia coli*. Fermentation, 9(1), 14.

Chapter 4 **Arab, B.**, Westbrook, A. W., Moo-Young, M., Liu, Y., & Chou, C. P. (2024). Bio-Based Production of Uroporphyrin in *Escherichia coli*. Synthetic Biology and Engineering, 2(1), 10002.

Chapter 5 **Arab, B.**, Westbrook, A., Moo-Young, M., Liu, Y., & Chou, C. P. (2024). High-Level Bio-Based Production of Coproporphyrin in *Escherichia coli*. Fermentation, 10(5), 250.

Chapter 6 **Arab, B.**, Moo-Young, M., Liu, Y., & Chou, C. P. Two-Step Bio-Based Production of Heme: In Vivo Cell Cultivation Followed by In Vitro Enzymatic Conversion. Fermentation, Submission No: 3544795 [Accepted].

Chapter 7 **Arab, B.**, Moo-Young, M., Liu, Y., & Chou, C. P. (2025). Manipulating Intracellular Oxidative Conditions to Enhance Porphyrin Production in *Escherichia coli*. Bioengineering, 12(1), 83.

Abstract

Bio-based production using microbial cell factories has emerged as a transformative approach to addressing the limitations of petrochemical processes, offering renewable, sustainable, and environmentally friendly alternatives for manufacturing valuable chemical compounds. Among various microbial systems, *Escherichia coli* (*E. coli*) has become a popular and versatile host for biomanufacturing due to its rapid growth, genetic tractability, and extensive history of industrial use. Through advances in synthetic biology, genome engineering, and metabolic engineering, *E. coli* can be tailored to produce a wide array of chemicals, including structurally complex compounds like porphyrins. Porphyrins and their derivatives, such as heme and chlorophyll, are critical for various biological and industrial applications, ranging from pharmaceuticals and diagnostics to renewable energy solutions. Despite their significance, challenges such as pathway bottlenecks, feedback inhibition, and intracellular toxicity of intermediates hinder microbial production of porphyrins. This thesis addresses these challenges by implementing integrated engineering strategies to enhance porphyrin biosynthesis in *E. coli* and establish a robust framework for scalable production.

A foundation of this work was the development of a genome engineering toolkit that integrates CRISPR-Cas9 and transposon-based methods. This system allowed for site-specific insertion of heterologous genes and precise inactivation of endogenous genes, achieving editing efficiencies exceeding 90%. Such flexible manipulation of the *E. coli* genome facilitated the construction of optimized strains for biomanufacturing. For example, the toolkit enabled the creation of a plasmid-free strain capable of producing polyhydroxyalkanoates, demonstrating its potential for industrial applications and providing the foundation for metabolic engineering efforts targeting porphyrin biosynthesis.

For the subsequent part of our thesis, the biosynthesis of uroporphyrin (UP), a key precursor for heme, was enhanced by implementing the Shemin/C4 pathway in *E. coli*. Strategies to increase intracellular succinyl-CoA availability and express a synthetic operon containing genes such as *hemA*, *hemB*, *hemC*, and *hemD* led to UP titers of 901.9 mg/L under batch bioreactor conditions. Furthermore, most of the UP produced was secreted extracellularly, simplifying downstream purification and demonstrating the feasibility of large-scale production. These advancements highlight the effectiveness of pathway optimization in overcoming metabolic bottlenecks.

We used the information obtained from previous chapter to enhance coproporphyrin (CP) biosynthesis. Dual synthetic operons controlled by strong promoters regulated key pathway genes, including *hemA*, *hemB*, *hemD*, *hemE*, and *hemY*. Bioreactor cultivation of the engineered strains using glycerol as the primary carbon source under aerobic conditions

led to CP titers of up to 353 mg/L with minimal byproduct formation. To the best of our knowledge this study marked the first targeted bio-based production of CP in *E. coli*, laying the groundwork for its industrial-scale synthesis and emphasizing the importance of precise gene regulation in pathway optimization.

Addressing the complexities of heme biosynthesis required a novel two-step strategy integrating in-vivo and in-vitro approaches. Engineered *E. coli* strains expressing the coproporphyrin-dependent (CPD) pathway produced ∼85 mg/L of coproheme and ∼18 mg/L of heme in-vivo. However, intracellular heme accumulation posed significant toxicity challenges due to limited secretion into the extracellular medium. These challenges were mitigated by developing an optimized in-vitro enzymatic conversion process, achieving a 77.2% reaction yield for the conversion of coproporphyrin III to coproheme and a 45.8% yield for the conversion of coproheme to heme. This integrated approach bypassed intracellular toxicity, enabling controlled and scalable production while addressing key bottlenecks in microbial production systems.

In our final study, to further enhance porphyrin biosynthesis, strategies were developed to mitigate reactive oxygen species (ROS)-induced stress and redirect dissimilated carbon flux toward type-III porphyrin biosynthesis. Antioxidant supplementation with ascorbic acid (up to 1 g/L) improved the UP-III/UP-I ratio from 0.62 to 2.57, enhancing the production of type-III porphyrins. Additionally, overexpression of ROS-scavenging genes such as *sodA* and *kat* significantly increased porphyrin yields. Notably, overexpression of *sodA* alone resulted in a 72.9% increase in total porphyrin production, reaching titers of 1.56 g/L, while improving the UP-III/UP-I ratio to 1.94. These findings underscore the importance of addressing oxidative stress to optimize metabolic fluxes and enhance type-III porphyrin biosynthesis in *E. coli*. The study provides a practical platform for improving bio-based porphyrin production at industrial scales.

Taken together, this thesis demonstrates the potential of integrating strain engineering, synthetic biology, and metabolic engineering to enhance porphyrin biosynthesis in *E. coli*. The innovative strategies developed provide scalable, sustainable, and economically viable solutions for producing porphyrins and their derivatives. These advancements open new avenues for industrial applications in pharmaceuticals, diagnostics, and renewable energy, establishing *E. coli* as a powerful platform for biomanufacturing complex biomolecules.

Acknowledgements

First and foremost, I extend my heartfelt appreciation to my supervisors Professors Perry Chou and Murray Moo-Young. Professor Chou provided invaluable mentorship, insightful guidance, and unwavering support throughout my doctoral studies. His expertise and commitment to excellence have been instrumental in pushing the boundaries of my research and shaping me as an independent researcher. I am deeply grateful for his dedication to promoting innovation and rigor, which has been a constant source of inspiration. Professor Moo-Young, a pioneer in the field of biotechnology, has offered profound wisdom and encouragement throughout this journey. His extensive knowledge and perspective have been invaluable, and I am honored to have had the opportunity to learn from someone whose contributions to the field are unparalleled.

I sincerely thank my committee members, Professors Yilan Liu, Valerie Ward, and Subha Kalyaanamoorthy, for their invaluable contributions to my research. Their constructive feedback, thought-provoking discussions, and expertise have greatly enriched this work and shaped it into its final form. I am also honored to have Professor Guneet Kaur from the University of Guelph as my external examiner and look forward to benefiting from her expertise during my PhD defense.

I am deeply thankful to the staff of the Department of Chemical Engineering for their administrative and technical support. Their professionalism and dedication ensured a smooth progression of my studies and experiments.

To my colleagues and associates, I am deeply thankful for the collaboration, support, and shared learning experiences throughout my journey. I extend a special thanks to Adam Westbrook for his guidance and expertise, particularly during my early days in the lab, which helped lay the groundwork for my research. I am also grateful to the co-op students and research assistants whose contributions and efforts streamlined many aspects of my experiments.

Finally, my deepest appreciation is reserved for my family. To my parents, Pari and Reza, who have been my unwavering pillars of support, thank you for your unconditional love, sacrifices, and encouragement. To my sister, Pegah, your constant support and encouragement have been an invaluable source of inspiration. And to my husband, Sadegh, no words can truly convey the depth of my gratitude for your endless patience, love, and unwavering confidence in me. Thank you for understanding me in ways no one else can, for supporting me through every step of this journey, and for being my anchor in moments of uncertainty. Your partnership has made every challenge seem conquered and every success more meaningful.

Dedication

To my Mother, Father, Pegah and Sadegh.

Table of Contents

Examining Committee ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgements vii

Dedication viii

List of Figures xiii

List of Tables xxii

List of abbreviations xiii

List of symbols xvi

1 Introduction 1
- 1.1 Research Objectives . . . 2
- 1.2 Thesis Organization . . . 3

2 Literature Review 4
- 2.1 Biomanufacturing . . . 4
  - 2.1.1 Comparison of Biological Methods with Chemical Synthesis . . . 5
  - 2.1.2 Genetic and Metabolic Engineering in Biomanufacturing . . . 6
- 2.2 Advantages of *E. coli* in Biomanufacturing . . . 7
  - 2.2.1 Historical Context and Model Organism Status . . . 7

2.2.2 Comparison with Alternative Hosts . . . 8
2.3 Genetic and Metabolic Engineering in *E. coli* . . . 8
2.3.1 Genome Engineering Methods in *E. coli* . . . 8
2.3.2 CRISPR-Cas9: Mechanism and Applications . . . 10
2.3.3 Transposons: Mechanism and Applications . . . 11
2.4 Overview of Porphyrin Biosynthesis . . . 12
2.4.1 Porphyrins . . . 12
2.4.2 Porphyrin Biosynthesis pathway . . . 13
2.4.3 Regulation of Porphyrin Biosynthesis . . . 16
2.5 Reactive Oxygen Species (ROS) and Their Impact . . . 17
2.5.1 Biochemical Basis of ROS Generation in Porphyrin Biosynthesis . . . 17
2.5.2 Cellular Impact of ROS . . . 17
2.5.3 Strategies to Mitigate ROS Effects . . . 17
2.5.4 Challenges of ROS Management in Industrial Bioprocesses . . . 18
2.6 Glycerol as an Ideal Feedstock . . . 19

3 A toolkit for effective and successive genome engineering of *Escherichia coli* 21
3.1 Introduction . . . 21
3.2 Materials and Methods . . . 28
3.2.1 Bacterial strains, plasmids, editing template, and oligonucleotides . . . 28
3.2.2 Competent cells preparation and transformations . . . 31
3.2.3 Media and bacterial cell cultivation . . . 31
3.2.4 PHBV extraction and analysis . . . 32
3.2.5 Statistical analysis . . . 33
3.3 Results . . . 33
3.3.1 Genomic knock-in based on CRISPR-Cas9 coupled with λ_{Red} recombineering . . . 33
3.3.2 Genomic knock-in based on transposon-associated CRISPR-Cas system 35
3.3.3 Gene knockout based on CRISPR-Cas9 coupled with λ_{Red} recombineering . . . 35
3.3.4 PHBV biosynthesis using genome-engineered strains . . . 37
3.4 Discussion . . . 38

4 Bio-Based Production of Uroporphyrin in *Escherichia coli* 42
4.1 Introduction . . . 42
4.2 Materials and Methods . . . 45
4.2.1 Bacterial Strains and Plasmids . . . 45

4.2.2 Media and Bacterial Cell Cultivation . . . 47
4.2.3 Analysis . . . 48
4.3 Results . . . 49
4.3.1 Implementation of the Shemin/C4 Pathway for UP Biosynthesis . . . 49
4.3.2 Effects of the Native HemB on UP Biosynthesis . . . 53
4.3.3 Effects of HemC from *E. coli* and *B. Subtilis* on UP Biosynthesis . . . 55
4.3.4 Effects of the Native HemD on UP Biosynthesis . . . 59
4.4 Discussion . . . 64

5 High-Level Bio-Based Production of Coproporphyrin in *Escherichia coli* 67
5.1 Introduction . . . 67
5.2 Materials and Methods . . . 71
5.2.1 Bacterial Strains and Plasmids . . . 71
5.2.2 Media and Bacterial Cell Cultivation . . . 73
5.2.3 Analysis . . . 74
5.3 Results . . . 75
5.3.1 Effects of the Native HemD for CP Biosynthesis . . . 75
5.3.2 Effects of the Native HemD/E for CP Biosynthesis . . . 78
5.3.3 Tuning of Expression Levels of HemD/E for CP Biosynthesis . . . 80
5.3.4 Effects of HemY on CP Biosynthesis . . . 82
5.3.5 HemD Overexpression Was Not Essential for CP Biosynthesis . . . 86
5.4 Discussion . . . 88

6 Two-Step Bio-Based Production of Coproheme and Heme in *Escherichia coli*: In Vivo and In Vitro Approaches 91
6.1 Introduction . . . 91
6.2 Materials and methods . . . 97
6.2.1 Bacterial strains and plasmids . . . 97
6.2.2 Media and bacterial cell cultivation . . . 100
6.2.3 Enzyme purification . . . 100
6.2.4 In vitro experiments reaction set up . . . 101
6.2.5 Intracellular porphyrin extraction . . . 101
6.2.6 Analysis . . . 102
6.3 Results . . . 103
6.3.1 Effects of HemH on Fe-CP-III biosynthesis . . . 103
6.3.2 Effects of HemQ on heme biosynthesis . . . 106
6.3.3 In vitro conversion of CP-III to Fe-CP-III via HemH . . . 108
6.3.4 In vitro conversion of Fe-CP-III to heme via HemQ . . . 111

6.4 Discussion . . . 113

7 Manipulating Intracellular Oxidative Conditions to Enhance Porphyrin Production in *Escherichia coli* 116
7.1 Introduction . . . 116
7.2 Materials and Methods . . . 120
7.2.1 Bacterial Strains and Plasmids . . . 120
7.2.2 Media and Bacterial Cell Cultivation . . . 124
7.2.3 Analysis . . . 124
7.3 Results . . . 125
7.3.1 Effects of Ascorbic Acid Supplementation on UP Biosynthesis . . . 125
7.3.2 Effects of SODs on UP Biosynthesis . . . 128
7.3.3 Effects of Catalases on UP Biosynthesis . . . 131
7.4 Discussion . . . 133

8 Original Contributions, Recommendations, and Future Prospects 136
8.1 Original Contributions . . . 136
8.1.1 Establishing a versatile genome engineering toolkit for manipulating *E. coli* . . . 136
8.1.2 Developing efficient production of uroporphyrin in *E. coli* . . . 137
8.1.3 Bioprocess engineering toward high-level coproporphyrin production in *E. coli* . . . 137
8.1.4 Developing a two-step strategy for coproheme and heme production via heterologous coproporphyrin-dependent (CPD) pathway . . . 138
8.1.5 Mitigating oxidative stress to enhance porphyrin biosynthetic efficiency 138
8.2 Recommendations and Future Prospects . . . 139

References 141

APPENDICES 171
Supplementary Tables . . . 172
Supplementary Figures . . . 189

List of Figures

2.1 DNA targeting and cleavage by Cas9-gRNA complex 11
2.2 Overview of heme biosynthesis pathways. The purple arrow represents the Chemin/C4 pathway; the gray arrows represents the PPD pathway found in Gram-negative bacteria and eukaryotes; the pink arrows represents the SHD pathway from archaea and sulfate-reducing bacteria; and the green arrows represents the CPD pathway found in Gram-positive bacteria. 15

3.1 Schematic representation of the two-plasmid system for genomic integration via transposon-associated CRISPR-Cas system. (**a**) CRISPR-Cas9 coupled with λ_{Red} recombineering mechanism: The guide RNA (gRNA) comprised of a protospacer (PS), Cas9-binding hairpin (CBH) and transcriptional terminator (ter), directs the Cas9 to the desired locus by recognizing the PAM sequence and the Cas9 cleaves the DNA through double strand breaks. Exo (lambda exonuclease) generates single-stranded 3' overhangs and Beta proteins mediate annealing of the homology arms (HA) to the complementary DNA target in the genome. (**b**) Transposition mechanism: pEffector plasmid harboring the crRNA, VcCascade and transposition machinery. The CRISPR array comprises two repeats (R) and a protospacer (PS) presented in violet. VcCascade consists of Cas6, Cas7 and Cas8 protein subunits presented in green. The transposition machinery comprises TniQ subunit and transposition proteins TnsA, TnsB and TnsC presented in orange. pDonor plasmid harboring a DNA cargo, i.e., *bktB*:*phaB* (2.2 kbp), *phaC*:*phaA* (3.3 kbp) or *cas9* (5.5 kbp), flanked by transposon right- (TR) and left- (TL) end sequences . The crRNA which is principally responsible for the PAM site recognition and R-loop formation, binds to the VcCascade. The TniQ homodimer is bound to the VcCascade to form the VcCascade-crRNA-TniQ complex. To insert the DNA cargo in the desired genomic site, the crRNA directs the VcCascade-crRNA-TniQ complex to the desired locus, TniQ communicates with TnsC, and TnsC mediates the transposition by employing TnsAB to excise the DNA from the its pDonor through double strand breaks, recognizing and binding to the TL and TR end sequences of the DNA cassette and integrates the DNA approximately 50 bp downstream of the targeted locus. 24

3.2 Schematic representation of engineered PHBV pathway in *E. coli* using glycerol as the carbon source. TCA cycle is presented in black. The reductive branch of the TCA cycle is show with purple arrows. Blue arrows represent the glyoxylate shunt in TCA cycle. The activated Sleeping beauty mutase (Sbm) pathway is presented in orange arrows. PHBV pathway and heterologous enzymes from *C. necator* are shown in green. Metabolite abbreviations: PHBV, Poly(3-hydroxybutyrate-co-3-hydroxyvalerate); (R)-3-HB-CoA, (R)-3- hydroxybutyryl-CoA; (R)-3-HV-CoA, (R)-3-hydroxyvaleryl-CoA; DHAP, dihydroxyacetone phosphate; PEP, phosphoenolpyruvate. Protein abbreviations: AceA, isocitrate lyase; AceB, malate syn- thase A; AceK, isocitrate dehydrogenase kinase/phosphatase; IclR, AceBAK operon repressor; LdhA, lactate dehydrogenase A; PckA, phosphoenolpyruvate carboxykinase; PPC, phosphoenolpyruvate carboxylase; PhaA, acetoacetyl-CoA thiolase; PhaB, acetoacetyl-CoA reductase; PhaC, PHA synthase; Sbm, methylmalonyl-CoA mutase; SdhA, succinate dehydrogenase subunit. 27

3.3 (**a**) Colony PCR screening of *cas9* and *bktB*:*phaB* integrations via λ_{Red} recombineering (column 2-3). To screen for *cas9* integration , primers P069/P070 amplified an 858 bp fragment (column 2). To screen for *bktB*:*phaB* integration, primers P071/P072 amplified a 930 bp fragment (column 3). (**b**) Integration efficiency based on colony PCR (ratio of colonies containing the desired mutation to the total number of colonies screened for colony PCR) and colony count (CFU/mL) of each transformation employing λ_{Red} recombineering system to integrate small- and large-size DNA cargos. (**c**) pgRNA plasmid eviction efficiencies confirmed by spectinomycin sensitivity (ratio of colonies with successful plasmid eviction to total number of colonies screened for antibiotic sensitivity). All values are reported as means ± SD (n = 3). . . 34

3.4 (**a**) Colony PCR screening of *cas9*, *bktB*:*phaB* and *phaC*:*phaA* integrations via Tn6677 transposons (column 2-4). To screen for *cas9* integration, primers P070/P073 amplified a 484 bp fragment (column 2). To screen for *bktB*:*phaB* integration, primers P072/P074 amplified a 641 bp fragment (column 3). To screen for *phaC*:*phaA* integration, primers P075/P076 amplified a 704 bp fragment (column 4). (**b**) Integration efficiency (ratio of colonies containing the desired mutation to the total number of colonies screened for colony PCR) and colony count (CFU/mL) of each transformation employing transposon-associated CRISPR-Cas system to integrate small-, medium- and large-size DNA cargos. (**c**) pDonor and pEffector plasmid eviction efficiencies (ratio of colonies with successful plasmid eviction to total number of colonies screened for antibiotic sensitivity). All values are reported as means ± SD (n = 3). . . 36

3.5 (**a**) Colony PCR screening of *iclR* and *sdhA* knockouts via λ_{Red} recombineering (column 2-3). To screen for *iclR* knockout, primers P077/P078 amplified a 790 bp fragment and successful recombination of the editing template generated products of 490 bp and 300 bp upon AseI digestion (columns 2). To screen for *sdhA* knockout, primers P079/P080 amplified a 1065 bp fragment and successful recombination of the editing template generated products of 779 bp and 286 bp upon AseI digestion (column 3). (**b**) Gene knockout efficiency based on colony PCR (ratio of colonies containing the desired mutation to the total number of colonies screened for colony PCR) and colony count (CFU/mL) of each transformation employing λ_{Red} recombineering system to knockout *iclR* and *sdhA* genes. (**c**) pgRNA plasmid eviction efficiencies confirmed by chloramphenicol sensitivity (ratio of colonies with successful plasmid eviction to total number of colonies screened for antibiotic sensitivity). All values are reported as means ± SD (n = 3). 37

3.6 PHBV biosynthesis of the metabolically manipulated strains (ST004, ST005, ST006) (**a**) presenting cell optical density (OD600) and dried cell weight (g/L) (**b**) biopolymer and 3-HV content represented by wt% and mol%, respectively and (**c**) Metabolite production (g/L). All values are reported as means ± SD (n = 2). 38

4.1 The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. Shemin pathway (in violet); Porphyrin biosynthesis (in grey). Pigments (in red); 5-ALA, 5-aminolevulinic acid; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; HMB, Hydroxymethylbilane; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HemY, coproporphyrinogen III oxidase; PBG, porphobilinogen; PCP-I, pentacarboxyl porphyrin I; PCPG-I, pentacarboxyl porphyrinogen I; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase. 44

4.2 Bioreactor cultivation of the control strain CPC-SbmΔ*iclR*Δ*sdhA*. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate production, (**b**) 5-ALA and PBG biosynthesis, and (**c**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). 51
4.3 Bioreactor cultivation of DMA for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). . . . 52
4.4 Bioreactor cultivation of DMB for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). . . . 54
4.5 Bioreactor cultivation of DMC for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). 56
4.6 Bioreactor cultivation of DMC^{B} for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). . . . 58
4.7 Bioreactor cultivation of DMD for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). . . . 60
4.8 Bioreactor cultivation of SMD for porphyrin biosynthesis under AM-II condition. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). 62

4.9 Bioreactor cultivation of SMD for porphyrin biosynthesis under AM-III condition. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$). 63

5.1 The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. 5-ALA, 5-aminolevulinic acid; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; Fe-CP-III, coproheme III; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin/coproporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HemQ, coproheme decarboxylase; HemY, coproporphyrinogen III oxidase; HMB, Hydroxymethylbilane; PBG, porphobilinogen; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase. 69
5.2 Aerobic bioreactor cultivation of BA001 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$). 77
5.3 Aerobic bioreactor cultivation of BA002 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$). 79
5.4 Aerobic bioreactor cultivation of BA003 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$). 81

5.5 Aerobic bioreactor cultivation of BA004 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD (n = 2). 83
5.6 Aerobic bioreactor cultivation of BA005 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD (n = 2). 85
5.7 Aerobic bioreactor cultivation of BA006 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD (n = 2). 87

6.1 The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. 5-ALA, 5-aminolevulinic acid; AhbA, siroheme decarboxylase α subunit; AhbB, siroheme decarboxylase β subunit; AhbC, Fe-coproporphyrin III synthase; AhbD, SAM-dependent heme synthase; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; Fe-CP-I, Fe-coproporphyrin I; Fe-CP-III, coproheme; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin/coproporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HemQ, coproheme decarboxylase; HemY, coproporphyrinogen III oxidase; HMB, Hydroxymethylbilane; PBG, porphobilinogen; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase. 93
6.2 Two step heme production using in vivo porphyrin production and in vitro cell-free enzymatic conversion. 96

6.3 Aerobic bioreactor cultivation of FECP1, FECP2 and FECP3 for porphyrin biosynthesis. Time profiles of (**a**,**d** and **g**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**b**,**e** and **h**) 5-ALA and PBG biosynthesis, and (**c**,**f** and **i**) porphyrin biosynthesis. All values are reported as means±SD (n=2). 105
6.4 Aerobic bioreactor cultivation of HEM1, HEM2 and HEM3 for porphyrin biosynthesis. Time profiles of (**a**,**d** and **g**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**b**,**e** and **h**) 5-ALA and PBG biosynthesis, and (**c**,**f** and **i**) porphyrin biosynthesis. All values are reported as means±SD (n=2). 107
6.5 SDS-PAGE analysis of HemH and HemQ expression and purification. Lane 1: Protein ladder; Lane 2: HemH insoluble fraction; Lane 3: HemH soluble fraction; Lane 4: Purified HemH from *B. subtilis*; Lane 5: HemQ insoluble fraction; Lane 6: HemQ soluble fraction; Lane 7: Purified HemQ from *B. subtilis*; Lane 8: Purified HemQ from *S. aureus*. 108
6.6 In vitro Fe-CP-III synthesis from CP-III. (**a**) Visible color change from red to orange, indicating the conversion of CP-III to Fe-CP-III. Reaction before adding enzyme (left) and after addition of enzyme (right) (**b**) Evaluation of buffer conditions for Fe-CP-III production. (**c**) Evaluation of pH conditions for Fe-CP-III production. (**d**) Tuning Fe^{2+} and HemH enzyme concentration, and buffer types for Fe-CP-III production. 110
6.7 In vitro Heme synthesis from Fe-CP-III. (**a**) Tuning H_2O_2 and HemQ concentration, and buffer types types for heme production. (**b**) HPLC analysis of Fe-CP-III conversion to heme at different time points: 1-hour (top) and overnight (bottom). 112

7.1 The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. 5-ALA, 5-aminolevulinic acid; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HMB, Hydroxymethylbilane; PBG, porphobilinogen; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase. 118

7.2 Aerobic bioreactor cultivation of DMB for UP biosynthesis with ascorbic acid supplementation. Time profiles of (**a**–**d**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**e**–**h**) 5-ALA and PBG biosynthesis, and (**i**) UP biosynthesis at 144 h. All values are reported as means ± SD (n = 2). 127
7.3 Aerobic bioreactor cultivation of SOD2 and SOD3 for UP biosynthesis with ascorbic acid supplementation. Time profiles of (**a**–**c**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**d**–**f**) 5-ALA and PBG biosynthesis, and (**g**) UP biosynthesis at 144 h. All values are reported as means ± SD (n = 2). 130
7.4 Aerobic bioreactor cultivation of KAT1, KAT2, KAT3, KAT4, and KAT5 for UP biosynthesis. Time profiles of (**a**–**e**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**f**–**j**) 5-ALA and PBG biosynthesis, and (**k**) UP biosynthesis at 144 h. All values are reported as means ± SD (n = 2). 132

S4.1 Aerobic bioreactor cultivation of ST-AB and ST-ABC strains for UP biosynthesis. Time profiles of (**a** and **b**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**c** and **d**) 5-ALA and PBG biosynthesis, and (**e**) UP biosynthesis at 144 h. All values are reported as means±SD (n=2). . 189
S4.2 Aerobic bioreactor cultivation of BL21(DE3)-derived strains for UP biosynthesis. Time profiles of (**a-c**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**d-f**) 5-ALA and PBG biosynthesis, and (**g**) UP biosynthesis at 144 h. All values are reported as means±SD (n=2). 190
S4.3 Aerobic bioreactor cultivation of BLD-derived strains for porphyrin biosynthesis. Time profiles of (**a-d**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**e-h**) 5-ALA and PBG biosynthesis, (**i**) UP biosynthesis time profile for BLD-AB, and (**j**) porphyrin biosynthesis at 144 h. All values are reported as means±SD (n=2). 191

List of Tables

2.1 Comparative Analysis of Chemical Synthesis and Biomanufacturing 6
2.2 Bioproducts derived from Glycerol Metabolism. 19

3.1 Strains and plasmids used in this study. 28

4.1 Strains and plasmids used in this study. 47

5.1 Strains and plasmids used in this study. 71

6.1 Strains and plasmids used in this study. 98

7.1 Strains and plasmids used in this study. 123

S3.1 Oligomers used in Chapter 3 172
S3.2 Statistical analysis data for integration efficiency based on transposon-associated CRISPR-Cas system vs. DNA cargo size 175
S3.3 Statistical analysis data for volumetric CFU based on transposon-associated CRISPR-Cas system vs. DNA cargo size 175
S3.4 Statistical analysis data for pDonor eviction efficiency based on transposon-associated CRISPR-Cas system vs. DNA cargo size 176
S3.5 Statistical analysis data for comparing the efficiency of the integrating methods with increase in the DNA cargo size 176
S4.1 Oligomers used in Chapter 4 177
S4.2 Strains and plasmids used in Supplementary Figures 178
S5.1 Oligomers used in Chapter 5 179
S6.1 Oligomers used in Chapter 6 181
S7.1 Oligomers used in Chapter 7 183
S7.2 Statistical analysis data for the UP-III/UP-I ratio for 0.5 g/L ascorbic acid supplementation vs the control strain 185

S7.3 Statistical analysis data for the UP-III/UP-I ratio for 1 g/L ascorbic acid supplementation vs the control strain 185
S7.4 Statistical analysis data for the UP-III/UP-I ratio for 2.0 g/L ascorbic acid supplementation vs the 0.5 g/L ascorbic acid supplementation 186
S7.5 Statistical analysis data for the UP-III/UP-I ratio for SOD2 vs the control strain 186
S7.6 Statistical analysis data for the UP-III/UP-I ratio for SOD3 vs the control strain 187
S7.7 Statistical analysis data for the UP-III/UP-I ratio for SOD2 + 1 g/L ascorbic acid supplementation vs the control strain 187
S7.8 Statistical analysis data for the UP-III/UP-I ratio for KAT1 vs the control strain 188
S7.9 Statistical analysis data for the UP-III/UP-I ratio for KAT2 vs the control strain 188

List of abbreviations

List of abbreviations

Abbreviation	**Description**
5-ALA	5 Aminolevulinic Acid
Amp	Ampicillin
B. subtilis	*Bacillus subtilis*
bktB	Beta-Ketothiolase
Cat	Chloramphenicol
CHO	Chinese Hamster Ovary
CFU	Colony Forming Units
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cas	CRISPR associated proteins
DCW	Dry Cell Weight
DSB	Double Stranded Break
dsDNA	Double stranded DNA
E. coli	*Escherichia coli*
FRT	Flippase Recognition Target
GC	Gas Chromatography
gRNA	Guide RNA
hemA	5-Aminolevulinate Synthase
hemB	Porphobilinogen Synthase
hemC	Porphobilinogen Deaminase
hemD	Uroporphyrinogen III Synthase
hemE	Uroporphyrinogen Decarboxylase
hemF	Coproporphyrinogen III Oxidase
hemG	Protoporphyrinogen Oxidase
hemH	Protoporphyrin/Coproporphyrin Ferrochelatase
hemN	Oxygen Independent Coproporphyrinogen III Oxidase
hemQ	Coproheme Decarboxylase

hemY	Coproporphyrinogen Oxidase
HMB	Hydroxymethylbilane
HPLC	High Performance Liquid Chromatography
iclR	Regulator of the AceBAK Operon
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
katE	Catalase HPII
katG	Catalase Peroxidase HPI
kat	Catalase
Km	Kanamycin
LB	Luria Bertani (a type of growth medium)
ldhA	Lactate Dehydrogenase A
NHEJ	Non Homologous End Joining
OD600	Optical Density at 600 nm
ori	Origin of replication
PAM	Protospacer Adjacent Motif
PBG	Porphobilinogen
PCR	Polymerase Chain Reaction
pckA	Phosphoenolpyruvate Carboxykinase
PHA	Polyhydroxyalkanoates
PHBV	Poly(3-hydroxybutyrate-co-3-hydroxyvalerate)
phaA	Acetoacetyl-CoA Thiolase
phaB	Acetoacetyl-CoA Reductase
phaC	Polyhydroxyalkanoate Synthase
PP-IX	Protoporphyrin IX
PPD	Protoporphyrin Dependent
PPG-IX	Protoporphyrinogen IX
ppc	Phosphoenolpyruvate Carboxylase
RBS	Ribosome Binding Site
ROS	Reactive Oxygen Species
rpm	Revolutions per Minute
S. aureus	*Staphylococcus aureus*
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis
sdhA	Succinate Dehydrogenase Subunit A
SOC Medium	Super Optimal Broth with Catabolite Repression
sodA	Superoxide Dismutase A
sodB	Superoxide Dismutase B
sodC	Superoxide Dismutase C
Spc	Spectinomycin
Succinyl-CoA	Succinyl Coenzyme A
TCA Cycle	Tricarboxylic Acid Cycle

TIR	Terminal Inverted Repeat
Tn6677	Transposon from Vibrio cholerae
UP	Uroporphyrin
UP-I	Uroporphyrin I
UP-III	Uroporphyrin III
UPG-I	Uroporphyrinogen I
UPG-III	Uroporphyrinogen III
ZFN	Zinc Finger Nuclease

List of symbols

List of symbols

Symbol	Description
Δ	Gene deletion
$^\circ$	Degrees of temperature
$\pm$	Plus-minus (precision of approximation)
α	Alpha subunit
β	Beta subunit
λ	Lambda (indicating wavelength)
μ	Mu (indicating "micro")
R	Resistance marker

Chapter 1

Introduction

The growing demand for sustainable alternatives to petrochemical-based processes has resulted in significant advancements in bio-based production. Bio-based production offers a renewable, environmentally friendly, and efficient solution to meeting global demands for valuable chemical compounds. This approach contrasts with traditional petrochemical methods, which are often associated with high environmental costs, non-renewable resource depletion, and complex processing requirements [1]. By leveraging microbial cell factories, bio-based production achieves remarkable specificity in converting simple substrates into high-value products through biological pathways [2, 3]. Among microbial hosts, *Escherichia coli* has emerged as a leading platform for biomanufacturing due to its well-characterized genetics, rapid growth rate, and adaptability to diverse bioprocessing conditions [4, 5].

The versatility of *E. coli* has been further enhanced by breakthroughs in synthetic biology and metabolic engineering. These fields have provided tools to reprogram the metabolic networks of microorganisms [6], allowing the synthesis of complex biomolecules with high efficiency. Among the diverse classes of molecules produced via microbial fermentation, porphyrins and their derivatives stand out due to their critical roles in biological systems and their broad industrial applications. Porphyrins, including heme and chlorophyll, are essential cofactors in processes such as oxygen transport, electron transfer, and photosynthesis. Industrially, they find use in pharmaceuticals, renewable energy technologies, and diagnostics [7, 8, 9]. However, producing porphyrins at scale using microbial systems remains challenging due to the inherent complexity of their biosynthetic pathways, issues with pathway regulation, feedback inhibition, and the toxicity of certain intermediates [10, 11, 12].

Porphyrin biosynthesis in *E. coli* involves a series of enzymatic reactions that transform simple precursors such as glycine and succinyl-CoA into complex tetrapyrrole compounds

[13]. While the natural pathways for porphyrin biosynthesis are well understood, their engineering to achieve industrial-scale production requires addressing multiple bottlenecks. These include enhancing precursor availability, optimizing the expression of biosynthetic enzymes, overcoming feedback inhibition [10, 14, 12], and managing intracellular oxidative stress. Moreover, the accumulation of toxic intermediates, such as heme, poses significant challenges, as these compounds can disrupt cellular homeostasis and reduce microbial viability [12, 15]. Therefore, the development of innovative strategies to bypass these limitations is crucial to unlocking the potential of *E. coli* as a biofactory for porphyrin production.

This thesis focuses on addressing these challenges through a multidisciplinary approach integrating genome engineering, metabolic pathway optimization, and innovative bioprocessing strategies. The research aims to establish a scalable framework for porphyrin production in *E. coli* by systematically addressing the limitations of the biosynthetic pathways and exploring solutions to enhance product yield, purity, and scalability. We hypothesized that by developing a robust genome engineering toolkit, implementing advanced metabolic engineering strategies, and mitigating production bottlenecks such as oxidative stress and intracellular toxicity, we could enhance porphyrin biosynthesis in *E. coli* and improve overall production efficiency.

1.1 Research Objectives

The primary objectives of this research are described as follows:

1. To develop a genome engineering toolkit for *E. coli* that enables precise and efficient genomic modifications to support strain optimization for bio-based production.

2. To engineer and optimize biosynthetic pathways for porphyrin intermediates such as uroporphyrin and coproporphyrin by improving precursor availability, fine-tuning gene expression, and addressing bottlenecks in the metabolic network.

3. To establish a two-step production strategy for heme, combining in-vivo biosynthesis with in-vitro enzymatic conversion to overcome challenges related to intracellular toxicity and enhance scalability.

4. To explore and implement strategies to mitigate oxidative stress to improve the efficiency of type-III porphyrin biosynthesis.

1.2 Thesis Organization

This thesis is structured into eight chapters. Following this introduction, Chapter 2 presents a comprehensive literature review that delves into the theoretical and practical aspects of biomanufacturing, genome engineering, and porphyrin biosynthesis. It examines the potential of *E. coli* as a microbial cell factory and highlights existing challenges and opportunities in the field.

Chapter 3 details the development of a versatile genome engineering toolkit that combines CRISPR-Cas9 and transposon technologies. This chapter explains the design and functionality of the toolkit and demonstrates its application in creating genetically engineered *E. coli* strains. The system's efficiency in knockouts and knock-ins is also discussed.

In Chapter 4, the focus shifts to the engineering of *E. coli* for uroporphyrin production. This chapter describes the implementation of the Shemin/C4 pathway, strategies for enhancing precursor availability, and the optimization of biosynthetic pathways to achieve high titers of uroporphyrin with minimal byproduct formation.

Chapter 5 expands on the biosynthesis of coproporphyrin, detailing the development of dual-operon systems regulated by strong promoters. The chapter highlights the successful production of coproporphyrin in *E. coli*, emphasizing the importance of pathway regulation and metabolic flux optimization in achieving efficient biosynthesis.

Chapter 6 introduces a two-step strategy for coproheme and heme production. This chapter outlines the integration of in-vivo and in-vitro approaches to overcome the toxicity challenges associated with intracellular heme accumulation. The development of enzymatic conversion processes to improve product yields is discussed in detail.

In Chapter 7, the role of oxidative stress in limiting porphyrin biosynthesis is explored. This chapter presents strategies to mitigate reactive oxygen species through antioxidant supplementation and the overexpression of ROS-scavenging enzymes, leading to enhanced type-III porphyrin production and improved scalability.

Finally, Chapter 8 highlights the original contributions of this research, summarizing the key findings and their implications for industrial biotechnology. Recommendations for future research directions are provided, emphasizing the potential for further advancements in metabolic engineering and synthetic biology to overcome remaining challenges and expand the applications of bio-based production systems.

Chapter 2

Literature Review

2.1 Biomanufacturing

Biomanufacturing, the use of biological systems to produce commercially valuable products, has revolutionized industries ranging from pharmaceuticals to renewable energy. This innovative approach capitalizes on the natural capabilities of biological entities, such as microorganisms and enzymes, to produce high-value compounds. For instance, the production of insulin, an essential treatment for diabetes, is now predominantly achieved through genetically engineered *Escherichia coli* (*E. coli)* or yeast [16, 17]. Similarly, biomanufacturing has enabled the large-scale production of monoclonal antibodies used in cancer therapies [18], as well as biofuels derived from algae and other renewable resources [19]. These examples illustrate the versatility and transformative potential of biomanufacturing across diverse sectors.

Other bioproducts have also significantly benefited from biomanufacturing. Antibiotics such as penicillin and streptomycin, originally extracted from natural sources, are now produced in industrial quantities using optimized strains of microorganisms [20]. Organic acids like lactic acid, a precursor for biodegradable plastics, and citric acid, widely used in food and beverages, are manufactured using microbial fermentation [21]. Moreover, biopolymers such as polyhydroxyalkanoates (PHAs) are produced by bacteria as environmentally friendly alternatives to conventional plastics [22]. These innovations highlight the breadth of applications biomanufacturing supports, showcasing its pivotal role in advancing sustainable industrial practices.

Unlike traditional chemical synthesis, which often relies on harsh reaction conditions and non-renewable resources, biomanufacturing leverages the precision and efficiency of

biological systems to catalyze complex reactions. This approach has gained significant attention due to its potential to address global challenges in sustainability, cost-effectiveness, and scalability [2, 3].

2.1.1 Comparison of Biological Methods with Chemical Synthesis

While biomanufacturing has proven to be a transformative approach, it is essential to understand how it compares with traditional chemical synthesis methods. Chemical synthesis has been a mainstay of industrial production, offering well-established methods for generating a wide range of compounds [23]. However, this approach faces significant limitations, particularly when synthesizing complex molecules. For example, chemical synthesis often requires high temperatures, toxic reagents, and multi-step processes with low overall yields [1]. Such requirements make chemical synthesis both energy-intensive and environmentally taxing, with the added burden of generating hazardous byproducts.

Biological methods, by contrast, offer a fundamentally different approach. Enzymes and microorganisms operate under mild conditions, such as ambient temperatures and pressures, which significantly reduce energy consumption [24]. Furthermore, biological systems utilize renewable feedstocks, such as sugars or agricultural byproducts, thereby aligning with principles of green chemistry [25]. A key advantage of biomanufacturing is its inherent specificity. Enzymatic processes are finely tuned to catalyze precise reactions, reducing the formation of undesired byproducts and simplifying downstream purification processes. These characteristics make biomanufacturing an attractive alternative for sustainable and efficient production [26]. A detailed comparison is presented in Table 2.1.

Table 2.1: Comparative Analysis of Chemical Synthesis and Biomanufacturing

Aspect	**Chemical Synthesis**	**Biomanufacturing**	**Ref.**
Raw Materials	Petroleum-based, non-renewable	Renewable feedstocks (e.g., sugars, agricultural byproducts)	[27, 28]
Reaction Conditions	High temperature and pressure; often harsh conditions	Mild (ambient temperature and pressure)	[29]
Energy Efficiency	Low due to energy-intensive processes	High due to efficient biological catalysts	[24, 30]
Environmental Impact	Generates hazardous byproducts; high carbon footprint	Minimal waste; aligns with green chemistry principles	[31, 32]
Scalability	High scalability for bulk chemicals; complex for intricate molecules	Highly scalable with advancements in fermentation and bioreactors	[24, 23]
Specificity of Products	Moderate; lacks fine-tuned specificity	High; enzymes and cells offer precision and selectivity	[33]
Production Costs	High for complex molecules; moderate for bulk chemicals	Lower for complex and high-value products; competitive for bulk products	[34]
Byproduct Management	Requires significant processing to handle toxic byproducts	Environmentally friendly; most byproducts are biodegradable	[35]

2.1.2 Genetic and Metabolic Engineering in Biomanufacturing

The ability to manipulate the genetic makeup of microorganisms has been a cornerstone of advancements in biomanufacturing. Genetic engineering allows the introduction of new pathways or the optimization of existing ones to increase the yield and efficiency of bioproduct synthesis [36]. For example, the overexpression of pathway-specific enzymes has been widely employed to enhance metabolic flux toward desired products. Additionally, genome editing tools, such as clustered regularly interspaced short palindromic repeats (CRISPR) coupled with Cas9, have enabled precise modifications, reducing unwanted byproducts and improving strain robustness [6, 18].

Metabolic engineering integrates genetic manipulation with a systems-level approach to redesign cellular metabolism. This involves optimizing carbon flux, balancing cofac-

tor availability, and alleviating bottlenecks within biosynthetic pathways [37, 33]. In the context of porphyrins metabolic engineering strategies include the amplification of rate-limiting enzymes, redirection of precursor pathways, and engineering of transport mechanisms to facilitate product secretion [38, 39, 40].

Moreover, synthetic biology has further revolutionized the field by introducing modular design principles. This allows the creation of synthetic gene circuits and the incorporation of non-natural pathways into microbial systems [41]. Such approaches have been instrumental in achieving high-level production of complex molecules like porphyrins.

In addition to strain engineering, advancements in omics technologies, such as transcriptomics, proteomics, and metabolomics, have provided deeper insights into cellular physiology. These tools enable the identification of potential targets for genetic manipulation and the evaluation of engineered strains under various conditions. As a result, researchers can iteratively improve microbial platforms for biomanufacturing applications [42].

2.2 Advantages of *E. coli* in Biomanufacturing

2.2.1 Historical Context and Model Organism Status

Since its first characterization in the late 19th century, *E. coli* has been a cornerstone of molecular biology and biotechnology. Its relatively simple genome, rapid growth rate, and ease of cultivation in laboratory settings have established *E. coli* as a model organism for biological research. The groundbreaking discovery of DNA replication mechanisms, gene expression regulation, and recombinant DNA technology was largely achieved using *E. coli* as the experimental system. Consequently, its utility in biomanufacturing has grown alongside advancements in molecular biology techniques [43, 4, 5].

One of the earliest commercial applications of *E. coli* was the production of human insulin, achieved through recombinant DNA technology in the 1980s [44]. This success demonstrated the feasibility of using genetically modified microorganisms to produce complex therapeutic proteins at scale. Since then, *E. coli* has been extensively utilized in the production of enzymes, biofuels, and small molecules, showcasing its adaptability and economic advantages [45].

2.2.2 Comparison with Alternative Hosts

While *E. coli* is a leading workhorse in biomanufacturing, other organisms such as *Bacillus subtilis*, yeast, and mammalian cells also play significant roles in industrial biotechnology. Each system offers unique strengths and limitations:

- ***Bacillus subtilis***: Known for its ability to secrete proteins directly into the growth medium, *B. subtilis* simplifies downstream processing for protein-based products [46]. However, it has less genetic tractability compared to *E. coli* and is less efficient for the production of certain metabolites [47].
- **Yeast**: *Saccharomyces cerevisiae* is a robust eukaryotic system capable of performing post-translational modifications essential for complex protein production [48]. However, its slower growth rate and the need for more complex media increase production costs [49].
- **Mammalian Cells**: Mammalian systems such as Chinese Hamster Ovary (CHO) cells are indispensable for producing biopharmaceuticals requiring human-like glycosylation patterns [50]. Yet, their high cost and stringent cultivation requirements limit their use for bulk production of simpler molecules [51].

Compared to these alternatives, *E. coli* offers unparalleled growth rates, scalability, and ease of genetic manipulation. However, its inability to perform certain post-translational modifications and the accumulation of endotoxins in its cell wall are notable limitations for applications requiring therapeutic-grade products [52].

2.3 Genetic and Metabolic Engineering in *E. coli*

2.3.1 Genome Engineering Methods in *E. coli*

Biomanufacturing using *E. coli* relies heavily on robust genome editing techniques to effectively manipulate the bacteria toward efficient biosynthesis of desired products. Precise genetic modifications allow the optimization of metabolic pathways, enhance production yields, and reduce unwanted byproducts. Over the years, several genome engineering methods have been developed for *E. coli*, ranging from foundational techniques to cutting-edge advancements.

Various methods have been employed for genome engineering in *E. coli*:

- **Recombineering Method**: Recombineering relies on homologous recombination to integrate or delete genetic elements using flanking homologous DNA sequences. Although highly specific, traditional homologous recombination is time-consuming and often requires extensive screening to identify successful edits [53]. It has, however, served as a cornerstone of early genetic engineering in *E. coli*. The efficiency of this method has been significantly enhanced by the λ Red recombinase system, which utilizes phage-derived enzymes to catalyze recombination at specific loci . This system accelerates the process compared to traditional homologous recombination and has been widely applied for genetic modifications, such as deletions and insertions. Despite its advantages, λ Red recombinase still requires careful design of homologous regions and is limited by the size of the DNA fragment being integrated [54].

- **Chemical Mutagenesis:** Historically, chemical mutagenesis has been employed to introduce random mutations into the *E. coli* genome. This technique involves exposing cells to UV radiation or chemical mutagens to induce mutations, followed by phenotype screening to identify beneficial traits. While simple and equipment-free, it lacks precision and requires extensive screening [55].

- **Zinc-Finger Nucleases (ZFNs) and Transcription Activator-Like Effectors (TALEs):** These protein-based genome editing tools enable targeted modifications by binding specific DNA sequences and introducing double-strand breaks. Although they provide more precision compared to earlier methods, their design complexity and cost have limited widespread adoption [56].

- **Base Editing and Prime Editing:** These methods allow precise, single-base changes without introducing double-stranded breaks. While base editing changes single nucleotides, prime editing enables small insertions or deletions, making these approaches particularly useful for fine-tuning genetic modifications [57, 58].

- **Genetic/Genome Recombination and Multiplex Automated Genome Engineering (MAGE):** Genetic recombination has been a cornerstone of genetic engineering. MAGE builds upon this concept by enabling the simultaneous introduction of multiple mutations, offering unparalleled flexibility for rapid strain optimization [59].

- **Phage Transduction and Bacterial Conjugation:** Phage-based systems like P1 transduction use bacteriophages to transfer genetic material between *E. coli* strains. Similarly, bacterial conjugation enables DNA transfer through direct cell-to-cell contact. Both techniques have been critical in facilitating horizontal gene transfer and genetic diversity [60].

- **CRISPR-Cas Systems:** CRISPR-based systems have revolutionized genome editing by enabling rapid and precise genetic modifications. These systems use RNA-guided endonucleases to target specific DNA sequences, allowing for highly efficient knockouts, knock-ins, and regulatory alterations. The versatility and accuracy of CRISPR-Cas systems make them valuable for metabolic engineering and strain optimization [61].

- **Transposon-Based Genome Engineering:** Transposon-based approaches facilitate stable genetic integrations by leveraging mobile genetic elements that can insert into specific or random genomic locations. These systems are particularly useful for introducing biosynthetic pathways, random mutagenesis, and operon integration, providing robust solutions for metabolic engineering and functional genomics [62].

2.3.2 CRISPR-Cas9: Mechanism and Applications

CRISPR-Cas9 is a revolutionary genome editing tool that has transformed genetic engineering due to its precision, simplicity, and adaptability [63, 61]. The system relies on two main components: a guide RNA (gRNA) and the Cas9 nuclease. The gRNA directs the Cas9 protein to specific genomic locations through complementarity to the target DNA sequence [64]. The gRNA consists of two parts: a CRISPR-derived RNA (crRNA) that recognizes the target sequence and a trans-activating crRNA (tracrRNA) that binds to Cas9, forming a functional RNA-protein complex [65]. Upon introduction into the cell, the Cas9-gRNA complex scans the genome for a protospacer adjacent motif (PAM), a short conserved sequence (typically NGG in *E. coli*) that is critical for initiating DNA binding (Figure 2.1) [66].

Once the PAM is identified, the Cas9 protein unwinds the adjacent DNA, allowing the gRNA to pair with the complementary strand. If the gRNA matches the target sequence with high fidelity, Cas9 induces a double-stranded break (DSB) at a precise location three base pairs upstream of the PAM site [67, 68]. This DSB activates the cell's DNA repair machinery, which resolves the break through one of two pathways. Homologous recombination (HR) utilizes a donor DNA template to introduce specific changes, such as gene insertions or point mutations, enabling precise editing. Alternatively, non-homologous end joining (NHEJ) repairs the break without a template, often resulting in small insertions or deletions that disrupt gene function and create knockouts [69].

The high efficiency of the Cas9-mediated DSB is due to its molecular architecture. The nuclease consists of two distinct domains: RuvC and HNH. The RuvC domain cleaves the non-complementary strand of DNA, while the HNH domain cleaves the complementary

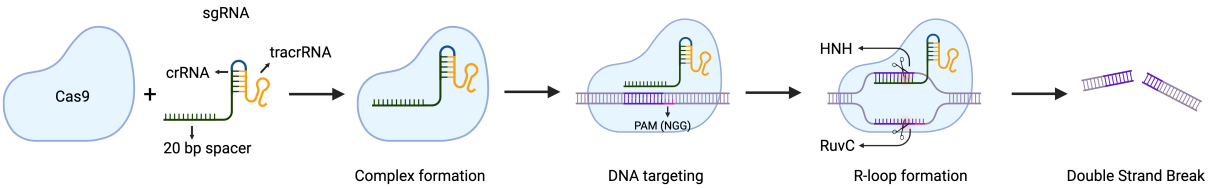


Figure 2.1: DNA targeting and cleavage by Cas9-gRNA complex

strand [66]. Together, these domains ensure a clean and accurate cut, making CRISPR-Cas9 a powerful tool for targeted genome modifications [70]. Advanced variations of Cas9, such as nickase Cas9 (nCas9) and dead Cas9 (dCas9), have further expanded its applications. nCas9 introduces single-strand breaks instead of DSBs, reducing the likelihood of off-target effects. dCas9 lacks nuclease activity and is used for transcriptional regulation by recruiting effector proteins to specific genomic loci [71].

CRISPR-Cas9 offers several advantages over traditional genome editing methods. Its precision is unmatched, as the high specificity of gRNA-Cas9 interactions minimizes off-target effects when designed carefully [70]. Additionally, its versatility allows for a wide range of applications, including gene knockouts, knock-ins, and transcriptional regulation, making it an indispensable tool in genetic engineering [72].

However, CRISPR-Cas9 is not without limitations. The requirement for a PAM sequence near the target site restricts potential editing locations, posing challenges for modifying certain genomic regions. Imperfect gRNA-DNA binding can result in off-target effects, leading to unintended mutations that may impact cellular function [69, 68]. To mitigate these issues, researchers have developed engineered Cas9 variants with relaxed PAM requirements or enhanced specificity, as well as transient delivery systems to minimize metabolic stress [70]. Despite these challenges, CRISPR-Cas9 remains a cornerstone of modern genetic engineering, continually evolving to address its limitations and expand its capabilities [72].

2.3.3 Transposons: Mechanism and Applications

Transposons, often referred to as "jumping genes," are DNA sequences that can move from one genomic location to another. This mobility is facilitated by transposase enzymes, which recognize specific sequences at the ends of the transposon and catalyze its excision and integration [62]. The mechanism of transposition involves several key steps. First, the transposase enzyme binds to the terminal inverted repeats (TIRs) flanking the transposon. These sequences are essential for the recognition and excision of the transposon from its

original genomic location. Once excised, the transposase forms a nucleoprotein complex which facilitates the search for a target site in the genome. The selection of the target site can be random or semi-specific, depending on the type of transposon used. Upon locating a suitable target, the transposon integrates into the genome by cleaving the target DNA and ligating its ends to the transposon ends, completing the insertion process [73, 74, 75].

In *E. coli*, transposons have been extensively utilized for genetic manipulation and metabolic engineering. One common application is the creation of random mutagenesis libraries, which allow researchers to identify genes involved in specific phenotypes or metabolic pathways [76]. Additionally, transposons have been employed to introduce entire biosynthetic pathways into the genome, enabling the production of complex molecules such as antibiotics and biofuels [75].

Despite their utility, transposons are not without limitations. The randomness of target site selection can lead to unintended disruptions of essential genes, necessitating careful screening to identify desired mutants. Furthermore, the stability of transposon insertions can vary, with some insertions being prone to excision or rearrangement. To address these challenges, researchers have developed engineered transposons with improved specificity and stability [77].

Transposons also offer unique advantages over other genome editing tools. Unlike CRISPR-Cas9, which relies on precise DNA targeting and repair mechanisms, transposons operate independently of host repair pathways [75], making them particularly useful in strains with compromised DNA repair systems. Additionally, their ability to integrate large DNA fragments enables the transfer of entire operons or metabolic pathways, a feature that is challenging to achieve with other editing technologies .

Recent advancements in transposon technology have further expanded their applications. Hybrid systems combining transposons with site-specific recombinases or CRISPR-Cas components have been developed to enhance targeting precision [78] making them a versatile tool in the field of synthetic biology and metabolic engineering.

2.4 Overview of Porphyrin Biosynthesis

2.4.1 Porphyrins

Porphyrins are a class of tetrapyrrolic compounds with essential roles in biological and industrial applications. Their derivatives, such as heme and chlorophyll, serve as vital cofactors in biological processes, including oxygen transport, electron transfer, and photosynthesis. In the biomedical field, heme derivatives are used in photodynamic therapy,

where they serve as photosensitizers to target cancer cells [7, 8]. Additionally, porphyrins play a central role in diagnostic imaging and as cofactors in various pharmaceutical compounds [9]. Industrial applications are equally diverse; for instance, porphyrins are used as catalysts in chemical synthesis, while heme is essential to the development of artificial blood substitutes and diagnostic kits for detecting blood-related disorders [79].

In renewable energy, porphyrin-based compounds have emerged as pivotal components in dye-sensitized solar cells, offering a sustainable solution for harnessing solar energy [80].

As global demand for porphyrins continues to grow, driven by advancements in healthcare, energy, and biotechnology, the need for scalable and cost-effective production methods has become increasingly evident. The global heme market alone is projected to reach approximately $530 million by 2026, with prices ranging from $482 to $548 per kilogram over the past six years, underscoring the economic value and commercial potential of porphyrin-based compounds [81, 82]. However, the structural complexity of porphyrins presents significant challenges for traditional synthesis. Harsh chemical conditions often degrade these molecules, further complicating their production [83]. Biomanufacturing overcomes these obstacles by leveraging microbial systems and enzymatic pathways to produce these compounds efficiently and at scale, preserving their functional integrity [84]. Tailoring fermentation conditions, such as nutrient composition, pH, and oxygen availability, can further maximize efficiency and minimize resource consumption [85]. The choice of microbial hosts is another critical factor; for instance, *E. coli* has become a preferred organism due to its rapid growth, well-characterized genetics, and ease of genetic manipulation [86]. Collectively, these strategies enable the development of robust bioprocesses capable of meeting the growing demand for porphyrins while addressing sustainability and economic considerations.

2.4.2 Porphyrin Biosynthesis pathway

The biosynthesis of porphyrins is a highly conserved process, beginning with the production of 5-aminolevulinic acid (5-ALA), a universal precursor. 5-ALA synthesis occurs via two distinct pathways: the C4 pathway and the C5 pathway [87, 13]. In the C5 pathway, used by archaea and most bacteria, glutamate is first ligated to glutamyl-tRNA by glutamyl-tRNA synthetase. This intermediate is then reduced to glutamate-1-semialdehyde by glutamyl-tRNA reductase, which is subsequently converted to 5-ALA by glutamate-1-semialdehyde aminotransferase. The C4 pathway, predominant in eukaryotes and some bacteria, involves the condensation of glycine and succinyl-CoA, catalyzed by 5-ALA synthase, to form 5-ALA (Figure 2.2). This step is tightly regulated as it often serves as the rate-limiting point of porphyrin biosynthesis [84].

Once 5-ALA is synthesized, it enters one of three downstream pathways: the protoporphyrin-dependent (PPD) pathway, the coproporphyrin-dependent (CPD) pathway, or the siroheme-dependent (SHD) pathway. These pathways diverge depending on the organism and environmental conditions [10, 39, 88, 89].

The PPD pathway is commonly found in gram-negative bacteria and eukaryotes. It progresses through a series of enzymatic steps. First, two molecules of 5-ALA condense to form porphobilinogen (PBG), catalyzed by 5-ALA dehydratase (HemB). Four PBG molecules are then polymerized into hydroxymethylbilane by porphobilinogen deaminase (HemC). This linear tetrapyrrole is cyclized into uroporphyrinogen III by uroporphyrinogen III synthase (HemD). Uroporphyrinogen III is decarboxylated by uroporphyrinogen decarboxylase (HemE) to form coproporphyrinogen III. Subsequently, coproporphyrinogen III is oxidatively decarboxylated by coproporphyrinogen oxidase (HemF/HemN) producing protoporphyrinogen IX.

Protoporphyrinogen IX is then oxidized to protoporphyrin IX by protoporphyrinogen oxidase (HemG). The final step involves the insertion of a ferrous ion (Fe^{2+}) into protoporphyrin IX, catalyzed by ferrochelatase (HemH), resulting in the production of heme [90, 91, 92, 87].

The CPD pathway, prevalent in gram-positive bacteria, also begins with the synthesis of 5-ALA and its conversion to PBG, hydroxymethylbilane, and uroporphyrinogen III. Following this, uroporphyrinogen III is decarboxylated to coproporphyrinogen III by uroporphyrinogen decarboxylase. However, in contrast to the PPD pathway, Coproporphyrinogen III is then oxidized to coproporphyrin III by coproporphyrinogen oxidase (HemY). Subsequently, coproporphyrin III undergoes iron (Fe^{2+}) insertion by ferrochelatase (HemH) to form Fe-coproporphyrin III (Coproheme). Finally, Fe-coproporphyrin III is converted to heme by HemQ, completing the pathway [93, 94].

In *E. coli*, the multifunctional enzyme siroheme synthase (CysG) catalyzes the conversion of uroporphyrinogen III to siroheme.. This process involves the methylation of uroporphyrinogen III to form precorrin-2, followed by its oxidation to sirohydrochlorin, and the subsequent chelation of Fe^{2+} to produce siroheme. However, in some archaea and sulfate-reducing or methanogenic bacteria, siroheme serves as an intermediate in the SHD pathway for heme biosynthesis. In these organisms, siroheme is converted to heme through additional enzymatic steps. Siroheme decarboxylase (AhbA and AhbB) first decarboxylate the acetate side chains of siroheme and AhbC then removes two acetate side chains to produce Fe-coproporphyrin III (coproheme). Finally, AhbD catalyzes the decarboxylation of the propionate side chains of Fe-coproporphyrin III, producing heme [95, 96, 97, 81].

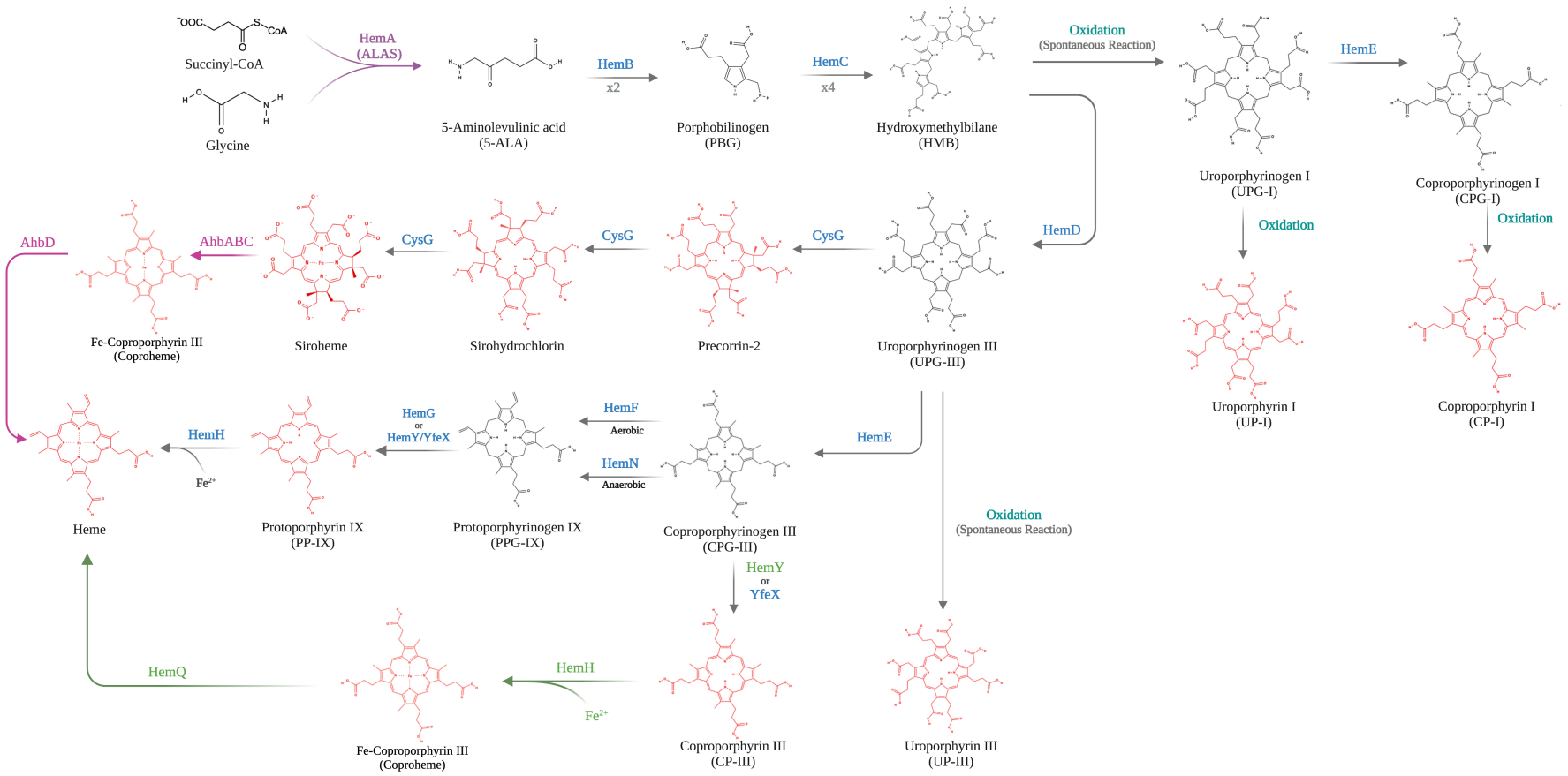


Figure 2.2: Overview of heme biosynthesis pathways. The purple arrow represents the Chemin/C4 pathway; the gray arrows represents the PPD pathway found in Gram-negative bacteria and eukaryotes; the pink arrows represents the SHD pathway from archaea and sulfate-reducing bacteria; and the green arrows represents the CPD pathway found in Gram-positive bacteria.

2.4.3 Regulation of Porphyrin Biosynthesis

The regulation of porphyrin biosynthesis in bacterial cells is a complex process influenced by various factors, including oxygen levels, iron availability, and post-translational modifications of enzymes involved in the biosynthetic pathways [10, 11, 98]. This intricate regulatory network ensures that porphyrin production aligns with cellular demands while minimizing the risks associated with the accumulation of toxic intermediates.

A key regulatory mechanism is feedback inhibition, particularly at the first and rate-limiting step of the pathway. 5-ALA synthase, responsible for the synthesis of 5-ALA, is tightly regulated by heme, the pathway's end product. Heme serves as a feedback inhibitor, ensuring the pathway is activated only when cellular heme levels are insufficient [12, 99]. This prevents excessive accumulation of porphyrin intermediates, such as 5-ALA, which are known to contribute to oxidative stress through the generation of reactive oxygen species (ROS) [15, 100].

Transcriptional regulation also plays a critical role. Genes encoding the biosynthetic enzymes are upregulated under conditions of heme demand, such as low oxygen or iron availability. This ensures that sufficient heme is produced to support essential cellular functions, particularly during stress or metabolic shifts [10, 101].

Post-translational modifications, including phosphorylation and acetylation, add another layer of control by dynamically altering enzyme activity, stability, or interactions [102]. Bottlenecks in the pathway highlight the need for balanced regulation to enhance overall flux. Experimental studies have shown that overexpression of rate-limiting enzymes like 5-ALA synthase and ferrochelatase can improve production yields [14, 103]. However, these approaches often require concurrent strategies to mitigate feedback inhibition and ROS accumulation for sustainable results.

By integrating these regulatory mechanisms, bacterial cells maintain a delicate balance in porphyrin biosynthesis, ensuring cellular integrity and meeting physiological needs. This understanding provides a framework for metabolic engineering aimed at optimizing porphyrin and heme production for industrial and therapeutic applications.

2.5 Reactive Oxygen Species (ROS) and Their Impact

2.5.1 Biochemical Basis of ROS Generation in Porphyrin Biosynthesis

ROS are chemically reactive molecules derived from molecular oxygen, including species such as superoxide anions ($O_2^{\bullet -}$), hydrogen peroxide (H_2O_2), and hydroxyl radicals ($OH^{\bullet}$) [104]. In biological systems, ROS play a dual role, acting as signaling molecules at low concentrations but causing cellular damage when present at high levels [105]. The biosynthesis of porphyrins inadvertently contributes to ROS generation, particularly during the oxidation and reduction reactions involved in heme synthesis [100].

Certain steps in porphyrin biosynthesis are highly oxygen-dependent and can lead to the generation of ROS as byproducts, contributing to oxidative stress within the cellular environment [94]. Additionally, the accumulation of porphyrin intermediates, such as protoporphyrin IX, can exacerbate oxidative stress resulting in ROS production, creating a feedback loop that impacts cellular metabolism and viability [106].

2.5.2 Cellular Impact of ROS

Excessive ROS levels can lead to oxidative damage of cellular macromolecules, including DNA, proteins, and lipids. DNA damage results from the formation of strand breaks and base modifications, impairing genetic integrity. Protein oxidation disrupts enzymatic functions, affecting metabolic pathways crucial for porphyrin biosynthesis and other cellular processes. Lipid peroxidation compromises membrane integrity, leading to increased permeability and ion leakage [107].

The accumulation of ROS during porphyrin biosynthesis also induces cellular stress responses, such as the upregulation of antioxidant defense systems. Elevated ROS levels can significantly reduce yield and productivity [108], necessitating targeted strategies to mitigate oxidative stress.

2.5.3 Strategies to Mitigate ROS Effects

To mitigate the detrimental effects of ROS, various genetic and process optimization strategies have been employed:

1. **Genetic Engineering of Antioxidant Systems**: Enhancing the expression of native or heterologous antioxidant enzymes is a common approach. Superoxide dismutase (SOD) converts superoxide anions to hydrogen peroxide, while catalase decomposes hydrogen peroxide into water and oxygen, reducing oxidative stress. Co-expression of these enzymes has been shown to improve microbial tolerance to ROS [109, 110].

2. **Optimization of Fermentation Conditions**: Limiting dissolved oxygen levels during fermentation can reduce ROS formation. Oxygen-limiting strategies, such as fed-batch cultivation or microaerophilic conditions, help maintain a balance between oxygen supply and cellular demand [109].

3. **Supplementation with Antioxidants**: Adding exogenous antioxidants, such as ascorbic acid, glutathione, or tocopherols, to the culture medium can scavenge ROS. These compounds protect cellular components from oxidative damage, improving cell viability and productivity [111].

4. **Engineering Redox Balance**: Modulating the expression of enzymes involved in NADH/NADPH generation, such as glucose-6-phosphate dehydrogenase, enhances redox homeostasis. Improved availability of reducing equivalents strengthens the cell's antioxidant defenses [112].

2.5.4 Challenges of ROS Management in Industrial Bioprocesses

Despite significant advancements, managing ROS in industrial-scale bioprocesses remains challenging. High cell density cultures, often employed to maximize yields, inherently generate elevated levels of ROS due to increased metabolic activity and oxygen consumption. This exacerbates oxidative stress and limits the scalability of engineered strains [113].

Furthermore, the heterogeneous nature of large-scale fermentations creates microenvironments with varying oxygen levels, complicating the consistent application of ROS mitigation strategies. These gradients can result in localized oxidative damage, reducing overall process efficiency [114]. Addressing these issues requires advanced bioreactor designs and real-time monitoring tools capable of assessing ROS levels at micro and macro scales [115].

2.6 Glycerol as an Ideal Feedstock

Glycerol has emerged as a highly advantageous feedstock for biomanufacturing processes due to its abundance, cost-effectiveness, and versatility. As a byproduct of biodiesel production, glycerol is readily available in large quantities, making it an economically viable option for industrial applications [116, 117]. Its chemical structure makes it a highly reduced carbon source, allowing microorganisms to efficiently convert it into a variety of value-added products with minimal metabolic energy expenditure [118].

One of the primary benefits of glycerol is its compatibility with diverse microbial systems. Many industrially relevant microorganisms, including *E. coli*, *Clostridium spp.*, and *Saccharomyces cerevisiae*, can metabolize glycerol through well-characterized pathways such as the glycerol kinase pathway or glycerol dehydrogenase pathway [119]. These pathways enable the production of biofuels, organic acids, polyols, and other biochemicals. For instance, *E. coli* has been engineered to utilize glycerol as a substrate for the production of ethanol, succinic acid, and polyhydroxyalkanoates (PHAs) [120, 121]. Furthermore, recent studies have demonstrated that glycerol can serve as a precursor for high-value compounds such as dihydroxyacetone, 1,3-propanediol, and biosurfactants, which have applications in cosmetics, pharmaceuticals, and biodegradable plastics [122]. Examples of bioproducts derived from glycerol metabolism, along with the microorganisms involved, are presented in Table 2.2.

Table 2.2: Bioproducts derived from Glycerol Metabolism.

Bioproduct	**Microorganism**	**Reference**
Ethanol	*E. coli*, *Saccharomyces cerevisiae*	[123, 124]
Succinic acid	*Actinobacillus succinogenes*, *E. coli*	[125, 126]
1,3-Propanediol	*Clostridium butyricum*, *Klebsiella pneumoniae*	[127, 128]
Polyhydroxyalkanoates	*Ralstonia eutropha*, *Pseudomonas putida*	[129, 130]
Biosurfactants	*Pseudomonas aeruginosa*	[131]
Citric acid	*Yarrowia lipolytica*	[132]
Hydrogen	*Enterobacter aerogenes*	[133]
Lactic acid	*Lactobacillus plantarum*, *E. coli*	[134, 135]
Propionic acid	*Propionibacterium acidipropionici*	[136]
Butanol	*Clostridium acetobutylicum*	[137]
Isoprenoids	*E. coli*	[138]

In addition to its metabolic efficiency, glycerol offers significant environmental advantages. Its utilization aligns with principles of green chemistry, as it reduces waste from

biodiesel manufacturing and minimizes reliance on non-renewable feedstocks [139]. Furthermore, glycerol-based bioprocesses often operate under mild conditions, reducing energy consumption and the generation of hazardous byproducts compared to traditional petrochemical processes [140]. By integrating glycerol as a feedstock, industries can achieve a circular economy model, transforming a biodiesel byproduct into a valuable resource.

However, the use of glycerol is not without challenges. Crude glycerol derived from biodiesel production contains impurities such as methanol, salts, and free fatty acids, which can inhibit microbial growth and fermentation efficiency. Pretreatment and purification steps are often required to enhance glycerol quality, adding to the overall production cost [141].

Another significant advantage of glycerol is its high reduction potential, which is particularly beneficial for the production of reduced metabolites such as fatty acids, bioethanol, and isoprenoids. Glycerol metabolism produces twice as many reducing equivalents, such as NADH, compared to glucose, thereby enhancing the efficiency of biosynthetic pathways. This property makes glycerol particularly suitable for microbial production of redox-intensive compounds [142, 118].

Continued research into optimizing microbial pathways and fermentation processes will further unlock the potential of glycerol as a feedstock. By addressing current limitations and leveraging its inherent advantages, glycerol could play a central role in the transition to sustainable and economically competitive biomanufacturing, supporting the production of a diverse array of bio-based chemicals, fuels, and materials.

Chapter 3

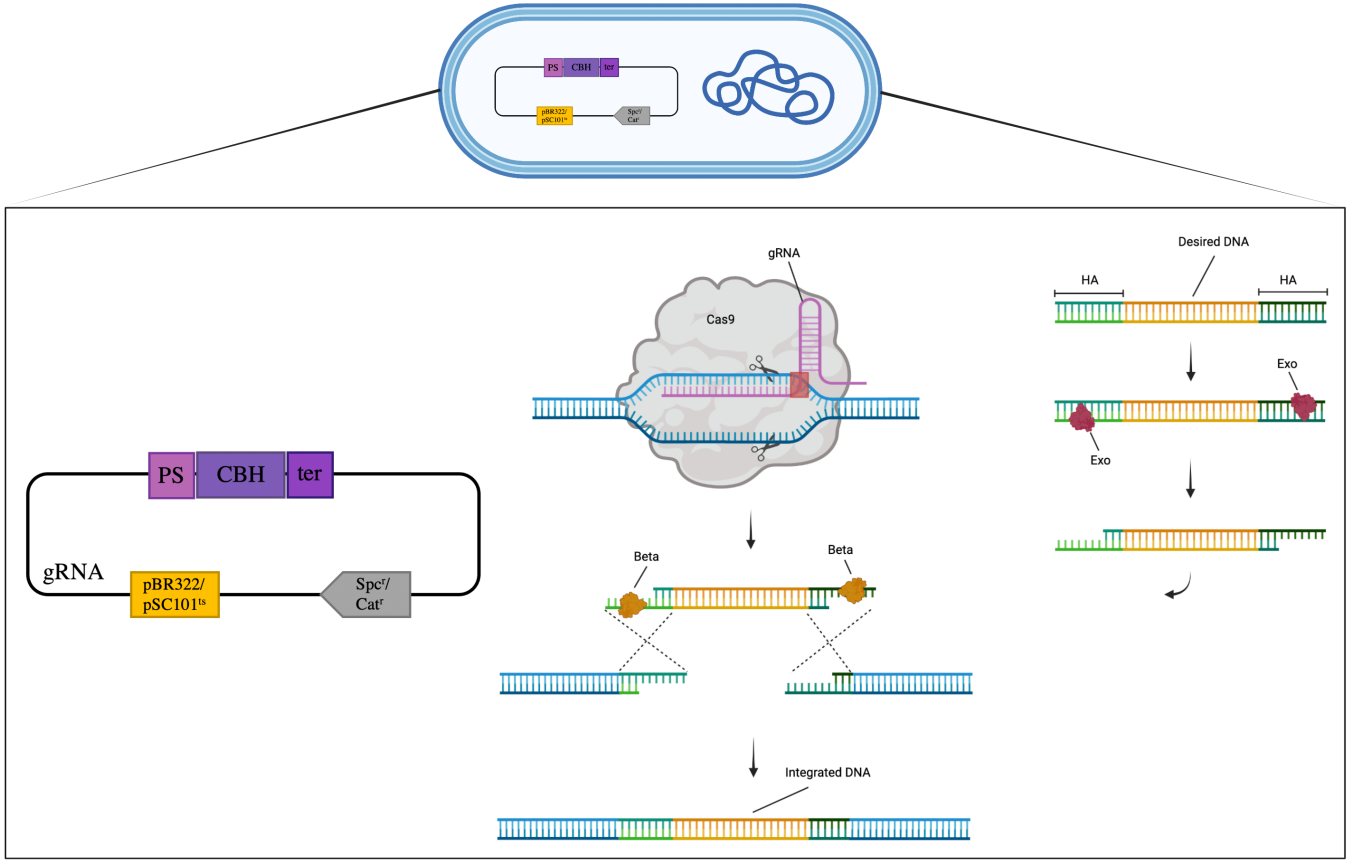
A toolkit for effective and successive genome engineering of *Escherichia coli*

3.1 Introduction

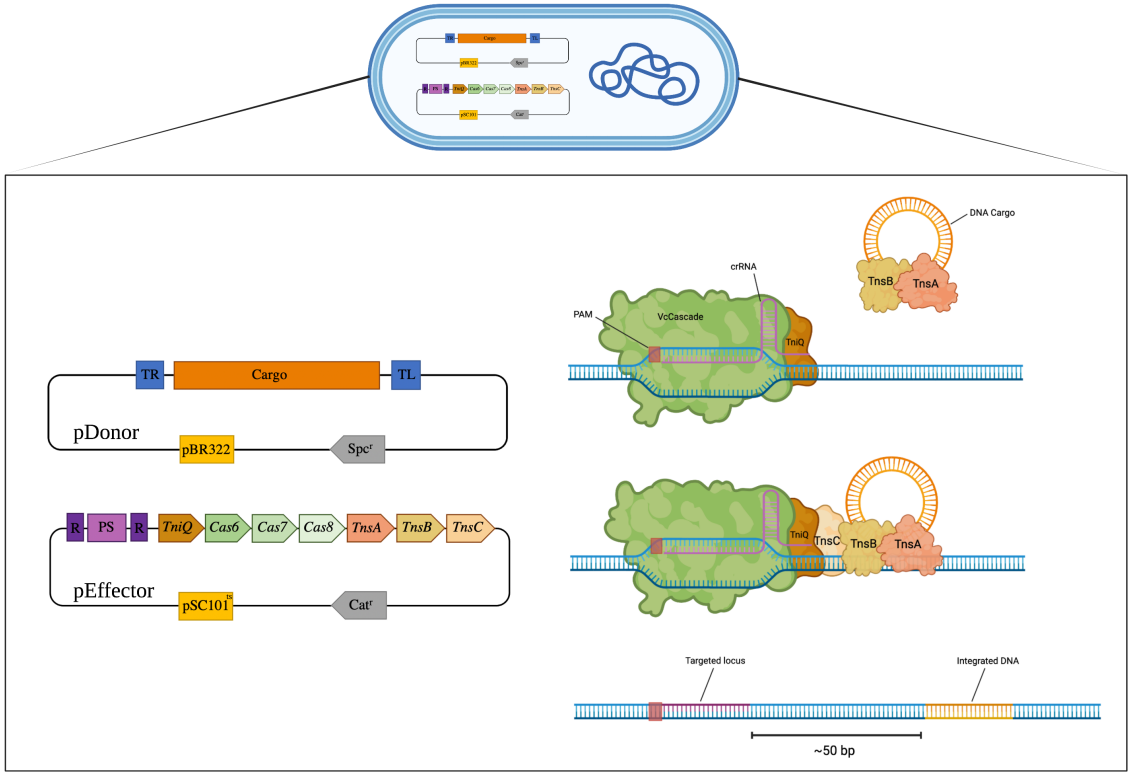
Over the past decades, bacterium *Escherichia coli* (*E. coli*) has been exploited as a workhorse for industrial applications [143], which often require the host strain to be genetically modified to develop proper phenotypes. Basically, genetic modification is conducted to change (i.e., increase, decrease or wipe out) the expression levels of key genes, either native or foreign ones. Traditionally, enhancing (or reducing) gene expression relies on the use of vector systems (e.g., plasmid or virus), whereas inactivating gene expression is conducted by targeting endogenous genes on the host genome for mutation (i.e., gene knockout). However, the presence of extrachromosomal plasmids containing foreign genes can adversely impact cell physiology [144, 145]. On the other hand, genomic integration of these foreign genes can enhance cell fitness and genetic stability [146], eliminating the necessity of using antibiotics to maintain plasmid stability [147]. While extrachromosomal plasmid systems appear to be relatively easy for implementation, targeting host genome for manipulation remains challenging and it is not even common until recently when novel biotechnologies for site-specific genome engineering become available.

Various genome engineering technologies were developed in early days based on (1) random mutagenesis via strain exposure to UV radiation or chemical mutagens followed by phenotype screening [148, 149], (2) zinc-finger nuclease (ZFN) [150] and transcription

activator-like effectors (TALEs) mutagenesis [151], (3) base editing [152] and prime editing [153], (4) genetic/genome recombination [154] and Multiplex automated genome engineering (MAGE) [59], (5) transposon mutagenesis [155], (6) phage transduction [156], and bacterial conjugation [157]. However, these traditional technologies are more limited to gene knockout and have various disadvantages, such as non-specific and non-reproducible mutations, hard-to-characterize mutations/mutants, labor-intensive and time-consuming protocols, etc. In early 21st century, technologies for site-specific genome engineering started to be developed, enabling targeted inactivation of endogenous genes in *E. coli*, and recombineering was one of them. Basically, recombineering relies on homologous recombination, which is mediated by phage recombinases such as RecET from Rac prophage or λ_{Red} proteins (i.e., Gam, Beta, Exo) [158], to integrate a double-stranded and single-stranded DNA fragment into the host genome [159]. In the λ_{Red} machinery, Gam prevents the degradation of the foreign DNA [160], Exo (lambda exonuclease) generates single-stranded 3' overhangs, and Beta mediates annealing of homology arms to the complementary DNA target in the genome [161] (Figure 3.1a). Recombineering has been extensively applied to inactivate endogenous genes and a comprehensive *E. coli* mutant library containing the single mutations to all non-essential genes, i.e., Keio collection, was constructed [162], significantly facilitating derivation of *E. coli* gene-knockout mutants. On the other hand, recombineering has major technological drawbacks, such as low recombination efficiencies, especially upon integrating large DNA cargos, and requiring multi-step processes [146, 163, 59], limiting its primary application to gene knockout.


PS
CBH
ter
gRNA
pBR322/
pSC101ts
Spcr/
Catr
Cas9
gRNA
Beta
Beta
Integrated DNA
Desired DNA
HA
HA
Exo
Exo

(a)



(b)

Figure 3.1: Schematic representation of the two-plasmid system for genomic integration via transposon-associated CRISPR-Cas system. (**a**) CRISPR-Cas9 coupled with λ_{Red} recombineering mechanism: The guide RNA (gRNA) comprised of a protospacer (PS), Cas9-binding hairpin (CBH) and transcriptional terminator (ter), directs the Cas9 to the desired locus by recognizing the PAM sequence and the Cas9 cleaves the DNA through double strand breaks. Exo (lambda exonuclease) generates single-stranded 3' overhangs and Beta proteins mediate annealing of the homology arms (HA) to the complementary DNA target in the genome. (**b**) Transposition mechanism: pEffector plasmid harboring the crRNA, VcCascade and transposition machinery. The CRISPR array comprises two repeats (R) and a protospacer (PS) presented in violet. VcCascade consists of Cas6, Cas7 and Cas8 protein subunits presented in green. The transposition machinery comprises TniQ subunit and transposition proteins TnsA, TnsB and TnsC presented in orange. pDonor plasmid harboring a DNA cargo, i.e., *bktB*:*phaB* (2.2 kbp), *phaC*:*phaA* (3.3 kbp) or *cas9* (5.5 kbp), flanked by transposon right- (TR) and left- (TL) end sequences . The crRNA which is principally responsible for the PAM site recognition and R-loop formation, binds to the VcCascade. The TniQ homodimer is bound to the VcCascade to form the VcCascade-crRNA-TniQ complex. To insert the DNA cargo in the desired genomic site, the crRNA directs the VcCascade-crRNA-TniQ complex to the desired locus, TniQ communicates with TnsC, and TnsC mediates the transposition by employing TnsAB to excise the DNA from the its pDonor through double strand breaks, recognizing and binding to the TL and TR end sequences of the DNA cassette and integrates the DNA approximately 50 bp downstream of the targeted locus.

More recently, site-specific genome engineering based on the clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins (Cas) was developed. The CRISPR-Cas system is a well-known machinery that reinforces immunity in bacteria and archaea by eradicating the invading foreign genetic elements [164, 165]. Basically, the CRISPR-Cas system employs sequence-specific protospacers that are expressed as CRISPR RNAs (crRNAs) to guide Cas nucleases to the DNA/RNA target for degradation through the recognition of the protospacer-adjacent motif (PAM) downstream of the targeted site [166, 61]. This natural defending mechanism in RNA-guided DNA cleavage has been repurposed as a powerful toolkit for genome engineering, particularly site-directed mutagenesis, in both prokaryotes and eukaryotes [167], successfully generating engineered cells for a wide range of applications [168] such as therapeutics [169] and engineered food crops [170].

Even with successful application to gene knockout and site-directed mutagenesis for genome editing, the CRISPR-Cas system has a limited editing efficiency upon inserting DNA cargos, particularly large ones, into the genome. The issue can be complemented by the new transposon-associated CRISPR-Cas system. Recent studies demonstrated the effective synergy between *Vibrio cholerae* Tn6677 transposon and type I-F CRISPR-Cas system for site-specific DNA transposition in *E. coli* [171, 172, 173]. Basically, *Vibrio cholerae* Cascade (VcCascade), encoded by *Vibrio cholerae* Tn6677, consists of Cas6, Cas7, and naturally fused Cas8-Cas5 (simply Cas8) protein subunits [174]. The hairpin on the 3' end of crRNA binds to Cas6, while the 5' end binds to Cas8 which is responsible for PAM site recognition and R-loop formation [175]. The transposition machinery comprises the TniQ subunit and transposition proteins of TnsA, TnsB, and TnsC. The TniQ homodimer is bound to the Cas6 protein of VcCascade to form the VcCascade-crRNA-TniQ complex [176]. To insert the DNA cargo, the crRNA directs the VcCascade-crRNA-TniQ complex to the targeted locus of the genome. TniQ communicates with TnsC to mediate the transposition by employing TnsAB to excise the DNA from its donor plasmid, recognizing and binding to the transposon right- (TR) and left- (TL) end sequences of the DNA cassette, and integrating it into the targeted locus [175, 177] (Figure 3.1b). Unlike most transposons that randomly insert the DNA cargo into the genome, the Tn6677 transposon associated with the CRISPR-Cas system precisely inserts the DNA cargo at a specified distance downstream of the targeted locus [178], i.e., approximately 50 base pairs downstream of the 3' end of the targeted site in either left to right or right to left orientation [171, 179]. In addition to the programmability, this system can insert kilobase-pair-sized DNA cargos with enhanced efficiency and accuracy, making it a promising molecular toolkit for large-scale and pathway-sized DNA integration into the genome [163, 180]. However, the transposon-associated CRISPR-Cas system is unsuitable to perform gene deletion or site-directed mutagenesis.

In this study, by integrating the three biomolecular mechanisms of CRISPR-Cas, Tn6677 transposon, and recombineering, we developed a programmable and efficient toolkit for extensive genome editing, including both site-specific inactivation of endogenous genes (gene knockout) and insertion of large DNA cargos (gene knock-in). The toolkit contains key engineered *E. coli* strains for convenient implementation of multiple rounds of genome editing and each editing round is based on one-step transformation with high editing efficiencies, enabling customized derivation of plasmid-free *E. coli* strains for industrial applications. As a demonstration, we used the developed toolkit to derive a plasmid-free engineered strain for the production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV) through metabolically implementing a heterologous PHBV-biosynthetic pathway (Figure 3.2) and then redirecting dissimilated carbon flux to the implemented PHBV-biosynthetic pathway.

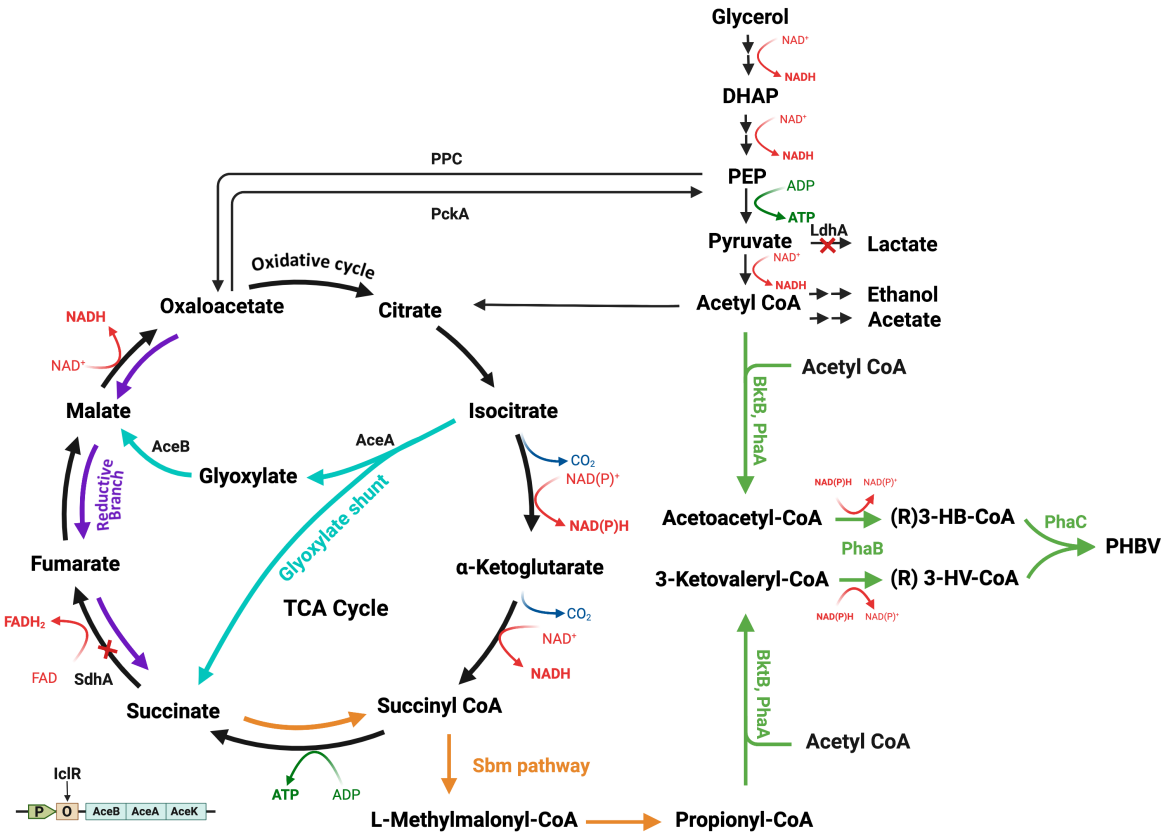


Figure 3.2: Schematic representation of engineered PHBV pathway in *E. coli* using glycerol as the carbon source. TCA cycle is presented in black. The reductive branch of the TCA cycle is show with purple arrows. Blue arrows represent the glyoxylate shunt in TCA cycle. The activated Sleeping beauty mutase (Sbm) pathway is presented in orange arrows. PHBV pathway and heterologous enzymes from *C. necator* are shown in green. Metabolite abbreviations: PHBV, Poly(3-hydroxybutyrate-co-3-hydroxyvalerate); (R)-3-HB-CoA, (R)-3- hydroxybutyryl-CoA; (R)-3-HV-CoA, (R)-3-hydroxyvaleryl-CoA; DHAP, dihydroxyacetone phosphate; PEP, phosphoenolpyruvate. Protein abbreviations: AceA, isocitrate lyase; AceB, malate syn- thase A; AceK, isocitrate dehydrogenase kinase/phosphatase; IclR, AceBAK operon repressor; LdhA, lactate dehydrogenase A; PckA, phosphoenolpyruvate carboxykinase; PPC, phosphoenolpyruvate carboxylase; PhaA, acetoacetyl-CoA thiolase; PhaB, acetoacetyl-CoA reductase; PhaC, PHA synthase; Sbm, methylmalonyl-CoA mutase; SdhA, succinate dehydrogenase subunit.

3.2 Materials and Methods

3.2.1 Bacterial strains, plasmids, editing template, and oligonucleotides

Bacterial strains and plasmids used in this study are listed in Table 3.1, while oligonucleotide sequences are presented in Table S3.1. Genomic DNA from bacterial cells was isolated using the Blood & Tissue DNA Isolation Kit (Qiagen, Hilden, Germany). *Taq* DNA polymerase was obtained from New England Biolabs (Ipswich, MA, USA). *E. coli* HI-Control 10G chemically competent cells (Lucigen, Middleton, WI, USA) were used as the host for plasmid cloning and propagation. The plasmids were extracted using Qiagen Miniprep kit according to the manufacturer's instructions and confirmed by DNA sequencing conducted by The Centre for Applied Genomics (TCAG) (Toronto, Ontario, Canada). All plasmids were constructed using Gibson assembly [181] except for the gRNAs which were constructed by ligation. All oligonucleotides were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA). *E. coli* strains were stored as glycerol stocks at -80°C.

Table 3.1: Strains and plasmids used in this study.

Name	**Description or Relevant Genotype**	**Source**
Host strains		
HI-Control 10G	*mcrA* Δ (*mrr-hsdRMS-mcrBC*) *endA1 recA1* φ80d*lac*ZΔM15 Δ*lacX74 araD139* Δ (*ara leu*)*7697* galU galK rpsL (Strr) *nupG* λ^- *tonA* Mini-F *lacIq1* (Gentr)	Lucigen
MG1655	K-12; F^- λ^- *rph-1*	Lab stock
CPC-Sbm	BWΔ*ldhA*, P*trc::sbm* (i.e., with the FRT-P*trc* cassette replacing the 204-bp up-stream of the Sbm operon)	Lab stock
ST001	CPC-Sbm *endA::*(P_{BAD}::γ:β:α)	Lab stock
ST002	CPC-Sbm *endA::*(P_{BAD}::γ:β:α) *yjcS::*(P_{tetA}::spc.-gRNA.P279T-*cas9*)	This study
ST003	CPC-Sbm *endA::*(P_{BAD}::γ:β:α) *yjcS::*(P_{tetA}::spc.-gRNA.P279T-*cas9*) *bcsA::*($P_{gracmax}$(T7.RBS)::*bktB*:*phaB*)	This study
ST004	CPC-Sbm *endA::*(P_{BAD}::γ:β:α) *yjcS::*(P_{tetA}::spc.-gRNA.P279T-*cas9*) *bcsA::* ($P_{gracmax}$(T7.RBS)::*bktB*:*phaB*) *intF::*($P_{gracmax}$(T7.RBS)::*phaC*:*phaA*	This study
ST005	ST004 Δ*iclR*	This study

ST006	ST004 Δ*iclR* Δ*sdhA*	This study
ST007	CPC-Sbm *yjcS::*(P_{tetA}::spc.-gRNA.P279T-*cas9*)	This study
Plasmids		
pDonor	Donor plasmid for *V. Cholerae* transposon system	[163]
pEffector	Effector expression for *V. Cholerae* transposon system	[163]
pUC19	Subcloning plasmid for generating Effector plasmids	Lab stock
pTrc99a	pBR322 ori, Amp^r	[182]
pSC101	$pSC101^{ts}$ ori, Cat^r	Lab stock
pDonor1	pBR322 ori, P_{tetA}::spc.-gRNA.P279T-*cas9*, Spc^r	This study
pDonor2	pBR322 ori, $P_{gracmax}$(T7.RBS)::*bktB*:*phaB*, Spc^r	This study
pDonor3	pBR322 ori, $P_{gracmax}$(T7.RBS)::*phaC*:*phaA*, Spc^r	This study
pEffector1	$pSC101^{ts}$ ori, P_{J23119}::*yjcS*-gRNA.P882T, Cat^r	This study
pEffector2	$pSC101^{ts}$ ori, P_{J23119}::*bcsA*-gRNA.P1249NT, Cat^r	This study
pEffector3	$pSC101^{ts}$ ori, P_{J23119}::*intF*-gRNA.P242T, Cat^r	This study
pbktB.phaB-Δ*bcsA*	pBR322 ori, $P_{gracmax}$(T7.RBS)::*bktB*:*phaB*, Amp^r	This study
pCas9-Δ*yjcS*	pBR322 ori, P_{tetA}::spc.-gRNA.P279T-*cas9*, Amp^r	This study
pgRNA1	pBR322 ori, $P_{xylA.SphI}$::*yjcS*-gRNA.P1032T, Spc^r	This study
pgRNA2	pBR322 ori, $P_{xylA.SphI}$::*bcsA*-gRNA.P1088T, Spc^r	This study
pgRNA3	$pSC101^{ts}$ ori, $P_{xylA.SphI}$::*iclR*-gRNA.P78T, Cat^r	This study
pgRNA4	$pSC101^{ts}$ ori, $P_{xylA.SphI}$::*sdhA*-gRNA.P392T, Cat^r	This study
pCas9	$pSC101^{ts}$ ori, P_{tetA}::spc.-gRNA.P279T-*cas9*, Cat^r	Lab stock

Standard recombinant DNA technologies were applied for molecular cloning [183]. Activation of the genomic Sbm operon in BWΔldhA to generate *E. coli* CPC-Sbm was described previously [184].

To construct pDonor1, the gRNA cassette spc.-gRNA.P279T and *cas9* were amplified using primer sets P001/P002 and P003/P004, respectively, from a lab-made plasmid for expression of *cas9* in *E. coli* harboring the spc.-gRNA.P279T gRNA cassette and the *cas9*. Primer sets P005/P006 and P007/P008 were used to amplify the TL and TR sequences, respectively, which are recognized by the integrase to excise the harboring DNA cassette. Primers P009/P010 were used to amplify the backbone from pDonor [163]. Finally, the five fragments were Gibson-assembled to form pDonor1. To construct pDonor2, *phaB* and *bktB* were amplified using primer sets P011/P012 and P013/P014, respectively, with the genomic DNA of wild-type *Cupriavidus necator* ATCC 43291 as the template. The backbone was amplified using primers P015/P016, and the TL and TR sequences were amplified using primer sets P017/P018 and P019/P020 from the pDonor template, respectively. The five fragments were Gibson-assembled to form pDonor2. Similarly, pDonor3 was constructed by

amplifying *phaC* and *phaA* using primer sets P021/P022 and P023/P024 from the genomic DNA of wild-type *C. necator* ATCC 43291. The backbone was amplified using primers P015/P016, and the TL and TR sequences were amplified using primer sets P025/P018 and P019/P020 from the pDonor template, respectively. Finally, the five fragments were Gibson-assembled to form pDonor3.

All pEffector-derived plasmids were constructed by assembling three common fragments amplified from the pEffector [163] plasmid using primer sets P026/P027, P028/P029, and P030/P031 and a specific protospacer amplified from a subcloning plasmid using P032/P033. To construct the subcloning plasmids, the protospacer sequences yjcS.P882T, bcsA.P1249NT and intF.P242T were first constructed by annealing the sense and antisense oligonucleotides using P034/P035, P036/P037, and P038/P039, respectively. The annealed oligos were then ligated into BsaI-digested pUC19-derived plasmid in which the ampicillin resistance marker was previously replaced with a spectinomycin cassette.

To construct the pbktB.phaB-Δ*bcsA* delivery vector, we began by amplifying the *bcsA* 5' and 3' homology lengths (HL-5' and HL-3') using primer sets P040/P041 and P042/P043, respectively, from the genomic DNA of *E. coli* MG1655. The DNA cassette including *bktB* and *phaB* was amplified from pDonor2 using primers P044/P045. The fragments were Gibson-assembled with the pTrc99a backbone amplified by primers P046/P047 to construct pbktB.phaB-Δ*bcsA*. The plasmid was then used as the template to obtain the PCR-amplified editing template using primers P040/P043. Similarly, to construct the pCas9-Δ*yjcS* delivery vector, we amplified the *yjcS* HL-5' and HL-3' with primer sets P048/P049 and P050/P051, respectively, from the genomic DNA of *E. coli* MG1655. The gRNA cassette spc.-gRNA.P279T and *cas9* were amplified from pDonor1 using primers P052/P053. The fragments were Gibson-assembled with the pTrc99a backbone amplified by primers P054/P055 to construct pCas9-Δ*yjcS*. The plasmid was then used as the template to obtain the PCR-amplified editing template using primers p048/P051.

The gRNA plasmids harboring gRNA cassettes *yjcS*.P1032T, bcsA.P1088NT, iclR.P87T, and sdhA.P392T were constructed by amplifying a single fragment using respective forward primers P056, P057, P058, and P059 containing unique protospacers and a common reverse primer, P060. A lab-made plasmid containing $P_{xylA.\mathrm{SphI}}$ upstream of a gRNA cassette (described previously [185]) with pBR322-ori was used as the template for pgRNA1 and pgRNA2. Another lab-made plasmid containing $P_{xylA.\mathrm{SphI}}$ upstream of a gRNA cassette with pSC101$^{\mathrm{ts}}$-ori was used as the template for pgRNA3 and pgRNA4. Finally, each fragment was self-ligated to form gRNA delivery vectors.

pCas9 was constructed by amplifying the gRNA cassette spc.-gRNA.P279T using primers P061/P062 and amplifying *cas9* using primers P063/P064 from pDonor1. The fragments were Gibson-assembled with the two fragments of backbone amplified using primer sets

P065/P066 and P067/P068 from pSC101 plasmid.

The genotypes of derived knockout strains were confirmed by colony polymerase chain reaction (PCR) using the appropriate verification primer sets listed in Table S3.1.

3.2.2 Competent cells preparation and transformations

To prepare chemically competent cells, bacterial cells on the overnight plate were used to inoculate the prewarmed Luria-Bertani (LB) medium with the initial optical density at 600 nm (OD_{600}) of 0.1. All experiments for CRISPR-Cas9 coupled with λ_{Red} recombineering were conducted at 30°C and the cultures were induced by 15 mM L-arabinose at OD_{600} of 0.3-0.4 to trigger the λ_{Red} operon. The cultures for transposon-associated CRISPR-Cas experiments were prepared at 37°C. Cells were harvested at OD_{600} of 0.5 and then washed with ice-cold 15% glycerol three times. Finally, cells were concentrated 230 times and 40 μl of the competent cells were used per transformation. Approximately 250 ng of the donor plasmid or editing template and 100 ng of the Effector or gRNA plasmid were mixed with 40 μl of electrocompetent cells. Mixture of the cells and DNA were electroporated and recovered in 1 mL SOC medium for 1.5 hours and then spread on selective LB agar plates and incubated at 30°C overnight.

For the transposition experiments, prior to restreaking, the bacteria were pooled from the plate by scraping the colonies, resuspending them in LB and appropriately diluting before plating them again. This step is beneficial for obtaining homogeneous colonies since it has been reported that some of the colonies growing on the first plate might be heterogeneous [171].

The derived mutations were confirmed by applying colony PCR toward colonies re-streaked from each plate using the appropriate primer sets listed in Table S3.1. The editing efficiencies of the mutations were reported as the ratio of the successful mutations to the total number of colonies screened for colony PCR (12 colonies per plate).

3.2.3 Media and bacterial cell cultivation

All media components were purchased from Sigma-Aldrich Co. (St Louis, MO, USA) except glucose, yeast extract, and tryptone which were obtained from BD Diagnostic Systems (Franklin Lakes, NJ, USA). The optical cell density measurements were made by appropriately diluting the samples with 0.15 M saline and using a spectrophotometer (GENESYS™·40/50·Vis/UV-Vis, Thermo Fisher Scientific Inc., Waltham, MA, USA).

When required, the media were supplimented with antibiotics at the following concentrations: 40 μg/mL chloramphenicol and 90 μg/mL spectinomycin. For double selection the spectinomycin concentration was reduced to 80 μg/mL.

E. coli mutants were streaked on LB agar plates and incubated at 30°C for 16 h. The plates were then used to to inoculate 20 mL prewarmed LB broth in 125 mL conical flasks as starter cultures at 30°C and 280 rpm for 8 h in a rotary shaker (New Brunswick Scientific, NJ, USA). The starter cultures were used to inoculate 215 mL prewarmed super broth (SB) medium (32 g/L tryptone, 20 g/L yeast extract, and 5 g/L NaCl) as seed cultures in 1 L conical flasks. The seed cultures were grown at 30°C and 280 rpm for 16 h. The cells were then harvested by centrifugation at 5000×g for 10 minutes and resuspended in 50 mL fresh LB medium which was eventually used to inoculate a 1 L stirred tank bioreactor (CelliGen 115, Eppendorf AG, Hamburg, Germany). The bioreactor cultivations operate at 30°C and 430 rpm for 24 h and cultivation medium consists of 30 g/L glycerol, 10 g/L yeast extract, 10 mM NaHCO3, 0.4 μM cyanocobalamin (vitamin B_{12}), 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), 0.23 g/L K_2HPO_4, 0.51 g/L NH_4Cl, 49.8 mg/L $MgCl_2$, 48.1 mg/L K_2SO_4, 2.78 mg/L $FeSO_4{\cdot}7H_2O$, 0.055 mg/L $CaCl_2$, 2.93 g/L NaCl, 0.72 g/L tricine and 1000^{th} dilution (i.e., 1 mL/L) trace elements (2.86 g/L H_3BO_3, 1.81 g/L $MnCl_2{\cdot}4H_2O$, 0.222 g/L $ZnSO_4{\cdot}7H_2O$, 0.39 g/L $Na_2MoO_4{\cdot}2H_2O$, 79 μg/L $CuSO_4{\cdot}5H_2O$, 49.4 μg/L $Co(NO_3)_2{\cdot}6H_2O$). Aerobic conditions were obtained by sparging air into the bulk culture at 1 vvm. The pH of the culture was maintained at 7.0 $\pm$ 0.1 using 3 M NH_4OH and 3 M H_3PO_4.

3.2.4 PHBV extraction and analysis

Intracellular polymer was extracted by applying the following procedure. The bacterial cells from the bioreactor cultivations were harvested and centrifuged at 8000×g for 10 minutes. The pellets were washed twice with distilled water and then lyophilized for 16 h. The dried cell weight (DCW) was measured prior to methanolysis. Methanolysis was carried out at 100°C for 4 h by resuspending the pellets in 2 mL chloroform and 1 mL PHA solution (i.e., 4 g/L benzoic acid and 15% sulfuric acid in methanol). The biphasic mixtures were then cooled down to room temperature and then were vortexed after adding 1 mL distilled water to them. Finally, after phase separation, the chloroform phase was filtered with 0.45 μm Polytetrafluoroethylene (PTFE) filters, appropriately diluted and then analyzed by Agilent 6890 series GC system (Agilent Technologies, Santa Clara, CA, USA) with a J & W Scientifics DB Wax column (30 m × 0.53 mm, film thickness 1 μM) (Agilent Technologies, Santa Clara, CA, USA). The oven program was set as following: initial temperature was set at 80°C for 5 minutes, then ramped to 230°C at 7.5°C/min,

and continued to ramp to 260°C at a faster rate 10°C/min followed by maintaining that temperature for the analysis. The calibration curves for methanolysis assay were prepared using the standards of methyl 3-hydroxybutyrate and methyl 3-hydroxyvalerate (Sigma-Aldrich Co., St Louis, MO, USA). The PHA content is reported as the ratio of PHA mass measured by GC to the DCW, expressed in weight percent. The 3-HV content is reported as the ratio of 3-HV to the total of 3-HV and 3-HB, expressed in mole percent.

3.2.5 Statistical analysis

All experiments in this study were conducted in either triplicate (for genome engineering) or duplicate (for bioreactor cultivations). We used t-test and ANOVA analysis to evaluate our methods and data. Readers are encouraged to check supplementray Material for details of each analysis.

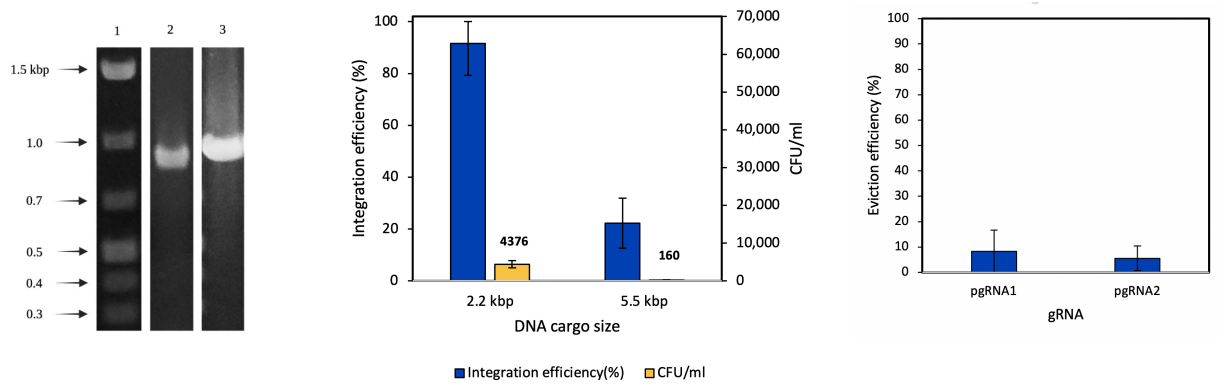
3.3 Results

3.3.1 Genomic knock-in based on CRISPR-Cas9 coupled with $\lambda_{\mathbf{Red}}$ recombineering

We first explored genomic integration of a DNA cargo (i.e., *cas9* as a large-size cargo and *bktB*:*phaB* as a small-size cargo) in *E. coli* using CRISPR-Cas9 coupled with λ_{Red} recombineering. To do this, strain ST001 was first derived by inserting the three λ_{Red} recombineering genes of *exo* (α), *beta* (β), and *gam* (γ), whose expression was regulated by the arabinose-inducible P_{BAD} promoter, into the *endA* locus of the *E. coli* genome. The three recombineering genes were annealed to the FRT-Kn^R-FRT cassette and the insertion was conducted by Flp-FRT recombination. The FRT-Kn^R-FRT cassette was then eliminated by transforming the strain with pCP20 [186], a temperature-sensitive plasmid expressing a flippase (Flp) recombinase. Upon Flp-mediated excision of the Kn^R cassette, a single Flp recognition site was generated. Plasmid pCP20 was then cured by growing cells at 42°C. Then, ST001 was transformed with pCas9 followed by two rounds of λ_{Red} recombineering. In the first round, ST001 harboring pCas9 was co-transformed with pgRNA1 expressing gRNA targeting *yjcS* and the large-size DNA cargo containing *cas9* flanked by 3' and 5' homologous arms. Genomic integration of *cas9* into the *yjcS* locus was confirmed by colony PCR using appropriate primer sets (Figure 3.3a). To prepare for the next round of genome engineering, only pgRNA1 was evicted to form ST002 harboring pCas9. Then, ST002 harboring pCas9 was co-transformed with pgRNA2 expressing gRNA

targeting *bcsA* and the small-size DNA cargo containing *bktB*:*phaB* flanked by 3' and 5' homologous arms. Finally, after confirming the successful integration of *bktB*:*phaB* into the *bcsA* locus by colony PCR (Figure 3.3a), both pCas9 and pgRNA2 were evicted to form ST003.

Efficiencies of 91.7% and 22.2% were obtained for genomic integration of the small- and large-size DNA cargo, respectively, with low volumetric CFU for both cases (Figure 3.3b). Note that both the integration efficiency and volumetric CFU decreased with an increased size of DNA cargo (t-test with p-values of $2.3e^{-3}$ and $1.5e^{-3}$ respectively), and the integration efficiency appeared to be limited by high numbers of unedited background colonies (also known as escaper colonies) evading the Cas9 break. In addition, low gRNA eviction efficiencies were observed (Figure 3.3c), potentially due to a high copy number of the plasmid with the pBR322 origin, limiting the applicability of this system for genome engineering.



(a) (b) (c)

Figure 3.3: (**a**) Colony PCR screening of *cas9* and *bktB*:*phaB* integrations via λ_{Red} recombineering (column 2-3). To screen for *cas9* integration , primers P069/P070 amplified an 858 bp fragment (column 2). To screen for *bktB*:*phaB* integration, primers P071/P072 amplified a 930 bp fragment (column 3). (**b**) Integration efficiency based on colony PCR (ratio of colonies containing the desired mutation to the total number of colonies screened for colony PCR) and colony count (CFU/mL) of each transformation employing λ_{Red} recombineering system to integrate small- and large-size DNA cargos. (**c**) pgRNA plasmid eviction efficiencies confirmed by spectinomycin sensitivity (ratio of colonies with successful plasmid eviction to total number of colonies screened for antibiotic sensitivity). All values are reported as means ± SD (n = 3).

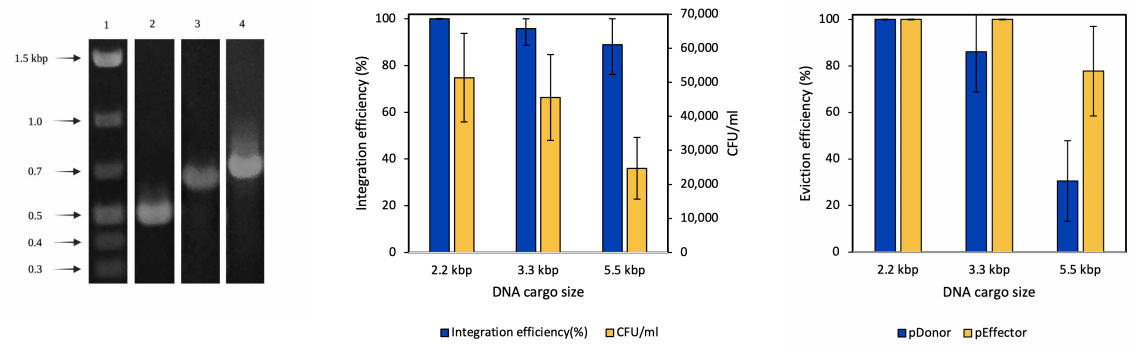
3.3.2 Genomic knock-in based on transposon-associated CRISPR-Cas system

In light of the above technical disadvantages for the CRISPR-Cas9 coupled with λ_{Red} recombineering, we further explored transposon-associated CRISPR-Cas system for genomic integration. To do this, ST001 was first used as the host for genomic insertion of *cas9* as a large-size cargo into the *yjcS* locus by co-transforming pDonor1 (which carries *cas9*) and pEffector1 (which carries the gRNA targeting *yjcS*). Genomic integration was confirmed by colony PCR using appropriate primer sets (Figure 3.4a). For the next round of genome engineering, both plasmids were evicted from the resulting engineered strain, generating ST002. Then, the DNA cargo containing *bktB*:*phaB* as a small-size cargo was integrated into the *bcsA* locus of ST002 by co-transforming pDonor2 and pEffector2. Both plasmids were evicted from the resulting engineered strain, generating ST003. Finally, for complete implementation of the PHBV biosynthetic pathway, another DNA cargo containing *phaC*:*phaA* as a medium-size cargo was inserted into the *intF* locus of ST003 by co-transforming pDonor3 and pEffector3. Both plasmids were evicted from the resulting engineered strain, generating ST004.

Efficiencies of 100%, 97.1%, and 88.9% were obtained for genomic integration of a small-, medium-, and large-size cargo, respectively, with high volumetric CFU for all cases (Figure 3.4b). Although the integration efficiency based on Tn6677 transposition appeared to decrease with an increased DNA cargo size, it was high enough to make the system a feasible toolkit for genome engineering. The ANOVA analysis also confirms that decrease in the integration efficiency is not significant (See Table S3.2, p-value = $2.7e^{-1}$). A similar decreasing trend, though not significant (See Table S3.3, p-value = $2.3e^{-1}$), was observed in volumetric CFU upon increasing the DNA cargo size. In addition, the efficiency of plasmid eviction, particularly for pDonor, also dropped considerably with an increased DNA cargo size (See Table S3.4, p-value = $2.2e^{-3}$) (Figure 3.4c), potentially limiting the applicability of this system.

3.3.3 Gene knockout based on CRISPR-Cas9 coupled with λ_{Red} recombineering

Next, we assessed the feasibility of the CRISPR-Cas9 coupled with λ_{Red} recombineering system for gene knockout using the *iclR* and *sdhA* genes for demonstration. To overcome the low efficiency for eviction of the pgRNA plasmids, the pBR322 origin was replaced with the temperature-sensitive $\text{pSC101}^{\text{ts}}$ one. To knockout *iclR,* ST004 was co-transformed with pgRNA3 expressing gRNA targeting *iclR* and a 60 bp oligonucleotide flanked by 3'

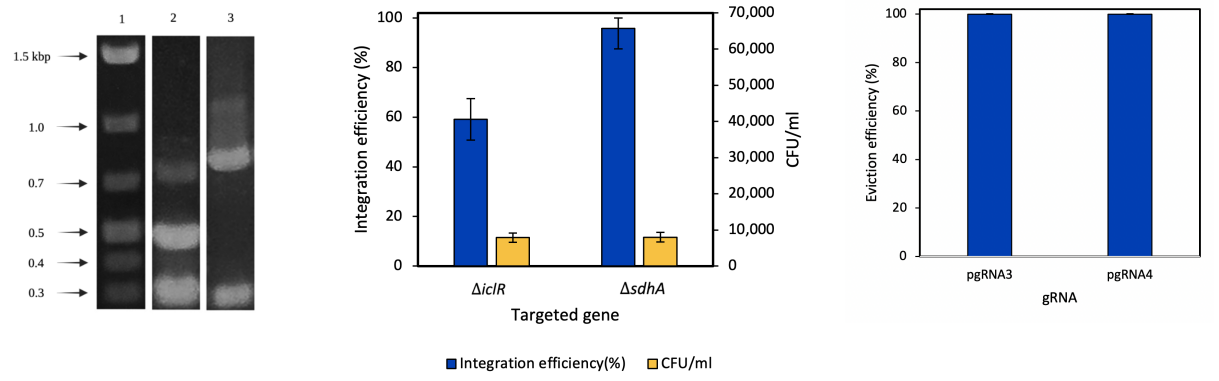


(a) (b) (c)

Figure 3.4: (**a**) Colony PCR screening of *cas9*, *bktB*:*phaB* and *phaC*:*phaA* integrations via Tn6677 transposons (column 2-4). To screen for *cas9* integration, primers P070/P073 amplified a 484 bp fragment (column 2). To screen for *bktB*:*phaB* integration, primers P072/P074 amplified a 641 bp fragment (column 3). To screen for *phaC*:*phaA* integration, primers P075/P076 amplified a 704 bp fragment (column 4). (**b**) Integration efficiency (ratio of colonies containing the desired mutation to the total number of colonies screened for colony PCR) and colony count (CFU/mL) of each transformation employing transposon-associated CRISPR-Cas system to integrate small-, medium- and large-size DNA cargos. (**c**) pDonor and pEffector plasmid eviction efficiencies (ratio of colonies with successful plasmid eviction to total number of colonies screened for antibiotic sensitivity). All values are reported as means ± SD (n = 3).

and 5' homology arms. The oligonucleotide containing six consecutive point mutations was designed to introduce two consecutive stop codons and an *Ase*I restriction site for screening of successful recombination, which was confirmed by colony PCR using appropriate primer sets (Figure 3.5a). For the next round of gene knockout, pgRNA3 was evicted to form ST005, which was then co-transformed with pgRNA4 expressing gRNA targeting *sdhA* and the respective oligonucleotide for recombination. After confirming successful knockout by colony PCR, pgRNA4 was evicted from the resulting strain to form ST006.

Efficiencies of 61.1% and 94.4% were obtained for knocking out the *iclR* and *sdhA* genes, respectively (Figure 3.5b). Note that pgRNA plasmids were evicted with 100% efficiency for both rounds of gene knockout (Figure 3.5c), suggesting the feasibility of this system for genome engineering. Also, note that the control experiments using ST007 (without three λ_{Red} recombineering genes on the genome) as the host had no genomic integration (data not shown), highlighting the importance of homologous recombination-mediated by the λ_{Red} system for efficient gene knockout.



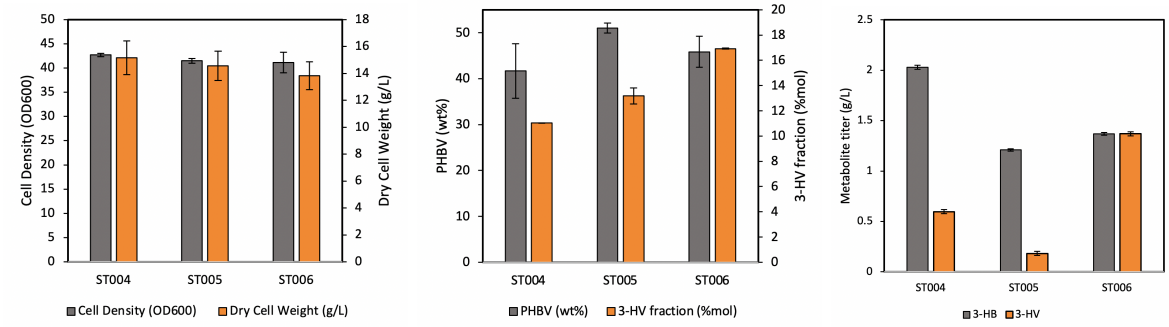
(a) (b) (c)

Figure 3.5: (**a**) Colony PCR screening of *iclR* and *sdhA* knockouts via λ_{Red} recombineering (column 2-3). To screen for *iclR* knockout, primers P077/P078 amplified a 790 bp fragment and successful recombination of the editing template generated products of 490 bp and 300 bp upon AseI digestion (columns 2). To screen for *sdhA* knockout, primers P079/P080 amplified a 1065 bp fragment and successful recombination of the editing template generated products of 779 bp and 286 bp upon AseI digestion (column 3). (**b**) Gene knockout efficiency based on colony PCR (ratio of colonies containing the desired mutation to the total number of colonies screened for colony PCR) and colony count (CFU/mL) of each transformation employing λ_{Red} recombineering system to knockout *iclR* and *sdhA* genes. (**c**) pgRNA plasmid eviction efficiencies confirmed by chloramphenicol sensitivity (ratio of colonies with successful plasmid eviction to total number of colonies screened for antibiotic sensitivity). All values are reported as means ± SD (n = 3).

3.3.4 PHBV biosynthesis using genome-engineered strains

PHBV production mainly relies on the availability of the monomer of 3HV-CoA, which is derived from succinyl-CoA in the tricarboxylic acid (TCA) cycle (Figure 3.2). It was previously shown that, by inactivating the oxidative TCA branch with the *sdhA* mutation and deregulating the glyoxylate shunt with the *iclR* mutation, the 3-HV monomeric fraction in the PHBV copolymer could be regulated [187]. A similar demonstration was conducted herein using the three genome-engineered and plasmid-free strains (i.e., ST004, ST005, and ST006) for bioreactor cultivation under aerobic conditions. The results suggest that cell growth was minimally affected by the *iclR* and *sdhA* mutations (Figure 3.6a). Although the DCW was slightly reduced (i.e., 4% and 9% for ST005 and ST006, respectively, compared to ST004), the PHBV content was higher for ST005 (51%) and ST006 (45.9%) compared to ST004 (41.7%). Moreover, the 3-HV monomeric fraction of PHBV copolymer also

increased in ST005 (13.2 mol%) and ST006 (17 mol%) compared to ST004 (11 mol%) (t-test with p-values of $4.1e^{-2}$ and $2.5e^{-5}$ respectively) (Figure 3.6b). In addition, the 3-HV titer was also incresed by inactivating both *iclR* and *sdhA* genes, confirming the successful flux channeling into the Sbm pathway (Figure 3.6c). Also, the titers of 3-HB and 3-HV monomers suggest that the PhaC is the limiting factor for PHBV synthesis, an outcome also described previously in *E. coli* [188]. The results suggest that not only PHBV could be produced by these genome-engineered and plasmid-free strains but also the 3-HV monomeric fraction of PHBV copolymer could be modulated by introducing *iclR* and *sdhA* mutations.



(a) (b) (c)

Figure 3.6: PHBV biosynthesis of the metabolically manipulated strains (ST004, ST005, ST006) (**a**) presenting cell optical density (OD600) and dried cell weight (g/L) (**b**) biopolymer and 3-HV content represented by wt% and mol%, respectively and (**c**) Metabolite production (g/L). All values are reported as means ± SD (n = 2).

3.4 Discussion

Today, strategies based on CRISPR have been developed for genome engineering of various microorganisms, particularly *E. coli*. While most of these strategies mainly focus on gene knockout, site-specific insertion of heterologous genes into the genome remains a challenge [189, 190] particularly for large-size genes. In this study, we developed an efficient toolkit for both site-specific inactivation of native genes based on a traditional CRISPR-Cas9 coupled with λ_{Red} recombineering system (i.e., gene knockout) and site-specific insertion of heterologous genes based on a transposon-associated CRISPR-Cas system (i.e., gene knock-in) in *E. coli*. With specific engineered strains being developed, the number of plasmids simultaneously used for genome editing was minimized. Also, the toolkit showed

high editing efficiencies with effective plasmid curing, facilitating multiple rounds for successive genome editing. Using the toolkit, customized engineered strains were developed for heterologous production of PHBV in plasmid-free *E. coli*.

To evaluate the system of CRISPR-Cas9 coupled with λ_{Red} recombineering for gene knock-in, we first constructed a plasmid-free strain ST001 by insertion of a DNA cassette containing the three λ_{Red} recombineering genes, i.e., *exo* (α), *beta* (β), and *gam* (γ) whose expression was regulated by the arabinose-inducible P_{BAD} promoter, into *E. coli* genome. With the expression of the λ_{Red} recombineering genes in ST001, homologous recombination was significantly enhanced for genomic insertion of DNA cargos while rejoining the excised genome by CRISPR. Though the λ_{Red} recombineering genes might not be critical for conducting CRISPR-mediated gene knockout which is often associated with the use of a small editing template, our results clearly showed that they were required for conducting CRISPR-mediated gene knock-in and the editing efficiency was affected by the size of inserted DNA cargo. It was previously reported that, upon large fragment recombineering via λ_{Red} recombineering, the dsDNA cargo was prone to mutations and mismatches, resulting in inefficient DNA recombination [191]. Our findings are also in agreement with a previous report of the size effects of DNA cargo, i.e., ineffective genomic integration of large DNA cargo up to 3 kbp and even failed integration when the size was increased to 3.8 kbp [192]. Similar reduction in integration efficiency was observed by Maresca et al. for DNA insertions greater than 3 kbp while slight decrease in efficiency was reported for gene deletion up to 50 kbp using λ_{Red} recombineering [193]. In another study, Jiang et al. attempt to integrate a 4.5 kbp DNA cassette into the *yjcS* locus yielded the maximum efficiency of 28% [166]. Li et al. also reported significant decrease in integration efficiency from 59% to 14% by increasing the DNA cassette size from 3 kbp to 8 kbp [194]. In addition, the volumetric CFU decreased with an increased size of DNA cargo with the appearance of many unedited escaper colonies. These colonies might have escaped the double-strand break due to homologous recombination between the cleaved and un-cleaved chromosome, deletion of targeted regions, or mutations in the protospacer sequence of the gRNA [195, 196], limiting the genome editing efficiency. A critical step for genome engineering is plasmid curing to form the strain for successive genome editing. We observed low plasmid curing efficiency (<10%) in the system of CRISPR-Cas9 coupled with λ_{Red} recombineering potentially due to the small size and high copy number of the plasmid. This issue was addressed by changing the plasmid origin to a low-copy one.

Given low editing efficiency upon using CRISPR-Cas9 coupled with λ_{Red} recombineering for gene knock-in, we further evaluated the transposon-associated CRISPR-Cas system with a pDonor plasmid carrying the DNA cargo and a pEffector plasmid carrying *V. cholerae* Tn6677 transposon, type I-F CRISPR-Cas machinery, and a protospacer tar-

geting the desired locus. Upon performing a two-way ANOVA analysis on our results, we concluded that though CRISPR-Cas9 coupled with λ_{Red} recombineering might be feasible for integrating DNA cassettes up to 3 kbp, it was rather inefficient compared to the transposon-associated CRISPR-Cas system (p-value = $3.1e^{-4}$) which appeared to be more suitable for integrating larger or pathway-sized DNA cassettes (t-test with p-value = $9.7e^{-4}$). The ANOVA analysis also suggest that there is an interation between the DNA cargo size and the method of integeration (See Table S3.5, p-value = $1.5e^{-3}$). The transposon-associated CRISPR system also had much higher volumetric CFU and effective plasmid curing upon genomic integration of large-size DNAs, compared to CRISPR-Cas9 coupled with λ_{Red} recombineering. In addition to high editing efficiency and effective plasmid curing, DNA integration via Tn6677 transposons did not require long homologous arms in the DNA cargo and the integration could occur in a single-step transformation.

Finally, we assessed the feasibility of combining the two systems for multiple rounds of genome engineering, i.e., employing the CRISPR-Cas9 coupled with λ_{Red} recombineering system for gene knockout whereas the transposon-associated CRISPR-Cas system for gene knock-in. To facilitate the application, plasmid-free ST002 harboring both the three λ_{Red} recombineering genes and *cas9* on the genome was used as the base strain such that all genome editing steps could be performed with one-step transformation and plasmid curing. To address the low eviction efficiency for plasmid curing, we used the temperature-sensitive $\text{pSC101}^{\text{ts}}$ origin to reduce the copy number of the plasmids for genome editing. Furthermore, genomic expression of the *cas9* gene will facilitate the eviction of pDonor plasmid when using Tn6677 transposons. For the demonstration using ST002 with an activated Sbm operon on the genome, we first knocked in four pathway genes (i.e., *bktB*, *phaA*, *phaB*, and *phaC* from *C. necator*) for heterologous biosynthesis of PHBV. β-ketothiolase which is encoded by *phaA* and *bktB,* is involved in two reactions to condensate two acetyl-CoA or an acetyl-CoA with propionyl-CoA to form acetoacetyl-CoA and 3-ketovaleryl-CoA, respectively. Previous studies have shown that *bktB* exhibits more selectivity towards ketovaleryl-CoA formation, thus more favorable for increasing the 3-HV content of the biopolymer [197]. Acetoacetyl-CoA reductase encoded by *phaB* reduces the acetoacetyl-CoA and 3-ketovaleryl-CoA to 3-HB-CoA and 3-HV-CoA, respectively. Eventually, polyhydroxyalkanoate synthase encoded by *phaC* mediates the polymerization of 3-HB-CoA and 3-HV-CoA monomers to form PHBV granules in *E. coli* cells (Figure 3.2) [198, 199]. Subsequently, we knocked out two native genes (i.e., *iclR* and *sdhA*) to direct more dissimilated carbon flux toward the Sbm pathway [200]. Since glycerol has a high degree of reduction and compared to glucose it produces approximately twice the number of reducing equivalents (i.e., NADH) [118], it was used as the carbon source for PHBV production. In addition, glycerol is an abundant byproduct in industry which makes it an inexpensive carbon source [201]. The intermediate plasmid-free ST004, with the PHBV pathway genes

being inserted into the genome of ST002, could heterologously produce PHBV at 41.7% of the DCW with 11 mol% of the 3-HV monomeric fraction. With further introduction of two knockouts of Δ*iclR* and Δ*sdhA* in ST006, the PHBV content was increased up to 45.9% of the DCW and the 3-HV monomeric fraction to 17 mol% without affecting cell growth. Such metabolic effects for PHBV biosynthesis were similar to those previously reported by Miscevic et al. using engineered *E. coli* strains with expression plasmids containing PHBV-biosynthetic genes [187]. They have reported slight impcats of *iclR* and *sdhA* knockouts on cell growth and total biopolymer content of PHBV (i.e., OD600 35 and DCW 12.5 g/L) in aerobic conditions. They also reported an increse in 3-HV monomeric fraction of the PHBV by knocking these two genes, however, they obtained the 3-HV monomeric fraction of 40.9 mol% which could possibly be due to higher expression levels of the PHBV-biosynthetic genes by expression plasmids.

Overall, our toolkit can be conveniently applied for flexible genome engineering by designing user-specific gRNAs and DNA cargos. Using our base engineered strain, ST002, we could facilitate genome engineering, either gene knockout or knock-in, through a single-step transformation. Moreover, our method relies on the use of curable plasmids which possess low copy number or temperature-sensitive origins. In addition, increasing the expression levels of inserted genes can be achieved by modifying the donor DNA cassette using strong promoters or ribosome binding sites, and even by integrating multiple copies of the donor DNA cassette into different loci in the genome.

Chapter 4

Bio-Based Production of Uroporphyrin in *Escherichia coli*

4.1 Introduction

Porphyrins are a group of colorful chemical compounds with an interconnected tetrapyrrole structure [202]. Among them, heme and chlorophyll are the two most well-known members which respectively give red color to blood and green color to plants, algae, and cyanobacteria. Due to versatile and novel applications in various sectors, including therapeutics, molecular electronics, and polymers, porphyrins have gained significant interest [203, 204]. Currently, most porphyrins are produced through chemical synthesis [205, 206, 207]. Extraction from natural sources such as animal tissues and plants are also reported [208, 209]. However, the chemical methods might not be technologically effective and environmentally friendly, and often result in low yields with no selective production of specific porphyrin isomers [210, 11]. Extraction from biological sources is also limited by low yield, seasonal variations, and complexities in purification [211]. The complex downstream processing for porphyrin purification further impacts the efficiency and economic feasibility of these methods. Bio-based production, which has been recognized as a more sustainable approach for the production of many value-added chemicals and fuels [212], can offer a potential solution to overcome these limitations. Bio-based systems have potentials for continuous production, thereby reducing time and resource consumption. Their scalability, higher yields, and reduced environmental impact make them a promising alternative to traditional methods.

Heme biosynthesis is highly conserved across biological kingdoms and the associated biosynthetic pathways, mechanisms, and enzymes are well documented [213]. Basically,

the common precursor for biosynthesis of tetrapyrroles is 5-aminolevulinic acid (5-ALA), which can be derived via either the C5 or C4/Shemin pathway with different upstream conversion steps for 5-ALA generation [214]. For post-5-ALA conversion steps, bacteria develop two distinct, i.e., classical (protoporphyrin-dependent; PPD) and non-classical (coproporphyrin-dependent; CPD), branches toward heme biosynthesis [213]. Portions of the PPD and CPD branches might coexist in an organism, leading to different mechanisms/pathways for heme biosynthesis under different environmental conditions and challenges. While bacterium *Escherichia coli* (*E. coli*) primarily utilizes the PPD branch for heme biosynthesis (Figure 4.1), the simultaneous presence of several enzymes and pathways diverging from the PPD branch offers potential opportunities for biosynthesis of specific porphyrins other than heme, such as uroporphyrin (UP) and coproporphyrin (CP). Herein, we focused on UP for its bio-based production in *E. coli*. Note that UP has potential applications in novel therapeutics and diagnostics, such as photodynamic therapy (PDT) for cancer treatment and antimicrobial therapy [215]. Traditional chemical synthesis of UP requires specific catalysts and involves complicated oxidation steps, making the process intricate and environmentally unfriendly [207]. To date, there has been no report on bio-based production of UP.

In *E. coli*, UP can be an end product in the metabolic branch divergent from the heme biosynthetic pathway at either hydroxymethylbilane (HMB; leading to uroporphyrin I (UP-I)) or uroporphyrinogen III (UPG-III; leading to uroporphyrin III (UP-III)) (Figure 4.1). Although *E. coli* has the native C5 pathway, it can be metabolically and energetically constrained for 5-ALA generation since this route requires the expression of multiple tightly regulated enzymes, the utilization of ATP and NADPH as limiting cofactors, and the formation of unstable glutamate-1-semialdehyde (GSA) intermediates [216]. To address these limitations in this study, we exploited the Shemin/C4 pathway by heterologous expression of *hemA* from *Rhodobacter sphaeroides* as we have previously developed an effective metabolic strategy [216] to redirect dissimilated carbon flux within the citric acid cycle toward succinyl-CoA, which is a key precursor/moiety to form 5-ALA. Such metabolic strategies were applied to enhance several value-added chemicals derived from succinyl-CoA, including propionate [200], 3-hydroxyvalerate [217], porphobilinogen (PBG) [218], and poly(3-hydroxybutyrate-co-3-hydroxyvalerate) [187].

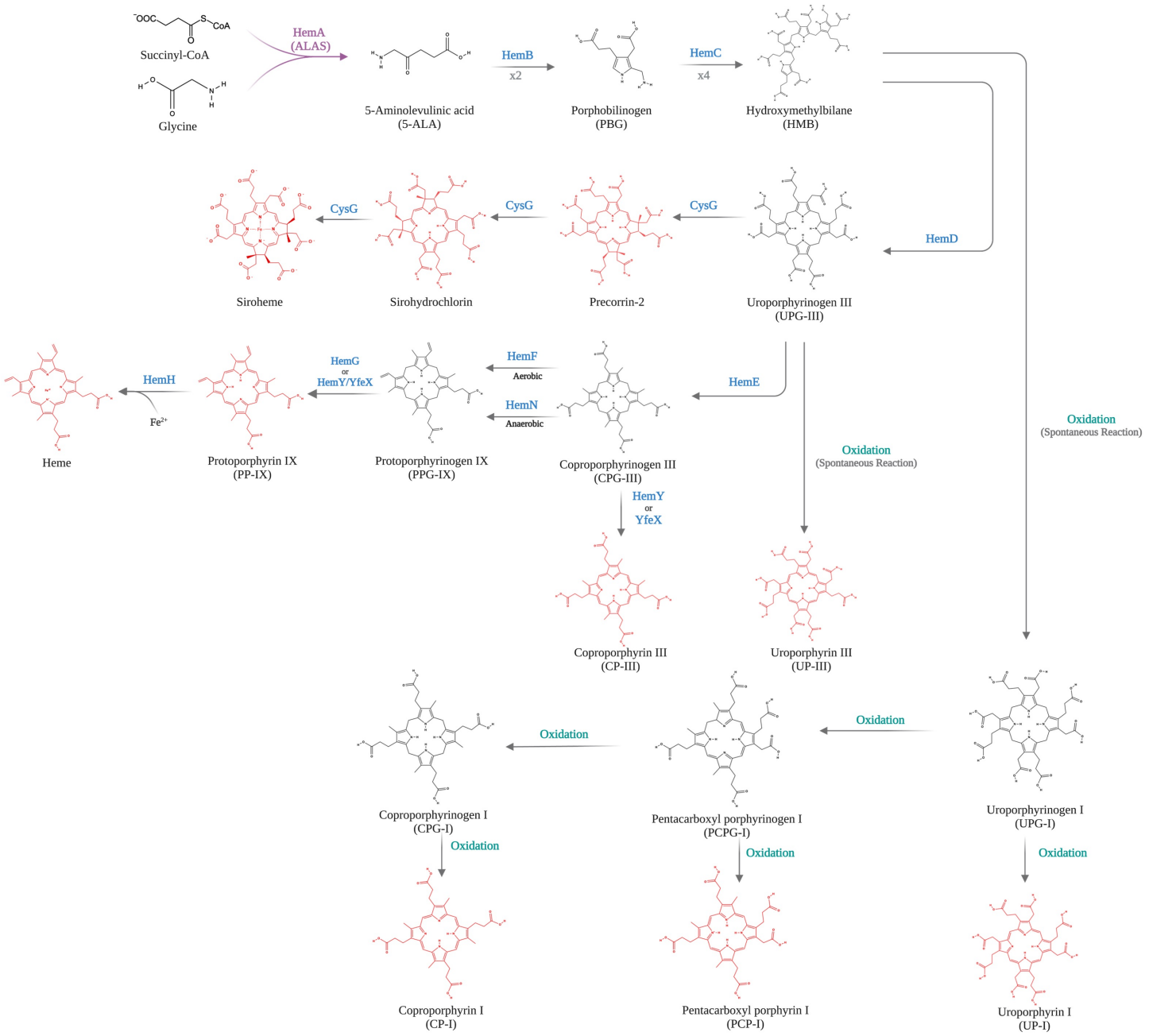


Figure 4.1: The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. Shemin pathway (in violet); Porphyrin biosynthesis (in grey). Pigments (in red); 5-ALA, 5-aminolevulinic acid; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; HMB, Hydroxymethylbilane; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HemY, coproporphyrinogen III oxidase; PBG, porphobilinogen; PCP-I, pentacarboxyl porphyrin I; PCPG-I, pentacarboxyl porphyrinogen I; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase.

After 5-ALA formation, two 5-ALA molecules are fused to form PBG by PBG synthase (HemB). Four PBG molecules are subsequently interconnected and cyclized by PBG deaminase (HemC) to form a linear tetrapyrrole HMB. At this point, the pathway can diverge to form either of the two stereoisomers, i.e., UPG-I via autoxidation or UPG-III via uroporphyrinogen III synthase (HemD). Both UPG-I and UPG-III can be spontaneously oxidized to form UP-I and UP-III, respectively, as metabolic end products. Alternatively, UPG-I/III can be enzymatically converted to coproporphyrinogen I/III (CPG-I/III) for subsequent formation of coproporphyrin I/III (CP-I/III) as metabolic end products [219, 220].

Modern biotechnologies, such as synthetic biology, metabolic engineering, and genome engineering, have been extensively applied for genetic construction of microbial cell factories [221, 222]. However, such bio-based production approaches remain challenging often due to the complex molecular and regulatory mechanisms associated with the engineered genetic and metabolic elements. For effective production of target metabolites, it is critical to implement not only proper metabolic pathways but also optimal tuning of the expression of various pathway genes, particularly those associated with the key conversion step(s) limiting the overall metabolite productivity. In this study, we heterologously implemented the Shemin/C4 pathway in an *E. coli* strain with an enlarged intracellular pool of succinyl-CoA to enhance UP biosynthesis. By exploring different combinations of selective pathway genes for heterologous expression, we were able to characterize individual gene effects and identify key step(s) limiting UP biosynthesis. To further increase the economic feasibility of the developed bioprocess, we used glycerol, which has emerged as a promising carbon source for bio-based production due to its low cost, abundance, and highly reduced nature [223], for cultivation of engineered *E. coli* strains with enhanced UP biosynthesis.

4.2 Materials and Methods

4.2.1 Bacterial Strains and Plasmids

The bacterial strains and plasmids utilized in this research are described in Table 4.1, whereas the primer sequences are provided in Table S4.1. Genomic DNA from bacterial cells was extracted with the Qiagen Blood & Tissue DNA Isolation Kit (Hilden, Germany). Taq DNA polymerase was obtained from New England Biolabs (Ipswich, MA, USA). All host cells in this study were derived from CPC-Sbm by introducing mutation(s) in the *sdhA* and/or *iclR* genes to direct the carbon flux in the tricarboxylic acid (TCA) cycle toward succinyl-CoA formation ensuring abundant supply of precursors [200]. Note that CPC-Sbm was derived from BW25113 by inactivating the *ldhA* gene [184]. *E. coli* HI-Control 10G (Lucigen, Middleton, WI, USA) was used for molecular cloning. Plasmids

were purified using the Qiagen Miniprep kit. DNA sequencing was performed by The Centre for Applied Genomics (TCAG) (Toronto, ON, Canada). Gibson assembly [181] was employed for plasmid construction. All oligonucleotides were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA).

Plasmid pK-hemABCD was first constructed to serve as the template for the other plasmids in this study. The *hemA* gene was amplified by polymerase chain reaction (PCR) from the genomic DNA of *R. sphaeroides* DSM 158 using primers P001/P002. Similarly, *hemB*, *hemC*, and *hemD* genes were amplified from the genomic DNA of *E. coli* MG1655 using primer sets P003/P004, P005/P006, and P007/P008, respectively. The backbone, consisting of p15A ori and P_{trc} with a kanamycin resistance marker, was amplified from a lab-made plasmid using primers P009/P010. Subsequently, these five fragments were Gibson-assembled to form pK-hemABCD. To construct pK-hemA, primers P011/P012 were first phosphorylated and then used to amplify a single fragment from pK-hemABCD as template. The resulting fragment was self-ligated to form pK-hemA. Similarly, a single fragment was amplified from pK-hemABCD with phosphorylated primers P011/P013. This fragment was self-ligated to form pK-hemAB. pK-hemABC was constructed using primer sets P014/P002 and P003/P015 to amplify the backbone and *hemA* in the first fragment and *hemB* and *hemC* in the second fragment. These two fragments were Gibson-assembled to form pK-hemABC. pK-hemABC^{B} was constructed by first amplifying *hemA* and *hemB* as a single fragment from pK-hemABCD using primers P001/P004. Primers P016/P017 were used to amplify *hemC* from the genomic DNA of *Bacillus subtilis* 168. The backbone was amplified using primers P018/P019 from pK-hemABCD. These three fragments were Gibson assembled to form pK-hemABC^{B}. Finally, pK-hemABD was constructed by amplifying *hemA* and *hemB* using primers P001/P004 and amplifying *hemD* along with the backbone using primers P020/P021 and pK-hemABCD as template. These two fragments were Gibson-assembled to form pK-hemABD.

Table 4.1: Strains and plasmids used in this study.

Name	Description or Relevant Genotype	Source
Host strains		
HI-Control 10G	*mcrA,* Δ(*mrr-hsdRMS-mcrBC*), *endA1, recA1,* ϕ80*dlacZ*ΔM15, Δ*lacX74, araD139,* Δ(*ara leu*)*7697, galU, galK, rpsL* (Str^R), *nupG,* λ^- , *tonA,* Mini-F $lacI^{q1}$ ($Gent^R$)	Lucigen
MG1655	K-12; F^-, λ^-, *rph-1*	Lab stock
Bacillus Subtilis 168	Wild type	Lab stock
CPC-Sbm*ΔiclRΔsdhA*	F^-, Δ*(araD-araB)567,* Δ*lacZ4787(::rrnB-3),* λ^-, *rph-1,* Δ*(rhaD-rhaB)568, hsdR514,* Δ*ldhA,* P_{trc}*::sbm* (i.e., with the FRT-P_{trc} cassette replacing the 204-bp upstream of the Sbm operon), Δ*iclR,* Δ*sdhA*	[200]
CPC-Sbm*ΔsdhA*	F^-, Δ*(araD-araB)567,* Δ*lacZ4787(::rrnB-3),* λ^-, *rph-1,* Δ*(rhaD-rhaB)568, hsdR514,* Δ*ldhA,* P_{trc}*::sbm* (i.e., with the FRT-P_{trc} cassette replacing the 204-bp upstream of the Sbm operon), Δ*sdhA*	[200]
DMA	CPC-Sbm*ΔiclRΔsdhA*/pK-hemA	This study
DMB	CPC-Sbm*ΔiclRΔsdhA*/pK-hemAB	This study
DMC	CPC-Sbm*ΔiclRΔsdhA*/pK-hemABC	This study
DMC^B	CPC-Sbm*ΔiclRΔsdhA*/pK-hemABC^B	This study
DMD	CPC-Sbm*ΔiclRΔsdhA*/pK-hemABD	This study
SMD	CPC-Sbm*ΔsdhA*/pK-hemABD	This study
Plasmids		
pK-hemABCD	p15A ori, Km^R, P_{trc}*::hemABCD*	This study
pK-hemA	p15A ori, Km^R, P_{trc}::*hemA*	This study
pK-hemAB	p15A ori, Km^R, P_{trc}::*hemAB*	This study
pK-hemABC	p15A ori, Km^R, P_{trc}::*hemABC*	This study
pK-hemABC^B	p15A ori, Km^R, P_{trc}::*hemABC -hemC* from *B. subtilis 168*-	This study
pK-hemABD	p15A ori, Km^R, P_{trc}::*hemABD*	This study

4.2.2 Media and Bacterial Cell Cultivation

All medium components were obtained from Sigma-Aldrich Co. (St. Louis), except for yeast extract and tryptone, which were obtained from BD Diagnostic Systems (Franklin Lakes). *E. coli* strains were preserved as glycerol stocks at -80 °C and streaked on lysogeny

broth (LB; 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl) agar plates, followed by incubation at 37 °C for 14–16 h.

For bioreactor cultivation, individual single colonies were picked from Luria-Bertani (LB) plates to inoculate 12 mL of super broth (SB) medium (32 g/L tryptone, 20 g/L yeast extract, and 5 g/L NaCl) in a 125 mL conical flask. The culture was incubated at 37 °C and 280 revolutions per min (rpm) using a rotary shaker (New Brunswick Scientific) for 4–6 h and subsequently used as a seed culture to inoculate 220 mL of SB medium at a 2% (v/v) concentration in a 1 L conical flask. This second seed culture was incubated at 37 °C and 280 rpm for 14–16 h. Cells were harvested by centrifugation at 4,500 g and 20 °C for 8 min and resuspended in 30 mL of fresh SB medium. The resuspended culture was used to inoculate a stirred-tank bioreactor (CelliGen 115, Eppendorf AG, Hamburg, Germany) with a working volume of 0.8 L at 37 °C and 430 rpm. The semi-defined production medium in the batch bioreactor contained 30.0 g/L glycerol, 0.23 g/L K_2HPO_4, 0.51 g/L NH_4Cl, 49.8 mg/L $MgCl_2$, 48.1 mg/L K_2SO_4, 1.52 mg/L $FeSO_4$, 0.055 mg/L $CaCl_2$, 2.93 g/L NaCl, 0.72 g/L tricine, 10.0 g/L yeast extract, 10.0 mM $NaHCO_3$, and 1.0 mL/L trace elements (2.86 g/L H_3BO_3, 1.81 g/L $MnCl_2{\cdot}4H_2O$, 0.222 g/L $ZnSO_4{\cdot}7H_2O$, 0.39 g/L $Na_2MoO_4{\cdot}2H_2O$, 79.0 μg/L $CuSO_4{\cdot}5H_2O$, 49.4 μg/L $Co(NO_3)_2{\cdot}6H_2O$) [224], supplemented with 0.05 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). To avoid glycine limitation during the cultivation, 2.0 g of glycine was supplemented into the bioreactor 30 h after the inoculation.

Aeration Level I (AL-I) with an aerobic condition was maintained by continuously purging the air into the bulk culture at 1 volume of air per volume of liquid per min (vvm). Aeration Level II (AL-II) was maintained by initially purging the air into the bulk culture at 1 vvm for the first 30 h, followed by switching to a microaerobic condition by purging the air into the headspace at 0.1 vvm. Aeration Level III (AL-III) was similar to AL-II during the initial aerobic phase; but followed by switching to an anaerobic condition by purging nitrogen (N_2) into the headspace. The pH of the bioreactor culture was maintained at 7.0 $\pm$ 0.1 using 3.0 M NH_4OH and 3.0 M H_3PO_4.

4.2.3 Analysis

Culture samples were appropriately diluted using a 0.15 M saline solution to measure cell density in OD_{600} with a spectrophotometer (DU520, Beckman Coulter, Fullerton, CA). All culture samples were washed one time with 0.15 M saline solution prior to OD_{600} absorbance measurement. To prepare cell-free medium, culture samples were centrifuged at 17,000 g for 1 min and then filter sterilized using a 0.2 μm syringe filter. Extracellular metabolites and glycerol were analyzed by high-performance liquid chromatography

(HPLC; LC-10AT; Shimadzu, Kyoto, Japan) with a refractive index detector (RID; RID-10A; Shimadzu) and a chromatographic column (Aminex HPX-87H; Bio-Rad Laboratories, CA, USA). The HPLC column temperature was maintained at 35 °C and the mobile phase was 5.0 mM H_2SO_4 (pH 2) running at 0.6 ml/min. The RID signal was analyzed by a data processing software (Clarity Lite, DataApex, Prague, Czech Republic). The concentrations of 5-ALA and PBG in the cell-free medium were determined using an assay with a modified Ehrlich's reagent [225]. Porphyrins were analyzed using an HPLC (2690 separation module, Waters™) with a photodiode array (PDA; 2996 PDA detector, Waters™) detector and a chromatographic column (Chromolith® HighResolution RP-18 endcapped, Supelco, Darmstadt, Germany). The UV absorbance was measured at 400 nm and the signal was analyzed using a data processing software (Empower 3, Waters™, Milford, USA). A previously described mobile-phase system was employed [11] with minor modifications. Briefly, the mobile phase consisted of two solvent mixtures: solvent A (7.7 g/L ammonium acetate, 125.0 mL/L acetonitrile, adjusted to pH 5.17 with glacial acetic acid) and solvent B (a methanol-glacial acetic acid mixture with a volume ratio at 10:1). The flow rate of the mobile phase was 1 mL/min. An isocratic elution step was used with 100% solvent A for 5 min, followed by a linear gradient from 0 to 100% solvent B over 30 min, and further followed by 100% solvent B for 5 min. The column temperature was maintained at 45 °C. The percentage yield of UP was calculated as the mole ratio of the produced UP to the theoretical maximum UP production based on the consumed glycerol, assuming a one-to-sixteen molar ratio (i.e., one mole UP produced per sixteen moles of glycerol consumed). Note that eight moles of succinyl-CoA (derived from 16 moles of glycerol) and eight moles of glycine are needed to generate one mole UP.

4.3 Results

4.3.1 Implementation of the Shemin/C4 Pathway for UP Biosynthesis

We implemented the Shemin/C4 pathway by heterologous expression of *hemA* from *R. sphaeroides* in CPC-SbmΔ*iclR*Δ*sdhA*, resulting in DMA strain. For abundant supply of precursors, the dissimilated carbon flux in the TCA cycle was directed toward succinyl-CoA under aerobic conditions based on the double mutation Δ*iclR*Δ*sdhA* [216]. We also supplemented all the cultures with 2.0 g/L glycine to ensure that the culture performance was not limited by glycine supply. Both DMA and the control strain (CPC-SbmΔ*iclR*Δ*sdhA*) were cultivated in a batch bioreactor under aerobic conditions. Unless otherwise noted, all titers reported here are associated with the last cultivation samples, and we calculated the

percentage yield as the ratio of the actual yield of the produced metabolite to the theoretical yield based on consumed glycerol. Compared to the control strain with relatively fast cell growth and glycerol consumption, DMA had a much lower biomass yield with substantial growth inhibition and acetogenesis (Figures 4.2a and 4.3a). Note that DMA generated much higher levels of 5-ALA, which peaked at 215.0 mg/L (22.6 mg/OD_{600}/L) at 71 h, and PBG, which peaked at 150.0 mg/L (15.9 mg/OD_{600}/L) (Figures 4.2b and 4.3b) with significant pigmentation (Figure 4.2d), producing 55.3 mg/L (0.32% yield) and 41.9 mg/L (0.25% yield) of UP-I and UP-III, respectively (Figure 4.2c). The results suggest that the implemented Shemin/C4 pathway was functional with an enhanced carbon flux redirected from the TCA cycle for UP biosynthesis in DMA.

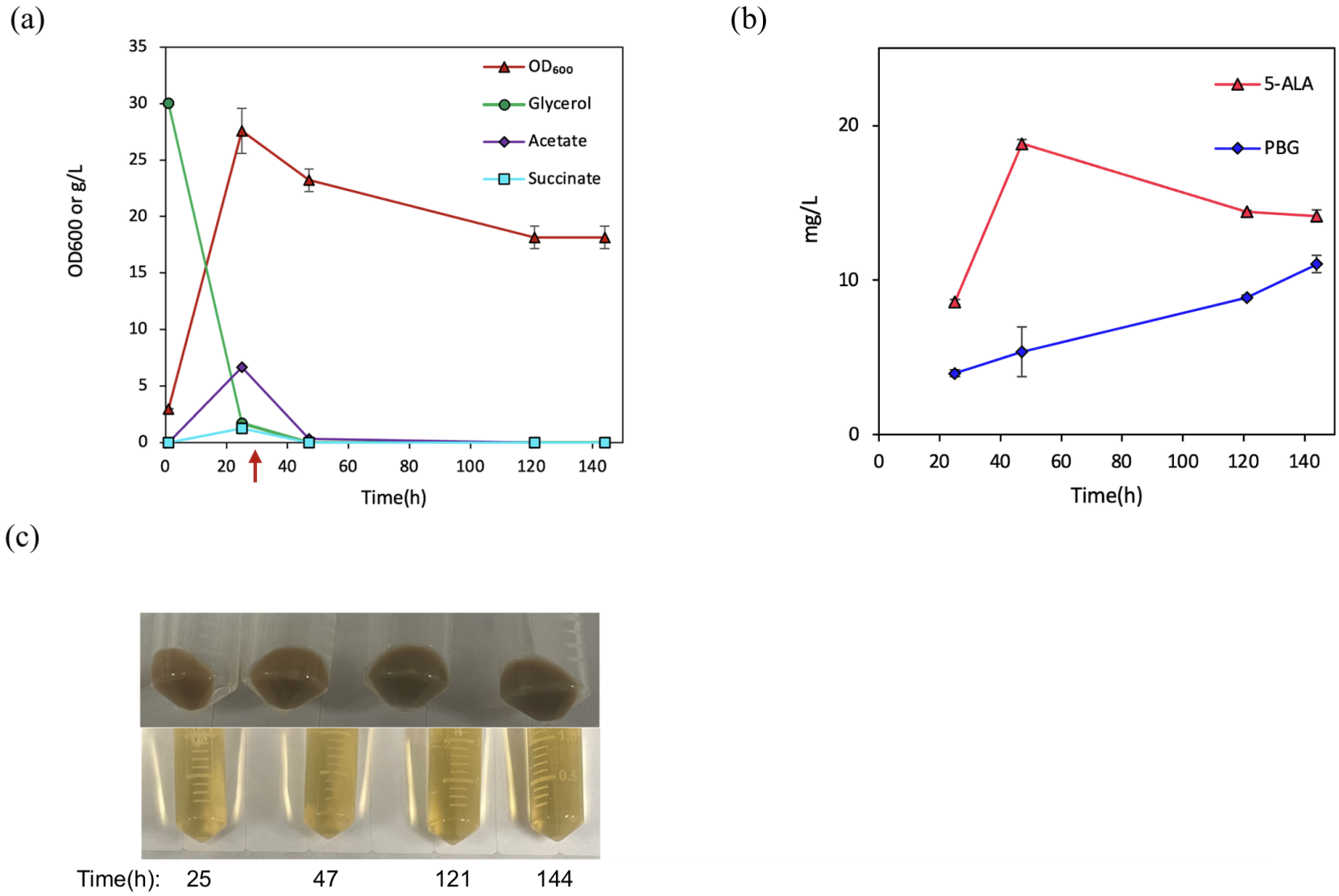


Figure 4.2: Bioreactor cultivation of the control strain CPC-SbmΔ*iclR*Δ*sdhA*. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate production, (**b**) 5-ALA and PBG biosynthesis, and (**c**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$).

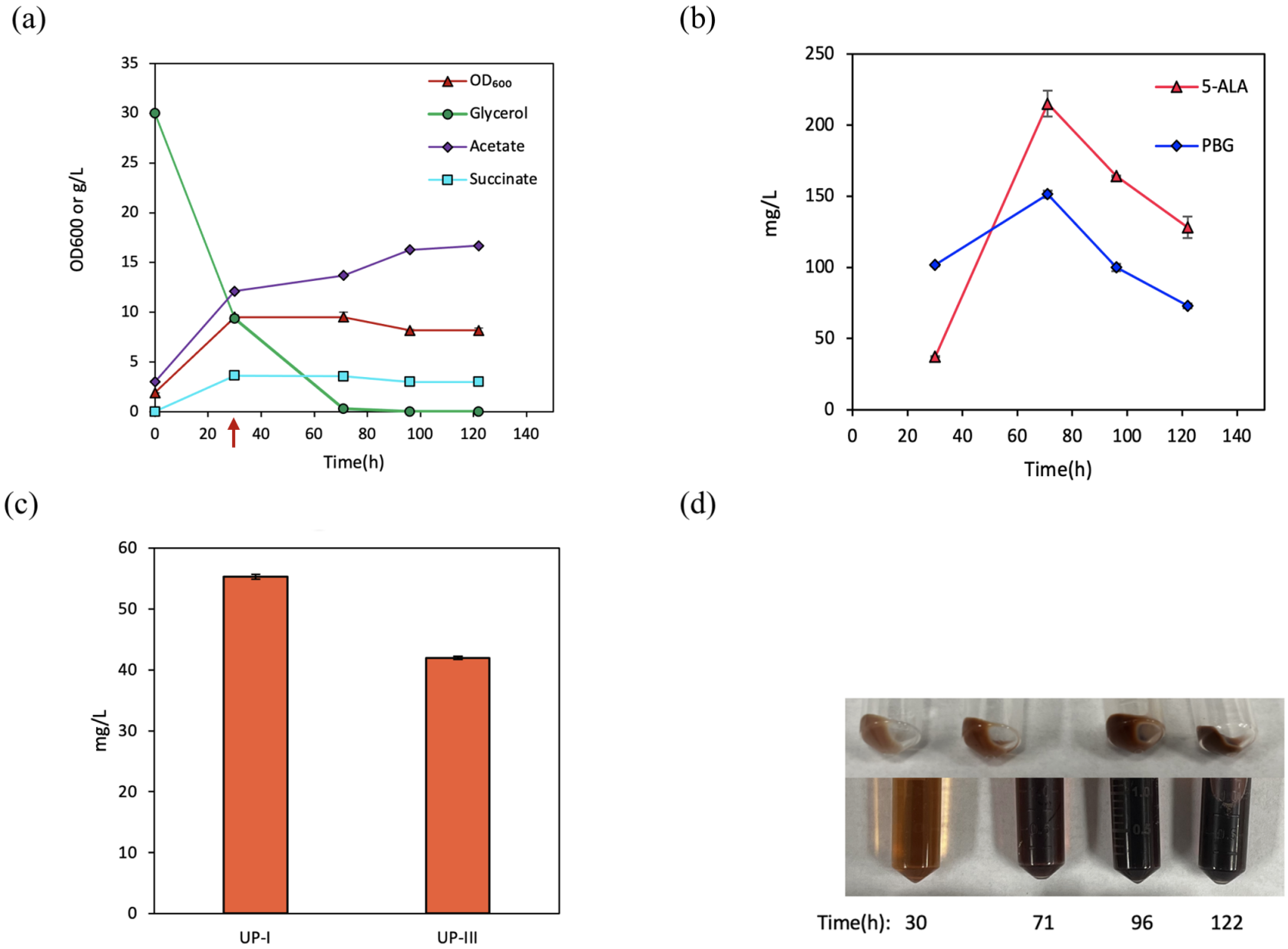


Figure 4.3: Bioreactor cultivation of DMA for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$).

4.3.2 Effects of the Native HemB on UP Biosynthesis

Next, to investigate the effects of native HemB on UP biosynthesis, we derived another UP-producing strain DMB by coexpression of *hemB* from *E. coli* and *hemA* from *R. sphaeroides*. Effective coexpression of *hemA* and *hemB* was achieved from a single operon, which was expressed from the strong *trc* promoter and included strong ribosomal binding sites (RBS) for each gene in plasmid pK-hemAB. Compared to DMA, aerobic cultivation of DMB strain showed a slightly better biomass yield with a slower glycerol consumption and a similar level of acetogenesis (Figure 4.4a). However, DMB produced much higher 5-ALA and PBG, with peak levels of 304.3 mg/L (20.9 mg/OD_{600}/L) and 1995.0 mg/L (193.0 mg/OD_{600}/L), respectively (Figure 4.4b). Additionally, the DMB cultivation had more intensified pigmentation, producing 555.3 mg/L (3.7% yield) and 346.6 mg/L (2.3% yield) of UP-I and UP-III, respectively (Figure 4.4c,d). The results suggest that, in addition to 5-ALA, the formation of PBG could be another key step limiting the overall UP biosynthesis in *E. coli*.

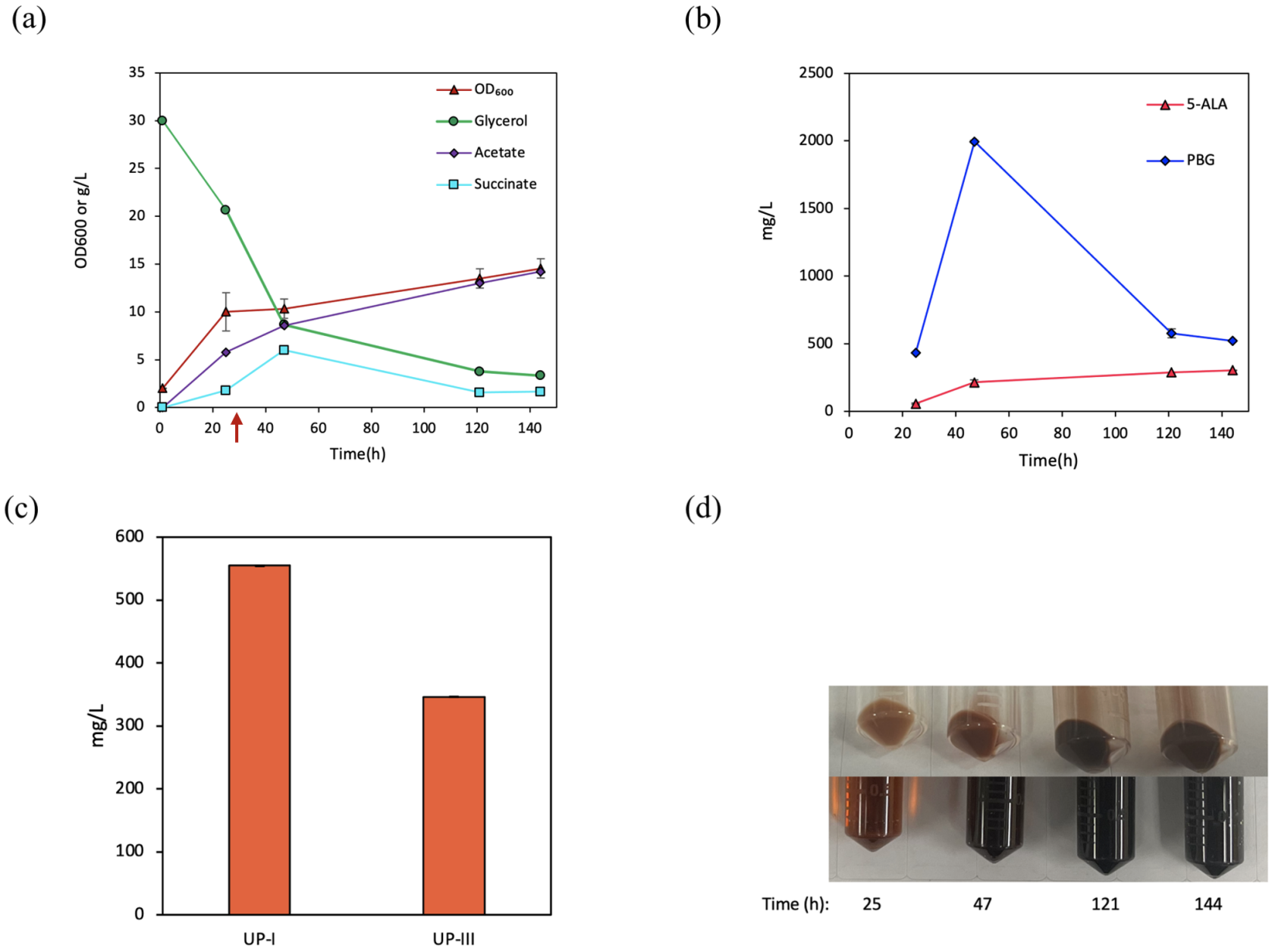


Figure 4.4: Bioreactor cultivation of DMB for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$).

4.3.3 Effects of HemC from *E. coli* and *B. Subtilis* on UP Biosynthesis

Next, we investigated the effects of the native HemC on UP biosynthesis. To do this, we derived another strain, DMC, by coexpression of the native *hemB* and *hemC* from *E. coli* and *hemA* from *R. sphaeroides*. Similar to pK-hemAB, we included all three genes in an operon as *hemABC* regulated by the common *trc* promoter with an individual strong RBS for each gene. Surprisingly, cultivation of the DMC strain exhibited no pigmentation (Figure 4.5c), similar to the control strain (CPC-SbmΔ*iclR*Δ*sdhA*) though the culture grew normally with a rapid glycerol consumption rate and minimal acetogenesis (Figure 4.5a). Notably, the low levels of 5-ALA and PBG for DMC, only up to 14.0 mg/L (0.84 mg/OD_{600}/L) and 11.0 mg/L (0.67 mg/OD_{600}/L), respectively (Figure 4.5b), were also similar to those of the control strain. These findings suggest that the overexpression of the native *hemC* gene potentially disrupted the metabolic activity of the Shemin/C4 pathway.

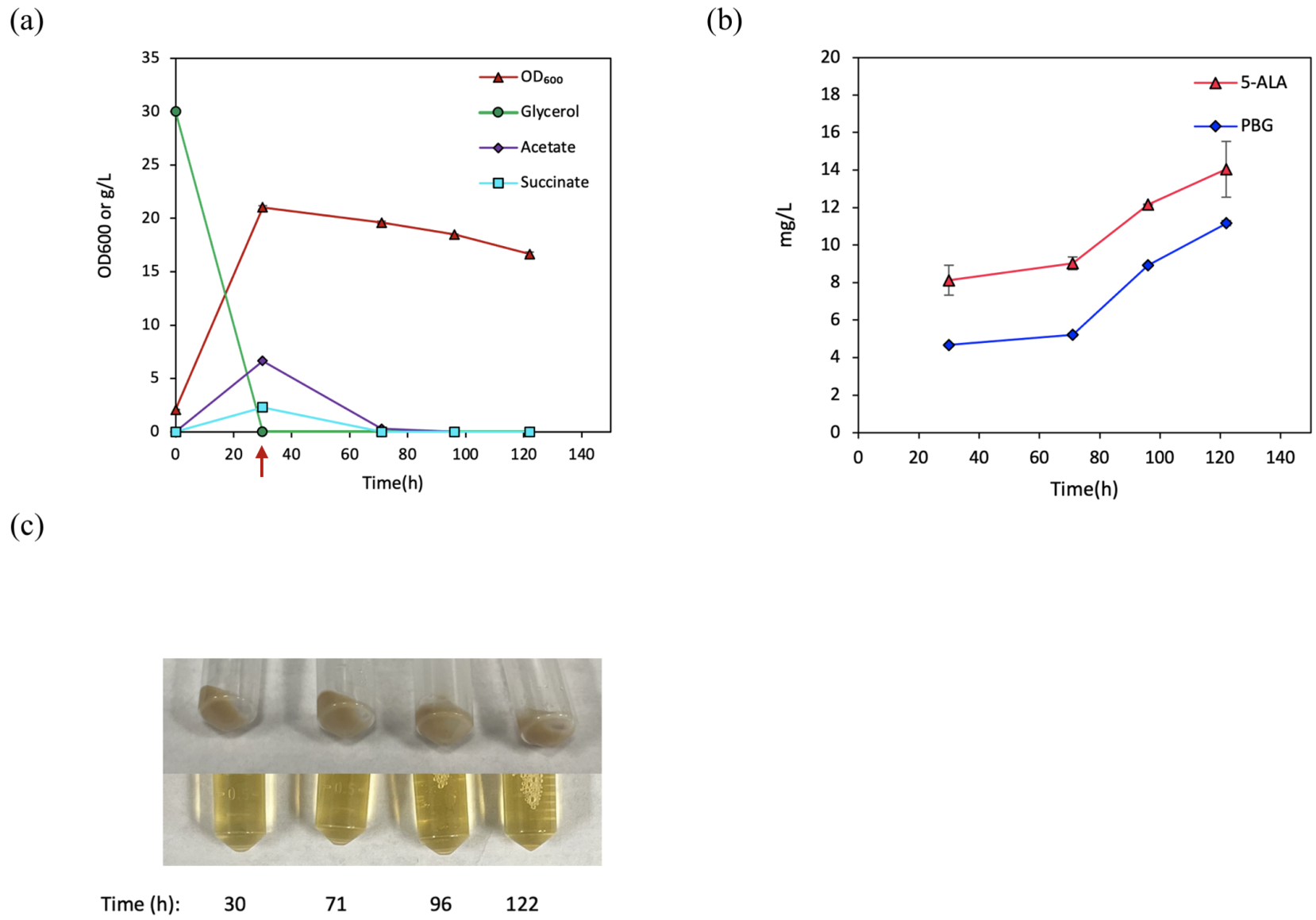


Figure 4.5: Bioreactor cultivation of DMC for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$).

In parallel, we examined the effects of HemC from B. subtilis on UP biosynthesis. To do this, we generated pK-hemABC^{B} with the same plasmid design as pK-hemABC except that the native *hemC* was replaced by *hemC* from *B. subtilis*. Unlike DMC strain, cultivation of the resulting strain DMC^{B} displayed a significant pigmentation (Figure 4.6d) with partially retarded cell growth (similar to DMB) and minimal acetogenesis (Figure 4.6a). Furthermore, DMC^{B} exhibited increased production of 5-ALA and PBG in comparison to DMC, with peak levels being 81.2 mg/L (5.25 mg/OD_{600}/L) and 108.3 mg/L (7.0 mg/OD_{600}/L), respectively (Figure 4.6b). However, the significant PBG accumulation (observed in DMB) was not observed in DMC^{B}, suggesting that the expressed HemC from *B. subtilis* effectively converted PBG to HMB in the early cultivation stage (Figures 4.4b and 4.6b). Nevertheless, DMC^{B} produced UP-I and UP-III significantly lower than DMB, with peak levels of 200.4 mg/L (1.18% yield) and 135.0 mg/L (0.8% yield), respectively (Figure 4.6c). These observations further suggest adverse effects of increasing *hemC* expression on UP biosynthesis.

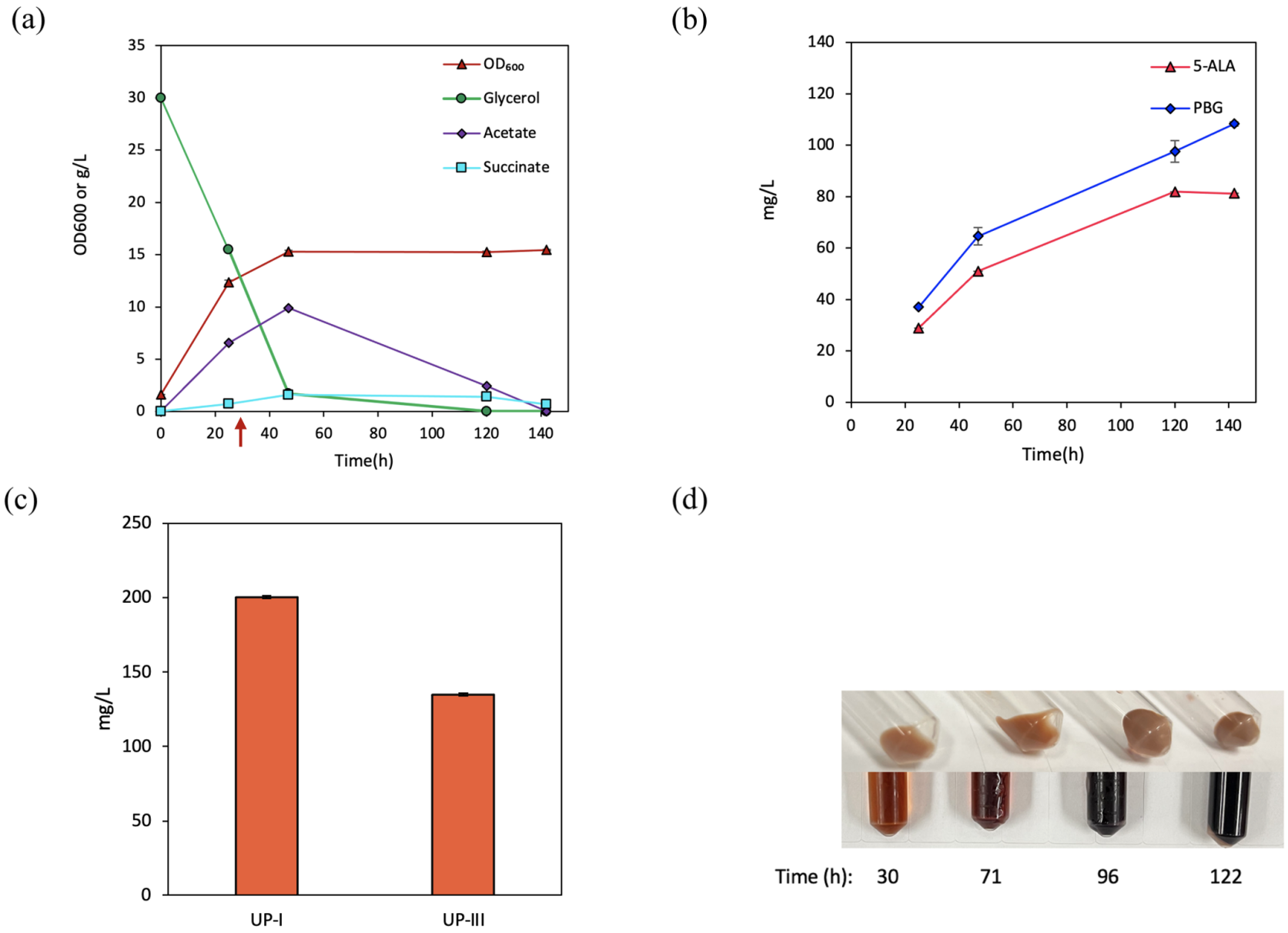


Figure 4.6: Bioreactor cultivation of DMC^{B} for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$).

4.3.4 Effects of the Native HemD on UP Biosynthesis

To investigate the effects of the native HemD on UP biosynthesis, we derived a plasmid pK-hemABD for overexpressing the native *hemB* and *hemD* with *hemA* from *R. sphaeroides* in CPC-SbmΔ*iclR*Δ*sdhA* and CPC-SbmΔ*sdhA*, resulting in DMD and SMD strains, respectively. We excluded *hemC* as its overexpression was detrimental to the Shemin/C4 pathway for porphyrin biosynthesis. DMD had a similar growth pattern and glycerol consumption rate to DMB during aerobic cultivation and significant acetogenesis (Figure 4.7a). Notably, DMD produced high levels of 5-ALA and PBG, peaking at 196.0 mg/L (18.2 mg/OD_{600}/L) and 1117.0 mg/L (83.16 mg/OD_{600}/L) at 57 h, respectively (Figure 4.7b). Similar to DMB with significant pigmentation, DMD produced a total UP of 826.45 mg/L (comprising 538.9 mg/L UP-I and 287.5 mg/L UP-III, at 3.18% and 1.7% yields, respectively) (Figure 4.7c,d). Note that the distribution of UP-I and UP-III in DMD resembled that of DMB, suggesting that *hemD* overexpression did not enhance the metabolic flux toward the classical pathway, suggesting that conversion of HMB via the autoxidation reaction was more effective than the enzymatic reaction by HemD even upon *hemD* overexpression.

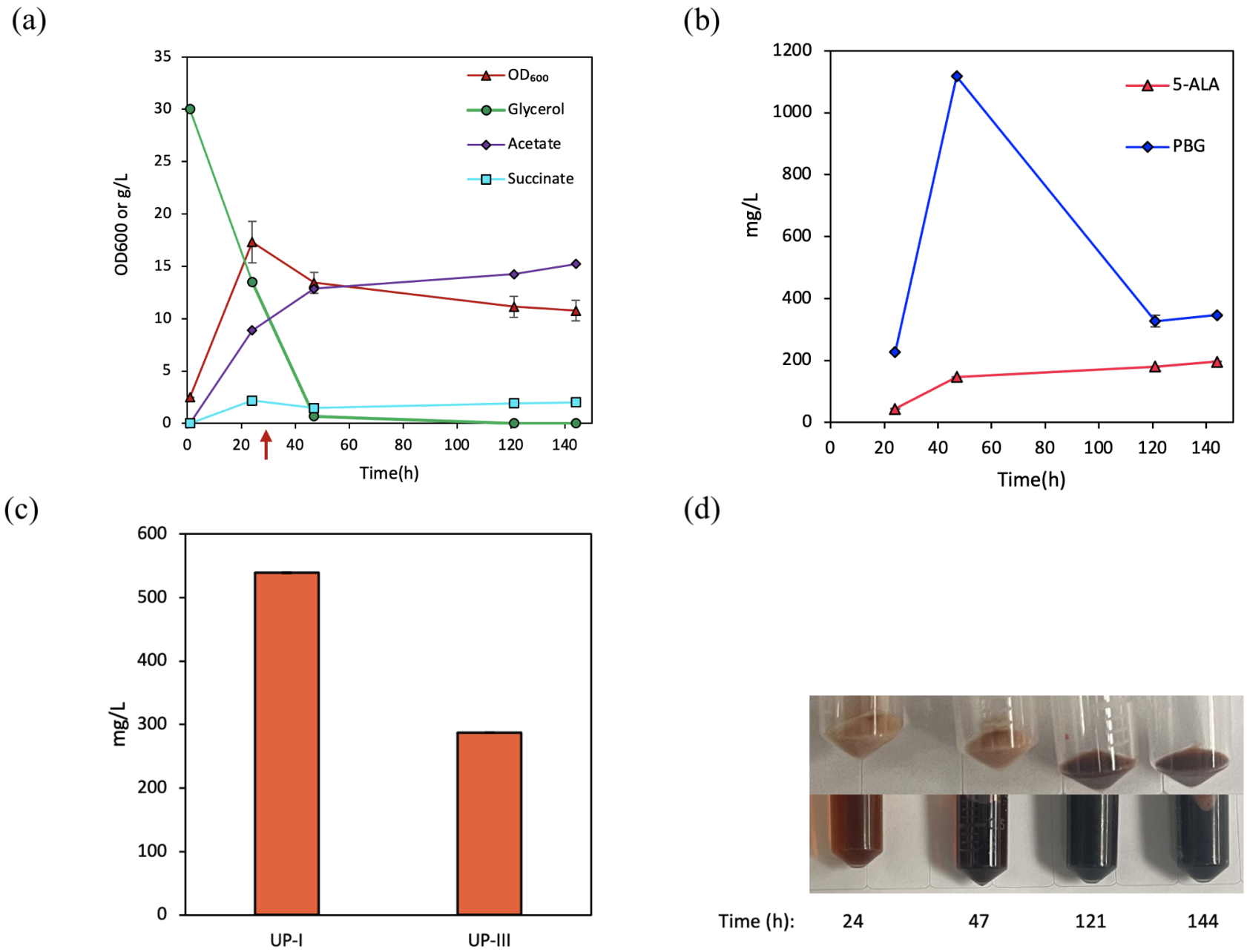


Figure 4.7: Bioreactor cultivation of DMD for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD ($n = 2$).

To investigate the effects of oxygenic condition on the autoxidation (converting HMB to UP-I), we conducted *hemD* overexpression in SMD under lower oxygenic conditions (i.e., AM-II/microaerobic and AM-III/anaerobic). While reducing oxygenic condition for SMD did not seriously affect cell growth and glycerol dissimilation compared to aerobic cultivation of DMD (Figures 4.8a and 4.9a), such environmental manipulation impacted metabolic activity associated with UP biosynthesis. Compared to DMD, SMD only produced 314.9 mg/L UP-I (1.86% yield) and 171.7 mg/L UP-III (1.01% yield) under AM-II, and 37.1 mg/L (0.22% yield) UP-I, 4.78 mg/L (0.02% yield) UP-III, and 19.6 mg/L CP-III (0.14% yield) under AM-III with much lower levels of 5-ALA and PBG (Figures 4.8b, 4.8c, 4.9b, and 4.9c). These results suggest that reducing the oxygenic condition impacted both the autoxidation and enzymatic reaction catalyzed by HemD for UP biosynthesis.

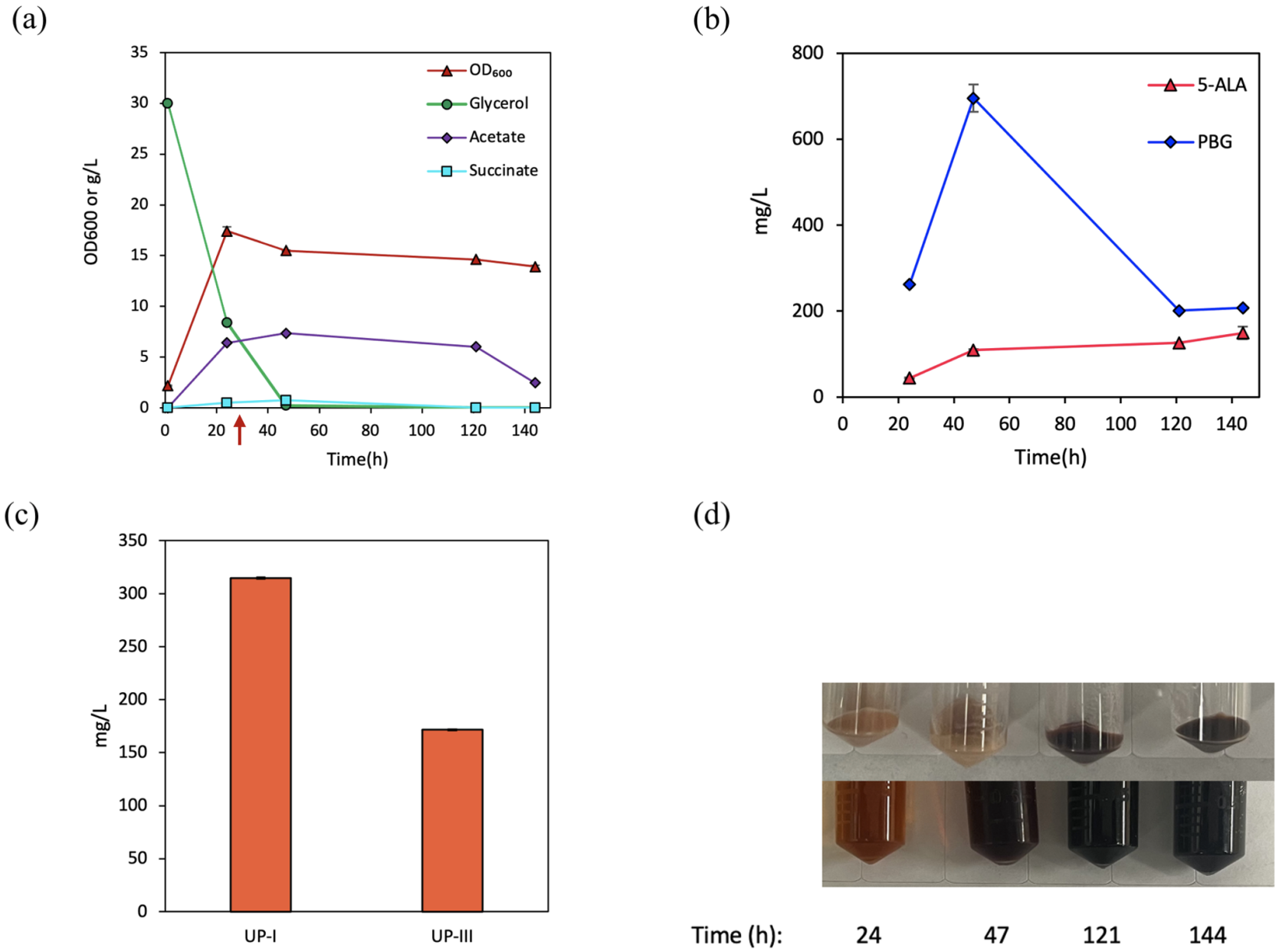


Figure 4.8: Bioreactor cultivation of SMD for porphyrin biosynthesis under AM-II condition. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD (n = 2).

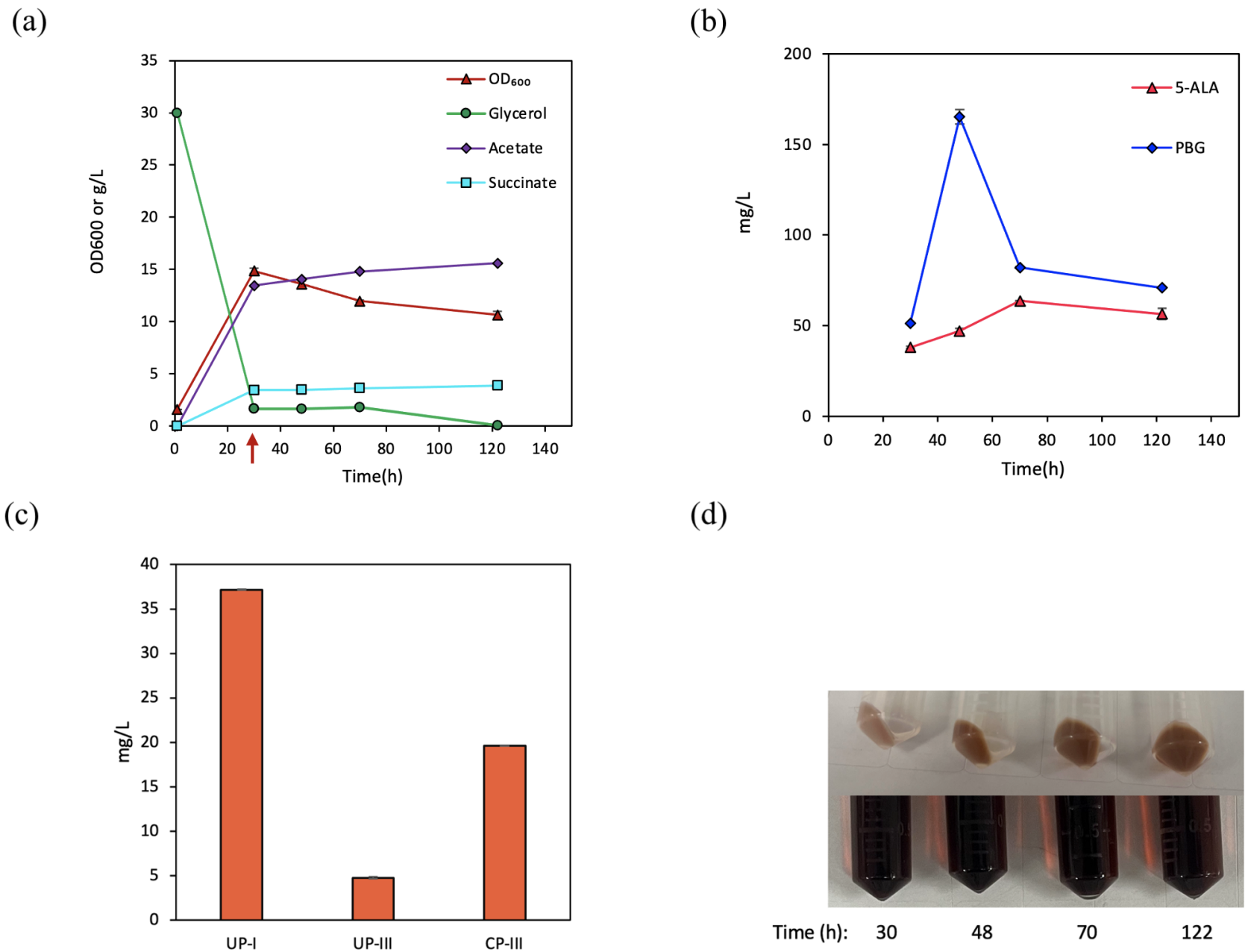


Figure 4.9: Bioreactor cultivation of SMD for porphyrin biosynthesis under AM-III condition. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. The red arrow shows glycine supplementation. All values are reported as means±SD (n = 2).

4.4 Discussion

In this study, we enhanced the biosynthesis of UP in *E. coli* via the Shemin/C4 pathway with a redirected carbon flux from the TCA cycle. Note that most of the heterologously synthesized UP was secreted extracellularly with minimal physiological impact on the producing cells, suggesting the bioprocess feasibility for large-scale production of this porphyrin. Compared to the control strain CPC-SbmΔ*iclR*Δ*sdhA*, heterologous expression of *hemA* resulted in high levels of 5-ALA and PBG in DMA with significant culture pigmentation and UP formation, suggesting that the overall UP biosynthesis in DMA was limited by step(s) after 5-ALA formation under this genetic background. The implementation of the Shemin/C4 pathway also resulted in significant acetogenesis, which reduced the available carbon for UP biosynthesis and impacted cell growth/physiology, similar to previous observations [216, 218]. The significant pigmentation exhibited by the DMA culture with notable UP formation suggests that *hemC* was actively expressed in this strain. Additionally, the higher production of UP-I than UP-III in DMA indicates a preference for the autoxidation of HMB (under aerobic conditions) over the enzymatic conversion by HemD. In addition, the autoxidation for conversion of UPG-I/III to UP-I/III appears to be effective as well such that both UP-I and UP-III can be metabolic end products extracellularly secreted.

Compared to DMA, heterologous expression of the native *hemB* along with *hemA* in DMB resulted in significant PBG accumulation with enhanced pigmentation and UP biosynthesis. The synergistic catalytic effects of HemA and HemB suggest that PBG formation was another potential step limiting UP biosynthesis, in addition to the previously identified one of 5-ALA formation. The significant culture pigmentation in DMB (and DMA) suggests that the genomic *hemC* was actively expressed and the conversion from PBG to HMB was rather effective. Similar to DMA, the relatively higher titer of UP-I than UP-III in DMB suggests that the autoxidation was favored over the enzymatic conversion by HemD, presumably due to aerobic cultivation conditions.

The HemC-mediated conversion of PBG to HMB, which is a critical step to form the initial tetrapyrrole structure, was effective in DMA and DMB based on the level of culture pigmentation, implying that the native genomic *hemC* was actively expressed. It was documented that HemC normally remains stable and abundant in *E. coli* [91]. In a previous study, the expression of *hemC* had to be repressed (via CRISPRi) in order to overproduce PBG, also suggesting a relatively high level of expression of native genomic *hemC* in *E. coli* [218]. Nevertheless, it was surprising to see that heterologous expression of the native *hemC* completely disrupted the metabolic activity of the Shemin/C4 pathway in DMC. Note that the lack of culture pigmentation and low levels of 5-ALA and PBG

(i.e., less than 20.0 mg/L), which were comparable to those of the control strain, suggest that HemA was somehow inactivated (presumably by overexpression of the native *hemC*) in DMC. Though the Shemin/C4 pathway remained active upon heterologous expression of *B. subtilis hemC* in DMC^B, UP biosynthesis was negatively impacted. Further research is required to identify the underlying cause of these *hemC* effects.

Our results showed that conversion of HMB to UPG-I via autoxidation was much more effective than the enzymatic conversion to UPG-III by HemD. Overexpression of *hemD* neither increased the overall UP biosynthesis nor the selectivity of HBM conversion to UPG-III. The results suggest that the overall UP biosynthesis was not limited by the step of HMB conversion. Reducing the oxygenic condition, which presumably retards autoxidation, only decreased the overall UP biosynthesis without significantly affecting the selectivity of HBM conversion to UPG-III. Such effective autoxidation for HBM conversion to UPG-I presented major challenges for biosynthesis of downstream porphyrins derived from UPG-III, such as CP, PP-IX, and heme, in *E. coli*. Interestingly, HMB conversion to UPG-III appeared to be enhanced by overexpression of *hemE* (data not shown), even without coexpression of *hemD*. While conversion of UPG-III to UP-III via autoxidation was more effective than conversion to CPG-III via HemE under aerobic conditions, reducing the oxygenic conditions appeared to reverse such conversion preference, leading to formation of CP-III more than UP-III.

Metabolically, structural linking of four molecules of PBG (to form the linear tetrapyrrole of HMB) and subsequent cyclization of HMB (to form either UPG-I or UPG-III) are critical for porphyrin biosynthesis, which can be observable by culture pigmentation. Apparently, both the two conversion steps were effective in *E. coli* based on the levels of culture pigmentation and UP biosynthesis in DMA and DMB, suggesting that genomic expression of *hemC* and *hemD* were abundant and active even under the conditions for enhanced UP biosynthesis. Heterologous overexpression of *hemD* did not enhance the overall UP biosynthesis, whereas heterologous overexpression of *hemC* completely disrupted the metabolic activity of the Shemin/C4 pathway (specifically the HemA-catalytical step for 5-ALA formation) with no UP biosynthesis. Our results demonstrated the importance of optimal tuning of the expression of various enzymes involved in the Shemin/C4 pathway for enhanced UP biosynthesis.

Notably, glycerol was added 30 hours after the start of cultivation to limit its utilization for biomass formation or cellular maintenance, which are more prominent during earlier growth phases. As porphyrin biosynthesis predominantly occurs in the late stationary phase, this delayed supplementation promotes its selective use for product formation.

In addition, the genome engineering toolkit introduced in Chapter 3 was employed in two distinct scenarios to enhance porphyrin biosynthesis in *E. coli*. In the first scenario,

We knocked out *iclR* and *sdhA* genes in the ST002 strain and subsequently integrated either the *hemAB* or *hemABC* operons into its genome. These engineered strains, named TS-AB and TS-ABC (Table S4.2), aimed to achieve stable expression of biosynthetic genes while reducing the metabolic burden associated with plasmid maintenance. However, UP production in these strains was minimal, indicating that genomic expression was insufficient to drive high-level porphyrin biosynthesis under the tested conditions (Figure S4.1).

In the second scenario, we shifted focus to *E. coli* BL21(DE3) strains, which are known for their robust growth and adaptability for industrial applications. Using the genome engineering toolkit, we created two knockout variants of BL21(DE3): *iclR* knockout (BLS) and *iclR* and *sdhA* double knockout (BLD). These strains were then transformed with the pK-HemAB plasmid to develop three experimental strains: BL-AB, BLS-AB, and BLD-AB (Table S4.2). Despite leveraging the plasmid-based system, the BL21-derived strains exhibited lower UP titers than the CPC-Sbm strains harboring similar plasmid. This discrepancy may be attributed to differences in metabolic flux, precursor availability, or regulatory mechanisms between the strain backgrounds (Figure S4.2 and S4.3).

These results guided the strategic decision to continue with plasmid-based systems in the CPC-Sbm strain for subsequent chapters, as they consistently yielded higher production levels. To provide a comprehensive comparison, supplementary figures (Figure S4.1- S4.3) present growth profiles, metabolite production, and porphyrin titers across the TS-AB/TS-ABC and BL21(DE3)-derived strains. These findings underscore the importance of strain selection and metabolic context in achieving efficient porphyrin biosynthesis.

Our study presents valuable insights into UP biosynthesis via the Shemin/C4 pathway in *E. coli*, demonstrating the feasibility of large-scale production. However, several areas offer potential for future research: Investigating the unexpected effects of HemC overexpression on HemA catalysis could provide valuable understanding of the heme pathway. Additionally, exploring strategies to avoid the autooxidation pathway might prevent the formation of unexpected byproducts, enhancing the efficiency of the biosynthesis process. Further investigation into the specific preferences of autoxidation versus enzymatic conversion pathways for HMB conversion would facilitate harnessing these pathways for targeted porphyrin biosynthesis.

Chapter 5

High-Level Bio-Based Production of Coproporphyrin in *Escherichia coli*

5.1 Introduction

Heme is an essential compound in various hemoproteins, such as hemoglobin, myoglobin, and cytochromes, and plays a critical role in oxygen transportation, energy generation, and several enzymatic activities [226]. In bacteria, heme can be naturally synthesized via two distinct pathways: i.e., protoporphyrin-dependent (PPD) and coproporphyrin-dependent (CPD) pathways [227], with the PPD pathway being predominantly present in Gram-negative bacteria and the CPD pathway in Gram-positive bacteria [228]. In both pathways, heme biosynthesis starts with the formation of 5-aminolevulinic acid (5-ALA) via either the C4 or C5 pathway [91] (Figure 5.1). Subsequently, the enzyme porphobilinogen synthase (HemB) mediates the fusion of two 5-ALA molecules, yielding porphobilinogen (PBG). Then, PBG deaminase (HemC) interconnects and cyclizes four PBG molecules to form a linear tetrapyrrole, hydroxymethylbilane (HMB) [213]. At this metabolic node, the heme biosynthetic pathway diverges to form either of the two stereoisomers, i.e., uroporphyrinogen I (UPG-I) via autoxidation or uroporphyrinogen III (UPG-III) via uroporphyrinogen III synthase (HemD). Uroporphyrinogen decarboxylase (HemE) decarboxylates the four acetate residues of UPG-I to form coproporphyrinogen I (CPG-I), which is spontaneously oxidized to form coproporphyrin I (CP-I) as a metabolic end product [229]. Similarly, HemE mediates the conversion of UPG-III to coproporphyrinogen III (CPG-III), which is enzymatically converted to coproporphyrin III (CP-III) via coproporphyrinogen III oxidase (HemY) or porphyrinogen peroxidase (YfeX). Note that CPG-III represents the final shared precursor in the PPD and CPD pathways, both of which lead to the formation of

heme via its respective precursors of protoporphyrin IX (PP-IX) and CP-III.

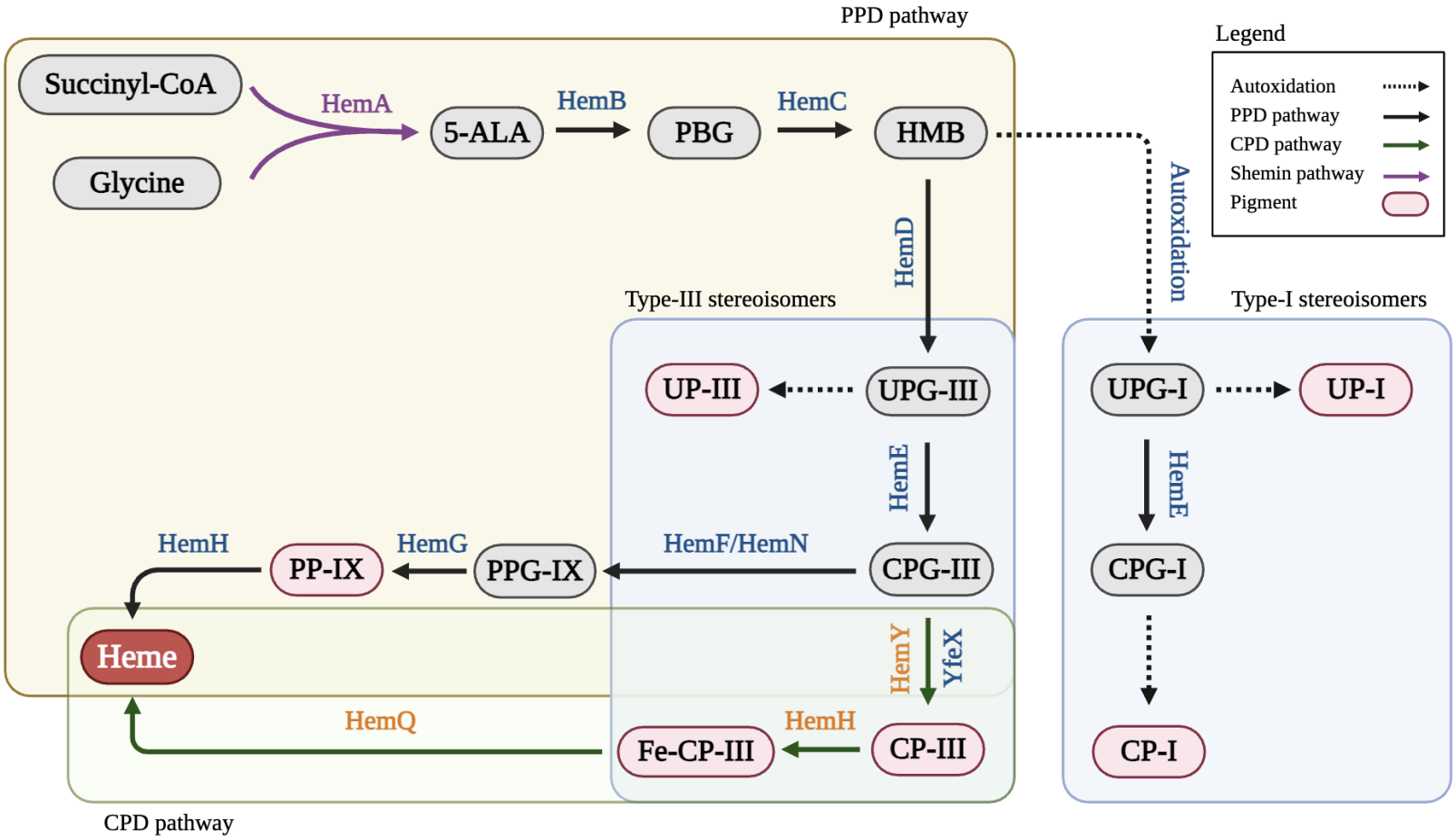


Figure 5.1: The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. 5-ALA, 5-aminolevulinic acid; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; Fe-CP-III, coproheme III; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin/coproporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HemQ, coproheme decarboxylase; HemY, coproporphyrinogen III oxidase; HMB, Hydroxymethylbilane; PBG, porphobilinogen; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase.

Porphyrin production in microbial hosts presents a sustainable and economically feasible alternative to traditional chemical synthesis methods, particularly in light of the growing demand for bio-based production strategies [230]. Modulating the metabolic capabilities of microorganisms, notably the most common bacterial host, *Escherichia coli*, facilitates the efficient and sustainable synthesis of porphyrins [231]. The well-characterized genetics and versatile metabolic engineering tools enable precise regulation of biosynthetic pathways in *E. coli* [232]. Given the commercial value and applicability of heme, its bio-based production has been explored [233]. While *E. coli* can be extensively engineered for enhanced heme biosynthesis, the overproduced heme is hard to secrete and tends to accumulate intracellularly, resulting in physiological toxicity to the production host and, therefore, limited production yield [38, 234]. Such technical limitation prompted us to look for alternative porphyrins that could be extracellularly secreted for overproduction in *E. coli*.

Coproporphyrin (CP), specifically CP-III, is a critical porphyrin intermediate and precursor for heme biosynthesis via the CPD pathway [235]. Beyond its significance in heme biosynthesis, CP holds noteworthy applications across diverse fields. Derivatives of CP are pivotal in clinical diagnostics such as the diagnosis of porphyria [236]. CP can also serve as a diagnostic biomarker [237, 238]. Moreover, due to unique optical properties and low cytotoxicity, CP has applications beyond healthcare, including use as fluorescent probes, in photodynamic therapy, and as bioimaging agents [239, 240, 241].

In *E. coli*, the PPD pathway is primarily used for heme biosynthesis (Figure 5.1). However, the simultaneous presence of enzymes and pathways diverging from the PPD pathway offers potential opportunities for the biosynthesis of porphyrins other than heme. Given the above-mentioned technical limitation associated with intracellular accumulation of heme, we herein focused on CP for its bio-based production in *E. coli*. While minimal CP production was observed in Chapter 4, it is noteworthy that targeted bio-based CP production in *E. coli* has yet to be reported. Note that the overproduced CP can be further processed, either chemically or biochemically, to form heme. In this study, we first heterologously implemented the Shemin/C4 pathway in an *E. coli* host strain with an expanded intracellular pool of succinyl-CoA, a key precursor for porphyrin biosynthesis via the Shemin/C4 pathway. Considering the complexity of the pathway and enzymes involved in CP biosynthesis, differential expression strategies for modulation of the levels of key enzymes, particularly those governing the sequential conversion of 5-ALA to CP, were applied as a strategic metabolic approach for enhanced CP biosynthesis in *E. coli*. Selective alteration of key enzyme levels enabled fine-tuning of the associated metabolic flux in the CP biosynthetic pathway, optimal substrate channeling, and minimal accumulation of intermediates or byproducts. Moreover, by applying differential expression strategies,

we can minimize the metabolic burden imposed by enzyme overexpression, ensuring efficient utilization of cellular resources to increase product yield [242]. By exploring diverse combinations of specific pathway genes for heterologous expression, we developed a flexible platform to engineer *E. coli* strains for high-level CP biosynthesis under aerobic conditions using glycerol as the primary carbon source.

5.2 Materials and Methods

5.2.1 Bacterial Strains and Plasmids

The bacterial strains and plasmids used in this study are provided in Table 5.1, and the primer sequences are listed in Table S5.1. Taq DNA polymerase was obtained from New England Biolabs (Ipswich, MA, USA). Genomic DNA of the bacterial cells was extracted using the Qiagen Blood & Tissue DNA Isolation Kit (Hilden, Germany). Plasmids were purified using the Qiagen Miniprep kit. All host cells in this study were derived from CPC-Sbm through the introduction of mutation(s) in the *sdhA* and/or *iclR* genes [200]. It is noteworthy that CPC-Sbm originated from BW25113 with the *ldhA* gene inactivated [184]. For molecular cloning, *E. coli* HI-Control 10G (Lucigen, Middleton, WI, USA) was used. DNA sequencing was performed by Plasmidsaurus (Eugene, OR, USA). All plasmids were constructed through Gibson assembly [181]. All oligonucleotides were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA).

Table 5.1: Strains and plasmids used in this study.

Name	**Description or Relevant Genotype**	**Source**
Host strains		
HI-Control 10G	*mcrA*, Δ(*mrr-hsdRMS-mcrBC*), *endA1*, *recA1*, *ϕ80dlacZ*ΔM15, Δ*lacX74*, *araD139*, Δ(*ara leu*)*7697*, *galU*, *galK*, *rpsL* (Str^R), *nupG*, λ^-, *tonA*, Mini-F $lacI^{q1}$ ($Gent^R$)	Lucigen
MG1655	K-12; F^-, λ^-, *rph-1*	Lab stock
Bacillus Subtilis 168	Wild type	Lab stock
CPC-SbmΔ*iclR*Δ*sdhA*	F^-, Δ*(araD-araB)567*, Δ*lacZ4787(::rrnB-3)*, λ^-, *rph-1*, Δ*(rhaD-rhaB)568*, *hsdR514*, Δ*ldhA*, P_{trc}*::sbm* (i.e., with the FRT-P_{trc} cassette replacing the 204-bp upstream of the Sbm operon), Δ*iclR*, Δ*sdhA*	[200]
BA001	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABD	Chapter 4

BA002	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABD-E	This study
BA003	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemAB-DE	This study
BA004	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABD-EY$_{B.s}$	This study
BA005	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABD-EY	This study
BA006	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemAB-E	This study
Plasmids		
pK-hemABCD	p15A ori, Km^R, P$_{trc}$*::hemABCD*	Chapter 4
pK-hemABCD-E	p15A ori, Km^R, P$_{trc}$*::hemABCD*-P$_{gracmax}$*::hemE*	This study
pK-hemABD	p15A ori, Km^R, P$_{trc}$*::hemABD*	Chapter 4
pK-hemABD-E	p15A ori, Km^R, P$_{trc}$*::hemABD*-P$_{gracmax}$*::hemE*	This study
pK-hemAB-DE	p15A ori, Km^R, P$_{trc}$::*hemAB*-P$_{gracmax}$*::hemDE*	This study
pK-hemABD-EY$_{B.s}$	p15A ori, Km^R, P$_{trc}$::*hemABD*-P$_{gracmax}$*::hemEY -hemY* from *B. subtilis 168*-	This study
pK-hemABD-EY	p15A ori, Km^R, P$_{trc}$::*hemABD*-P$_{gracmax}$*::hemEY -hemY* from MG1655-	This study
pK-hemAB-E	p15A ori, Km^R, P$_{trc}$::*hemAB*-P$_{gracmax}$*::hemE*	This study

Plasmid pK-hemABCD was used from Chapter 4 to serve as the template for the remaining plasmids. More information is available in the supplementary materials. Plasmid pK-hemABCD-E was also constructed to serve as the template for the other plasmids containing two operons. pK-hemABCD-E was constructed by amplifying *hemA* along with the backbone, consisting of p15A ori and *trc* promoter with a kanamycin resistance marker, using primers P002/P011 and amplifying *hemB*, *hemC*, and *hemD* using primers P003/P012 and pK-hemABCD as a template. The *hemE* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P013/P014. The *grac*max promoter was amplified from a previous lab-made plasmid using primers P015/P016. Subsequently, these four fragments were Gibson-assembled to form pK-hemABCD-E. For effective coexpression, *hemA*, *hemB, hemC*, and *hemD* were aligned to form an operon *hemA/B/C/D* regulated by a common strong *trc* promoter and *hemE* was included on the second operon regulated by the *grac*max promoter with an individual strong ribosomal binding site (RBS) for each gene.

pK-hemABD was used from Chapter 4. pK-hemABD-E was constructed by amplifying *hemA* and *hemB* using primers P001/P004 and amplifying *hemD* and *hemE* along with the backbone using primers P017/P018, using pK-hemABCD-E as a template. These two fragments were Gibson-assembled to form pK-hemABD-E. For effective coexpression, *hemA*, *hemB*, and *hemD* were aligned to form an operon *hemA/B/D* regulated by a common strong *trc* promoter and *hemD* was included on the second operon regulated by the *grac*max promoter with an individual strong RBS for each gene.

pK-hemAB-DE was constructed by amplifying *hemA* and *hemB* using primers P001/P019 and pK-hemABCD as a template. *grac*max promoter and *hemD* were amplified from pK-hemABD-E separately using primer sets P020/P021 and P022/P023, respectively. *hemE* was amplified along with the backbone, using primers P018/P024 and pK-hemABD-E as a template. These four fragments were Gibson-assembled to form pK-hemAB-DE. *hemA* and *hemB* were aligned to form an operon *hemA/B* regulated by a common strong *trc* promoter and *hemD/E* were included on the second operon regulated by the *grac*max promoter with an individual strong RBS for each gene.

pK-hemABD-$EY_{B.s}$ was constructed by amplifying *hemA*, *hemB*, *hemD*, and *hemE* using primers P001/P025 and pK-hemABD-E as a template. The *hemY* gene was amplified from the genomic DNA of *Bacillus subtilis* 168 using primers P026/P027. The backbone was amplified from pK-hemABD-E using primers P028/P029. These three fragments were Gibson-assembled to form pK-hemABD-$EY_{B.s}$. *hemA, hemB*, and *hemD* were aligned to form an operon *hemA/B/D* regulated by a common strong *trc* promoter and *hemE/Y* were included on the second operon regulated by the *grac*max promoter with an individual strong RBS for each gene.

pK-hemABD-EY was constructed by amplifying *hemA*, *hemB, hemD*, and *hemE* using primers P001/P025 and pK-hemABD-E as template. The *hemY* gene was amplified from the genomic DNA of *Escherichia coli* MG1655 using primers P030/P031. The backbone was amplified from pK-hemABD-E using primers P032/P029. These three fragments were Gibson-assembled to form pK-hemABD-EY. *hemA, hemB*, and *hemD* were aligned to form an operon *hemA/B/D* regulated by a common strong *trc* promoter and *hemE/ Y* were included on the second operon regulated by the *grac*max promoter with an individual strong RBS for each gene.

pK-hemAB-E was constructed by amplifying *hemA* and *hemB*, and amplifying *grac*max promoter and *hemE* along with the backbone, using primer sets P019/P033 and P020/P034 with pK-hemABCD-E as a template. These two fragments were Gibson-assembled to form pK-hemAB-E. *hemA* and *hemB* were aligned to form an operon *hemA/B* regulated by a common strong *trc* promoter and *hemE* was included on the second operon regulated by the *grac*max promoter with an individual strong RBS for each gene.

5.2.2 Media and Bacterial Cell Cultivation

All medium components were obtained from Sigma-Aldrich Co. (St. Louis, MO, USA), with the exception of yeast extract and tryptone, which were obtained from BD Diagnostic Systems (Franklin Lakes). *E. coli* strains were preserved as glycerol stocks at -80 °C

and streaked on lysogeny broth (LB; 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl) agar plates, then incubated at 37 °C for 14–16 h. Bioreactor cultivations were conducted following the methodology outlined in Chapter 4. Briefly, for bioreactor cultivation, individual single colonies were picked from LB plates to inoculate 12 mL of super broth (SB) medium (32 g/L tryptone, 20 g/L yeast extract, and 5 g/L NaCl) in a 125 mL conical flask. The culture was incubated at 37 °C and 280 revolutions per min (rpm) using a rotary shaker (New Brunswick Scientific) for 4–6 h and subsequently used as a starter culture to inoculate 220 mL of SB medium at a 2% (vol/vol) concentration in a 1 L conical flask. The seed culture was incubated at 37 °C and 280 rpm for 14–16 h. Cells were harvested by centrifugation at 4500× g and 20 °C for 8 min and resuspended in 40 mL of fresh SB medium. The resuspended culture was used to inoculate a stirred tank bioreactor (BioFlo 115, Eppendorf AG, Hamburg, Germany) with a working volume of 0.8 L at 37 °C and 430 rpm. The semi-defined production medium in the batch bioreactor contained 30 g/L glycerol, 0.23 g/L K_2HPO_4, 0.51 g/L NH_4Cl, 49.8 mg/L $MgCl_2$, 48.1 mg/L K_2SO_4, 1.52 mg/L $FeSO_4$, 0.055 mg/L $CaCl_2$, 2.93 g/L NaCl, 0.72 g/L tricine, 10 g/L yeast extract, 10 mM $NaHCO_3$, and 1 mL/L trace elements (2.86 g/L H_3BO_3, 1.81 g/L $MnCl_2{\cdot}4H_2O$, 0.222 g/L $ZnSO_4{\cdot}7H_2O$, 0.39 g/L $Na_2MoO_4{\cdot}2H_2O$, 79 μg/L $CuSO_4{\cdot}5H_2O$, 49.4 μg/L $Co(NO_3)_2{\cdot}6H_2O$); (Neidhardt, Bloch, and Smith, 1974), supplemented with 0.05 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). To avoid glycine limiting during the cultivation, 2 g of glycine was supplemented into the bioreactor 30 h post inoculation. The aerobic condition was maintained by continuously purging the air into the bulk culture at 1 volume of air per volume of liquid per min (vvm). The pH of the bioreactor culture was maintained at 7.0 ± 0.1 using 3 M NH_4OH and 3 M H_3PO_4.

5.2.3 Analysis

To measure cell density at OD600, all culture samples were washed once and then appropriately diluted with a 0.15 M saline solution and analyzed using a spectrophotometer (GENESYS^{TM} 40/50 Vis/UV-Vis, Thermo Fisher Scientific Inc., Waltham, MA, USA). To prepare a cell-free medium, culture samples were centrifuged at 17,000× g for 1 min, followed by filter sterilization using a 0.2 μm syringe filter. Extracellular metabolites and glycerol were analyzed by high-performance liquid chromatography (HPLC; LC-10AT; Shimadzu, Kyoto, Japan) with a refractive index detector (RID; RID-10A; Shimadzu) and a chromatographic column (Aminex HPX-87H; Bio-Rad Laboratories, CA, USA). The HPLC column temperature was maintained at 35 °C and the mobile phase was 5 mM H_2SO_4 (pH 2) running at 0.6 mL/min. The RID signal was analyzed by data processing software (Clarity Lite v. 7.4.1.88 software, DataApex, Prague, Czech Republic). The levels of 5-ALA and PBG in the cell-free medium were assessed using a modified Ehrlich's reagent

assay [225]. Porphyrins were analyzed using an HPLC (2690 separation module, Waters™, Milford, MA, USA) equipped with a photodiode array (2996 PDA detector, Waters™, Milford, MA, USA) detector and a chromatographic column (Chromolith® HighResolution RP-18 endcapped, Supelco, Darmstadt, Germany). The UV absorbance was detected at 400 nm and data were processed using Empower 3 software (Waters™, Milford, MA, USA). A mobile-phase system previously outlined [11] was used with minor modifications. Briefly, the mobile phase consisted of two solvent mixtures: solvent A (7.7 g/L ammonium acetate, 125 mL/L acetonitrile, adjusted to pH 5.17 with glacial acetic acid) and solvent B (a methanol-glacial acetic acid mixture with a volume ratio at 10:1). The flow rate of the mobile phase was 1 mL/min. An isocratic elution step was used with 100% solvent A for 5 min, followed by a linear gradient from 0 to 100% solvent B over 30 min, and further followed by 100% solvent B for 5 min. The column temperature was maintained at 45 °C. The CP percentage yield was determined by calculating the ratio of the produced CP to the maximum theoretical CP production based on the consumed glycerol. This calculation assumed a one-to-sixteen molar ratio, meaning 1 mole of CP was produced for every 16 moles of glycerol consumed.

5.3 Results

5.3.1 Effects of the Native HemD for CP Biosynthesis

Results of Chapter 4 showed that heterologous coexpression of *hemA* from *Rhodobacter sphaeroides* and the native *hemB* can increase the metabolic flux into the Shemin/C4 pathway, resulting in enhanced UP biosynthesis with significant culture pigmentation. On the other hand, the inclusion of the native *hemC* for coexpression was detrimental to porphyrin biosynthesis and even completely blocked the culture pigmentation. To investigate the effects of the native HemD on CP biosynthesis, we derived a plasmid pK-hemABD by heterologous coexpression of *hemA* from *R. sphaeroides* and *hemB/D* from *E. coli* in CPC-SbmΔ*iclR*Δ*sdhA*, resulting in BA001. For effective coexpression, the three genes of *hemA*, *hemB*, and *hemD* were cloned to form an operon regulated by the strong *trc* promoter and each gene had an individual strong RBS. The increased *hemD* expression could presumably direct more carbon flux toward the classical heme-biosynthetic pathway by enhancing the specific conversion of hydroxymethylbilane (HMB) to Uroporphyrinogen III (UPG-III) (Figure 5.1), rather than Uroporphyrinogen I (UPG-I) which occurs via autoxidation. We also supplemented all the cultures with 2 g/L glycine to ensure that porphyrin biosynthesis was not limited by glycine availability. For an abundant supply of

succinyl-CoA for porphyrin biosynthesis, the production host with a double mutation of $\Delta iclR\Delta sdhA$ was used and all CP-producing strains were cultivated aerobically [216].

Cultivation of BA001 showed normal cell growth with rapid glycerol consumption and significant acetogenesis (Figure 5.2a). Notably, BA001 produced high levels of 5-ALA and PBG, peaking at 257.7 mg/L (19.6 mg/OD_{600}/L) and 1472 mg/L (112 mg/OD_{600}/L) at 57 h, respectively (Figure 5.2b). This strain also exhibited significant pigmentation, producing 430.7 mg/L (2.55% yield) UP-I and 363 mg/L (2.2% yield) UP-III, respectively (Figure 5.2c,d). The results suggest that overexpression of *hemA/B/D* primarily led to UP production, but hardly any CP.

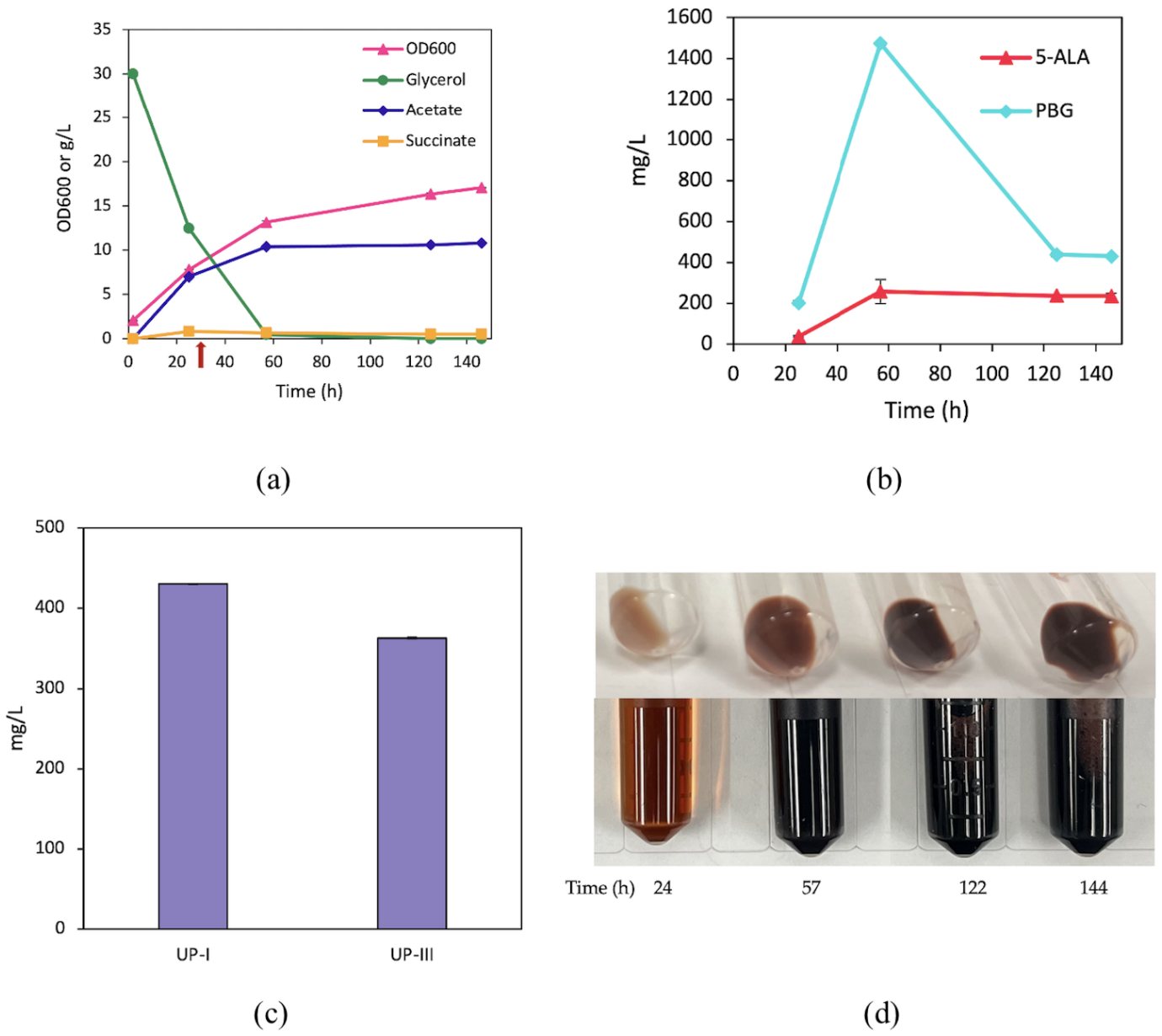


Figure 5.2: Aerobic bioreactor cultivation of BA001 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) UP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$).

5.3.2 Effects of the Native HemD/E for CP Biosynthesis

To investigate the effects of the native *hemD/E* on CP biosynthesis, we derived a CP-producing strain, BA002, by heterologous coexpression of *hemA* from *R. sphaeroides* and *hemB/D/E* from *E. coli*. For effective coexpression, the four genes of *hemA*, *hemB*, *hemD,* and *hemE* were cloned in a single plasmid to generate pK-hemABD-E, where the *hemA/B/D* genes formed the first operon regulated by the strong *trc* promoter, *hemE* formed the second operon regulated by the strong *grac*max promoter, and each gene had an individual strong RBS.

Cultivation of BA002 showed normal cell growth with rapid glycerol consumption and significant acetogenesis (Figure 5.3a). The strain produced 5-ALA and PBG, with peak levels of 99.7 mg/L (5.4 mg/OD_{600}/L) and 665.7 mg/L (35.7 mg/OD_{600}/L), respectively (Figure 5.3b). Additionally, the BA002 culture had intensified pigmentation, producing 84.3 mg/L (0.6% yield) and 182.1 mg/L (1.4% yield) of CP-I and CP-III, respectively (Figure 5.3c,d), with 17.5 mg/L and 20.1 mg/L of UP-I and UP-III being produced as side products, respectively (Figure 5.3c). The results suggest that a higher expression level of the native HemE was critical for CP production.

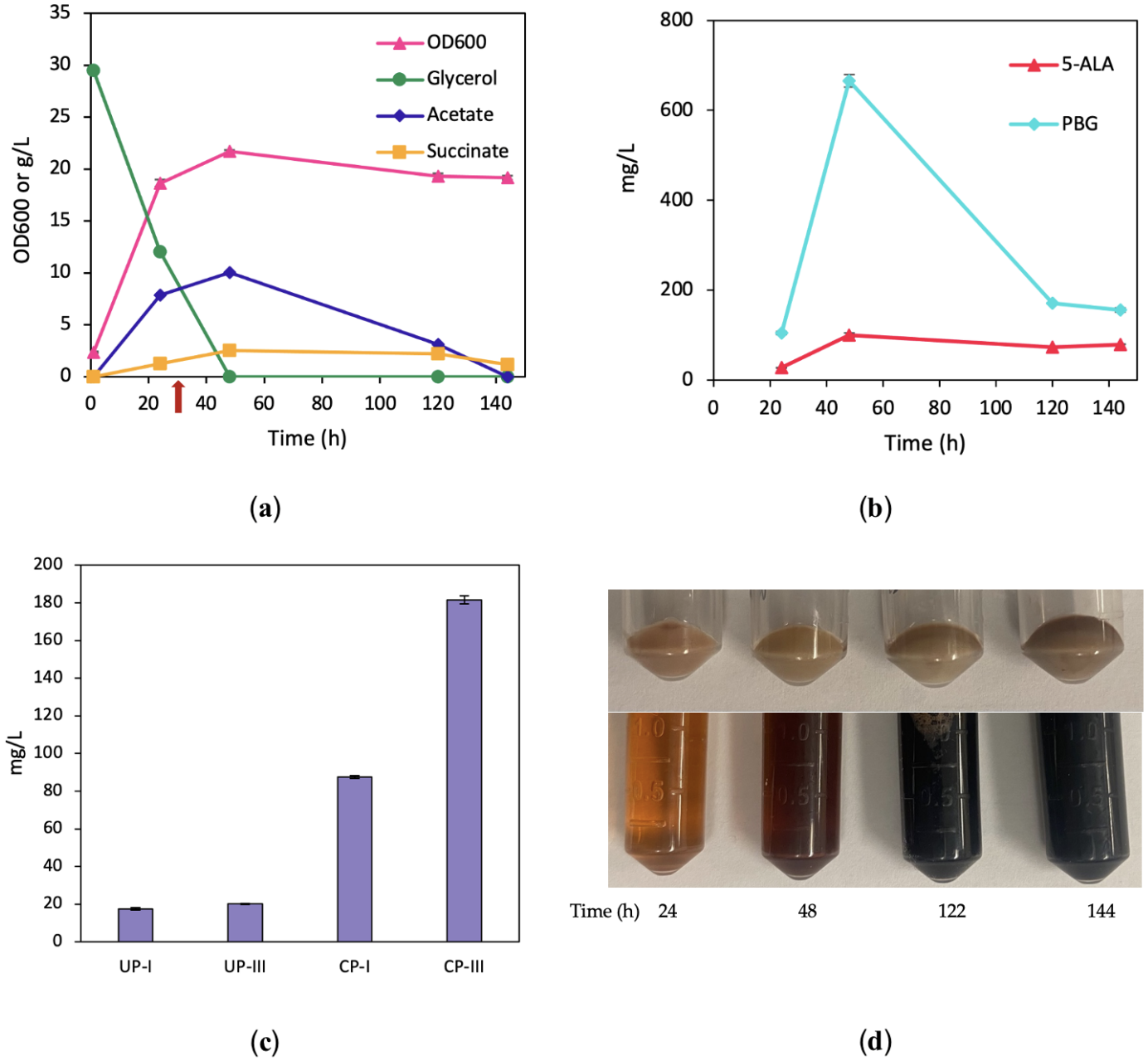


Figure 5.3: Aerobic bioreactor cultivation of BA002 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$).

5.3.3 Tuning of Expression Levels of HemD/E for CP Biosynthesis

To further evaluate the effect of *hemD/E* on CP biosynthesis, we constructed another plasmid pK-hemAB-DE, in which the two genes, *hemA/B*, were in the first operon regulated by the *trc* promoter and *hemD/E* were in the second operon regulated by the *grac*max promoter, resulting in CP-producing strain BA003. Such designs can potentially lead to a higher expression level of *hemD* and a lower expression level of *hemE*, in comparison to pK-hemABD-E. The tuning of the expression levels of the key enzymes in the Shemin/C4 pathway can potentially enable the identification of the step(s) limiting the overall CP biosynthesis.

Similar to BA002, cultivation of BA003 displayed a significant pigmentation (Figure 5.4d) but cell growth was somewhat retarded with significant acetogenesis (Figure 5.4a). Also, BA003 exhibited a similar pattern in the production of 5-ALA and PBG in comparison to BA002, with peak levels of 93 mg/L (10 mg/OD_{600}/L) and 867 mg/L (92.7 mg/OD_{600}/L), respectively (Figure 5.4b). Nevertheless, BA003 produced CP-I and CP-III slightly less than BA002, peaking at 68.2 mg/L (0.51% yield) and 119 mg/L (0.89% yield), respectively (Figure 5.4c). In addition, 18.6 mg/L UP-I and 20.3 mg/L UP-III were also produced. These observations further suggest the importance of the expression level of the native HemE for CP biosynthesis.

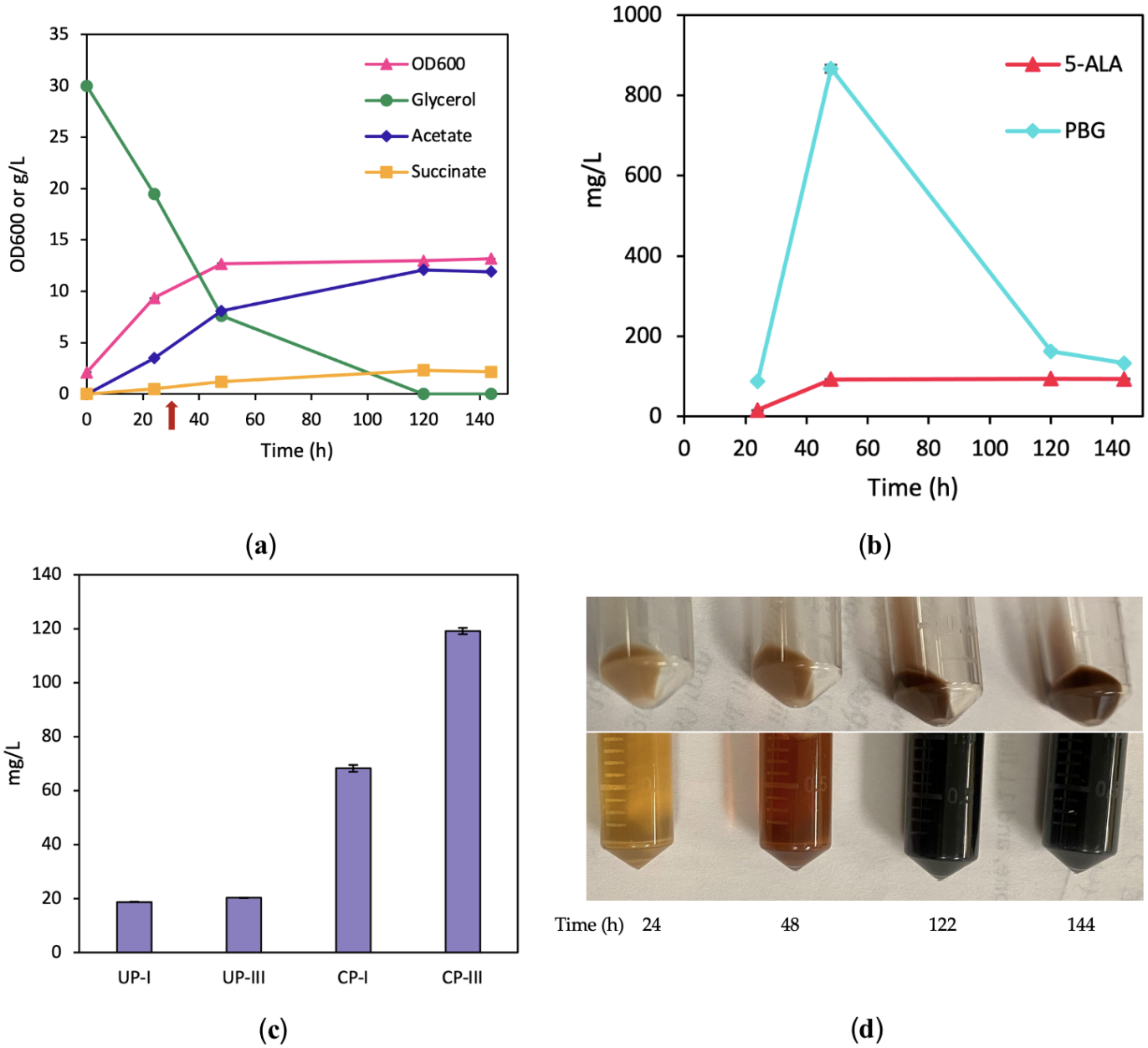


Figure 5.4: Aerobic bioreactor cultivation of BA003 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$).

5.3.4 Effects of HemY on CP Biosynthesis

The enzymatic conversion of CPG-III to CP-III is mediated by HemY (Figure 5.1). To investigate the effects of *hemY* on CP biosynthesis, we derived two plasmids, pK-hemABD-$EY_{B.s}$ and pK-hemABD-EY, by overexpressing *hemA/B/D* in the first operon regulated by *trc* promoter and overexpressing *hemE* and *hemY* from either *B. subtilis* or *E. coli* in the second operon regulated by *grac*max promoter, resulting in the CP-producing strains BA004 and BA005.

Upon cultivation of BA004, we observed a reduced pigmentation (Figure 5.5d) though cells grew normally with similar acetogenesis to BA002 and more rapid consumption of glycerol (Figure 5.5a). Also, BA004 exhibited reduced 5-ALA and PBG production in comparison to BA002, with peak levels of 104 mg/L (8.4 mg/OD_{600}/L) and 207 mg/L (16.7 mg/OD_{600}/L) at 120 h, respectively (Figure 5.5b). As a result, BA004 produced significantly less CP-I and CP-III than BA002, with peak concentrations reaching only 43.8 mg/L (0.33% yield) and 62.2 mg/L (0.47% yield), respectively. In addition, 6.2 mg/L and 10.6 mg/L of UP-I and UP-III were produced as side products, respectively (Figure 5.5c).

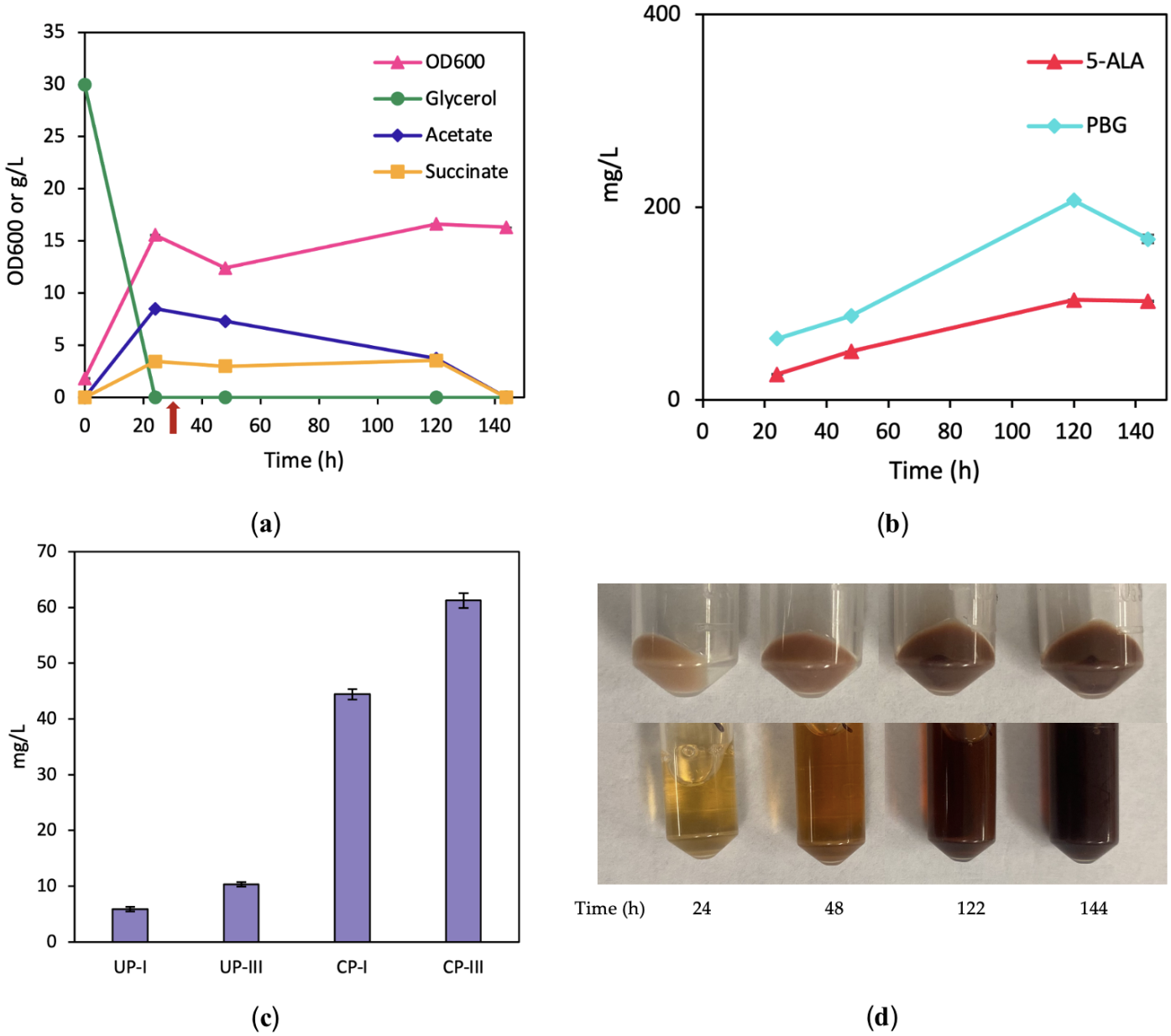


Figure 5.5: Aerobic bioreactor cultivation of BA004 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$).

Similarly, a reduced pigmentation was observed in cultivation of BA005 (Figure 5.6d) though cells grew normally with similar acetogenesis to BA002 (Figure 5.6a). Additionally, BA005 exhibited reduced 5-ALA and PBG production in comparison to BA002, with peak levels of 52.7 mg/L (3.7 mg/OD_{600}/L) and 60.4 mg/L (4.2 mg/OD_{600}/L) at 144 h, respectively (Figure 5.6b). As a result, BA005 yielded significantly lower amounts of CP-I and CP-III compared to BA002, with peak concentrations of 61.1 mg/L (0.46% yield) and 47.2 mg/L (0.35% yield), respectively. In addition, 5.8 mg/L and 6.2 mg/L of UP-I and UP-III were produced as side products, respectively. (Figure 5.6c). These observations suggest adverse effects upon *hemY* overexpression for CP biosynthesis. Consequently, we excluded *hemY* for coexpression in the subsequent experiments.

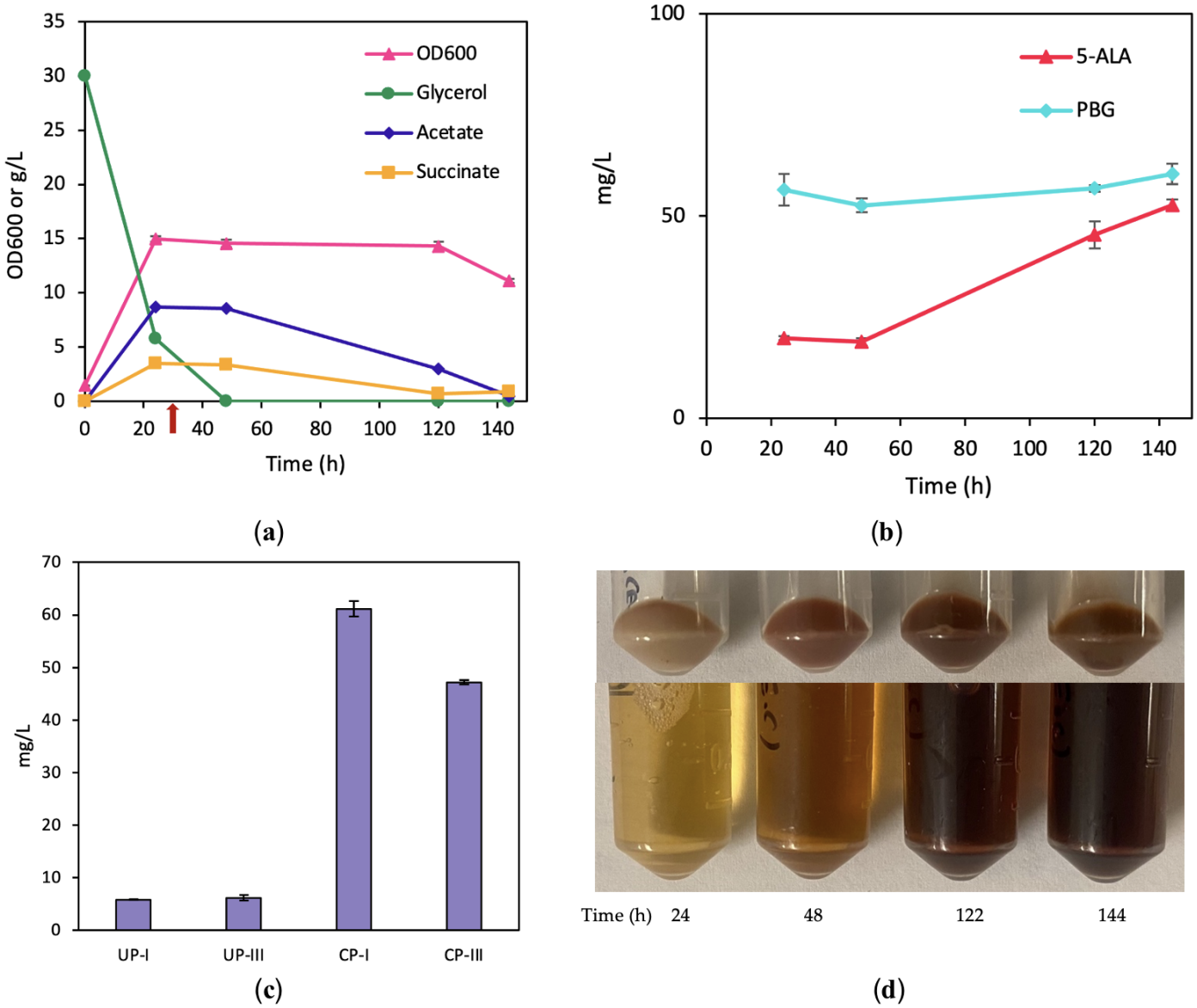


Figure 5.6: Aerobic bioreactor cultivation of BA005 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$).

5.3.5 HemD Overexpression Was Not Essential for CP Biosynthesis

Based on the cultivation results of BA002 and BA003, we further investigated whether the overexpression of the native *hemD* was essential for CP biosynthesis. To do this, we excluded *hemD* in the next derived plasmid pK-hemAB-E by overexpressing *hemA/B* in the first operon regulated by the *trc* promoter and *hemE* in the second operon regulated by the *grac*max promoter, resulting in the CP-producing strain BA006.

BA006 exhibited a glycerol consumption rate similar to BA002, accompanied by a progressively increasing cell density and significant acetogenesis (Figure 5.7a). Notably, BA006 produced high levels of 5-ALA and PBG, peaking at 408 mg/L (36.9 mg/OD_{600}/L) and 1156 mg/L (104.6 mg/OD_{600}/L) at 48 h, respectively (Figure 5.7b). Most importantly, BA006 produced even more CP than BA002 with a total titer of 353 mg/L comprising 102.4 mg/L and 250.6 mg/L of CP-I and CP-III, respectively (0.77% and 1.88% yields) (Figure 5.7c,d). Also, note that the product ratio/distribution of CP-I and CP-III in BA006 was similar to BA002, suggesting that *hemD* overexpression did not change the metabolic flux distribution between the classical pathway and autoxidation pathway at the node of HMB. Finally, it is noteworthy that other porphyrins were not detected as side products in this cultivation.

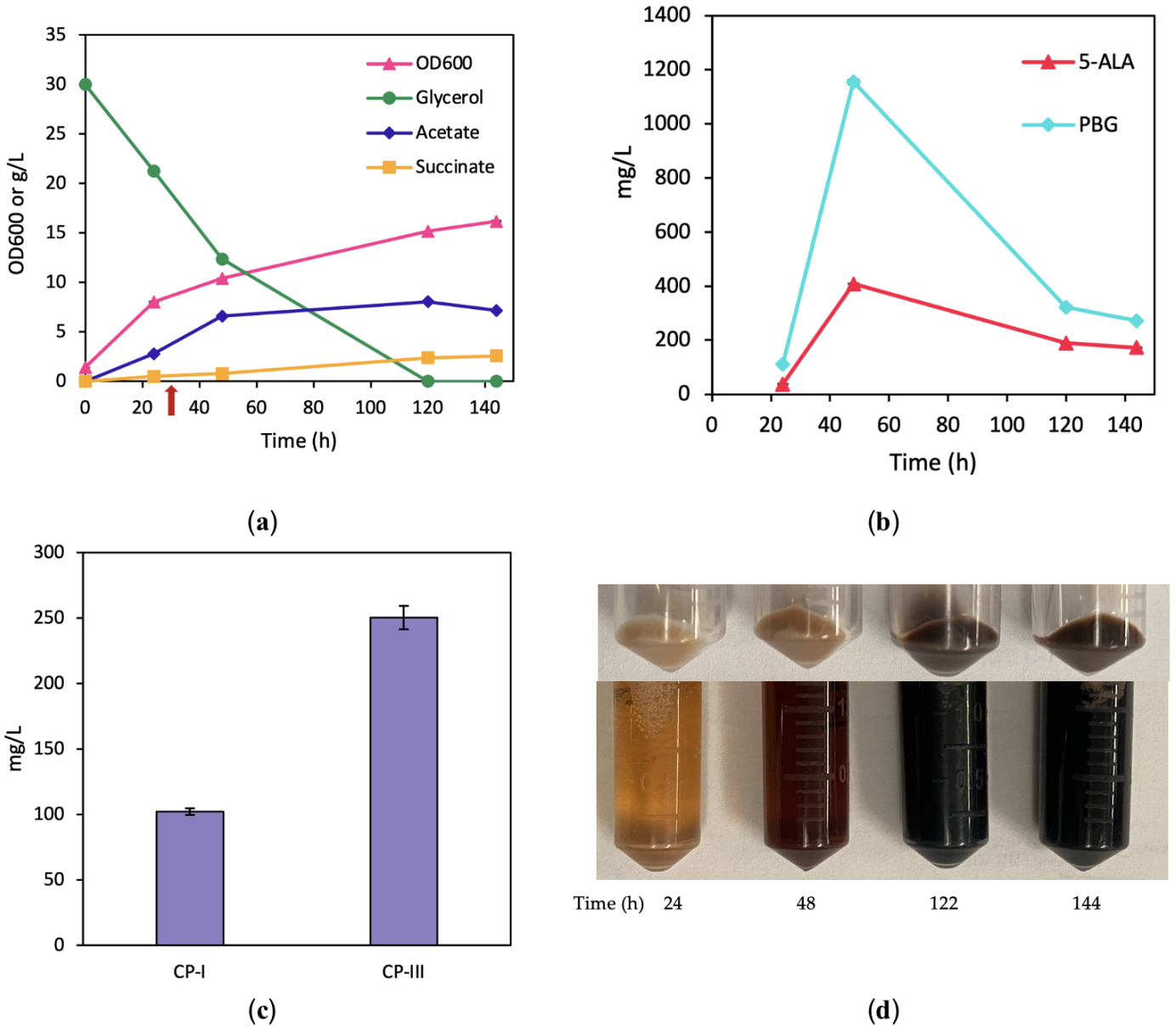


Figure 5.7: Aerobic bioreactor cultivation of BA006 for porphyrin biosynthesis. Time profiles of (**a**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (the red arrow shows glycine supplementation), (**b**) 5-ALA and PBG biosynthesis, and (**c**) CP biosynthesis. (**d**) Image of CFM and cell paste samples at different time points. All values are reported as means ± SD ($n = 2$).

5.4 Discussion

In this study, we enhanced CP biosynthesis in *E. coli* by strategically implementing and manipulating the Shemin/C4 pathway. In the production host with a double mutation of $\Delta iclR\Delta sdhA$ [216], the carbon flux from the tricarboxylic acid (TCA) cycle was redirected toward succinyl-CoA, a critical precursor in the CP biosynthesis pathway. Such redirection of carbon flux not only provided an abundant supply of succinyl-CoA but also enabled CP biosynthesis under aerobic conditions. The implementation of the Shemin/C4 pathway into the host strain was achieved through the heterologous expression of *hemA* from *R. sphaeroides*. Among the six genes, i.e., *hemA/B/C/D/E/Y*, contributing to CP biosynthesis (Figure 5.1), we previously identified the importance of *hemA/B* overexpression for UP biosynthesis, suggesting that the overall porphyrin biosynthesis via the Shemin/C4 pathway could be potentially limited by the two steps converting succinyl-CoA and glycine to PBG. In addition, we previously observed that the genomic expression level of *hemC* was sufficient for effective UP biosynthesis and *hemC* overexpression could potentially disrupt the metabolic activity of the Shemin/C4 pathway, resulting in the complete disappearance of culture pigmentation (Chapter 4). Additionally, prior studies have demonstrated that downregulation of *hemC* can result in extracellular accumulation of PBG [218]. Hence, in this study, *hemA/B* were included whereas *hemC* was excluded for overexpression in all the engineered strains for CP biosynthesis.

To evaluate the effects of the other three hem-genes, i.e., *hemD/E/Y*, on CP biosynthesis, we first derived strain BA001 by overexpressing *hemA/B/D*. Cultivation of BA001 revealed substantial pigmentation with extracellular production of UP (both UP-I and UP-III) only. Note that we previously observed that *hemD* overexpression hardly affected UP biosynthesis, suggesting that the genomic expression level of *hemD* was sufficient to drive the conversion of HBM to UPG-III. The absence of CP biosynthesis in BA001 also indicated a potential limitation in the reactions downstream of CPG-III catalyzed by HemE/Y. Specifically, HemE, which is responsible for catalyzing the conversion of UPG-III to CPG-III, can potentially play a critical role in the overall CP biosynthesis.

In the next attempt, we overexpressed the *hemA/B/D/E* genes from the plasmid pK-hemABD-E in BA002. Cultivation of BA002 resulted in a significant pigmentation with extracellular production of CP (both CP-I and CP-III), suggesting the critical role of HemE for CP biosynthesis. The produced CP can be extracellularly secreted, posing minimal intracellular toxicity on cell physiology. The production of UP-I in BA001 and CP-I in BA002 indicates the involvement of autooxidation as a side reaction converting HMB to UPG-I in *E. coli*. The relatively high titer ratio of UP-I/UP-III at 1.19 in BA001 suggests that conversion of HMB was more favorable to UPG-I via autooxidation than UPG-III via

HemD (even though *hemD* was overexpressed). The observation was unsurprising since bacterial cultivation for porphyrin production was conducted aerobically. However, such carbon flux distribution appears to be reversed, i.e., conversion of HMB was more favorable to UPG-III/CPG-III than UPG-I/CPG-I when *hemE* was also included for overexpression in BA002, even though bacterial cultivation was still conducted aerobically. Note that the total porphyrin titer in BA002 was significantly reduced compared to BA001 for unknown reasons. Nevertheless, the results suggest that the overexpressed HemE could play a critical role by directing more carbon flux toward the classical Shemin/C4 pathway for CP-III biosynthesis in *E. coli*. Also, it appears that more UPG-I was converted to CPG-I, resulting in increased CP-I titer in BA002. The results suggest that HemE can potentially mediate the conversion of UPG-I to CPG-I in a similar mechanism to its native activity converting UPG-III to CPG-III.

For overexpressing *hemA/B/D/E*, we also made another plasmid pK-hemAB-DE in which *hemD/E* genes were fused to form the second operon whose expression was regulated by the *gracmax* promoter, resulting in another engineered strain BA003. Compared to the plasmid pK-hemABD-E (in BA002), in which *hemE* was the only gene in the second operon regulated by the *gracmax* promoter, *hemE* expression for pK-hemAB-DE would presumably be lower whereas *hemD* expression for pK-hemAB-DE would be higher. The upregulated *hemD* expression and downregulated *hemE* expression for pK-hemAB-DE can be attributed to the transcription process for a gene operon, resulting in the formation of multiple immature cistronic mRNAs with the first mRNA being more stable than the latter mRNA(s) [243, 244]. Compared to BA002, we observed a slight reduction in CP biosynthesis with a similar titer ratio of CP-III/CP-I upon cultivation of BA003. The reduced CP biosynthesis may be associated with the reduced expression level of *hemE* in BA003.

Next, we investigated the effects of the final step on CP biosynthesis, i.e., the conversion of coproporphyrinogen III (CPG-III) into CP-III, which is catalyzed by HemY. To do this, we included *B. subtilis hemY* and native *E. coli hemY* for heterologous expression in BA004 and BA005, respectively. Unexpectedly, we noted significantly reduced culture pigmentation and CP titers for both engineered strains. Our observations align with a prior report [11], in which adverse effects of *B. subtilis hemY* overexpression on porphyrin titers were observed in multiple instances. The results suggest a possible inhibitory effect of HemY on the earlier enzymatic conversions for CP biosynthesis. In the CPD pathway of actinobacteria or firmicutes, coproporphyrinogen III oxidase (HemY) primarily catalyzes the oxidation of CPG-III to CP-III, which is a heme precursor unique to the CPD pathway [93]. However, HemY in *E. coli* is often annotated as a protoporphyrinogen oxidase gene, whose biological activity is different from the protoporphyrinogen oxidase found in

actinobacteria and firmicutes, and its role in heme biosynthesis is yet to be demonstrated [245]. Although *E. coli* possesses porphyrinogen peroxidase (YfeX), which can also convert CPG-III to CP-III [246], HemY from actinobacteria or firmicutes was identified to be highly efficient for such conversion [91].

Based on the cultivation results of BA001, BA002, and BA003, it appears that increasing *hemD* expression did not contribute to CP biosynthesis. Hence, we decided to exclude *hemD* from the next plasmid construct pK-hemAB-E in BA006. Compared to BA002 and BA003, cultivation of BA006 resulted in a substantial increase in the overall CP production with a more favorable CP-III/CP-I distribution. Interestingly, the conversion of HMB to UPG-III was significantly enhanced by *hemE* overexpression, even in the absence of *hemD* coexpression. The results corroborate the critical role that the overexpressed HemE played in directing more carbon flux toward the classical Shemin/C4 pathway for enhanced CP-III (with minimal UP) biosynthesis. This was particularly important for our developed engineered strains (which were cultivated aerobically) in terms of directing dissimilated carbon flux away from several autooxidations which lead to various side products (such as UP-I, CP-I, and UP-III). Note that cultivation of BA006 only produced CP without detectable UP formation and with considerably reduced acetogenesis compared to the other strains. The results also support our previous observation in Chapter 4 that the genomic expression level of *hemD* was sufficient enough to maintain the overall metabolic activity of the Shemin/C4 pathway, even under various genetic backgrounds and cultivation conditions for enhanced production of UP and CP. Based on the results of this study, it appears that HemE is the sole enzyme (after the HMB node) critically limiting the overall CP biosynthesis in *E. coli*.

It is notable that we observed significant acetogenesis, reaching up to 10.8 g/L, alongside CP production. Interestingly, in strains like BA002, BA004, and BA005, acetate was eventually metabolized by the cells as a secondary carbon source, particularly once the initial carbon source (i.e., glycerol) was depleted. Achieving complete elimination of acetogenesis during aerobic cultivation of *E. coli* can be particularly challenging due to metabolic constraints linked to carbon overflow metabolism. In conclusion, this study presents a significant advancement in the bio-based production of CP using *E. coli* as a production host, achieving high-level CP production with minimal byproduct formation. Future research could focus on the underlying reason behind the adverse effects of HemY on CP biosynthesis, providing valuable insights into the possible interactions of the pathway enzymes. Additionally, exploring the metabolic pathways and regulatory mechanisms of acetogenesis under aerobic conditions may uncover strategies to mitigate acetate overflow and enhance CP production efficiency.

Chapter 6

Two-Step Bio-Based Production of Coproheme and Heme in *Escherichia coli*: In Vivo and In Vitro Approaches

6.1 Introduction

Heme, an iron-containing compound within the tetrapyrrole class, is integral to various biological systems across all domains of life [247]. Its complex architecture enables it to function in numerous biochemical processes essential for life, such as oxygen transport in hemoglobin, electron transfer in cytochromes, and catalysis in peroxidases and other heme enzymes [90, 226]. Beyond its biological roles, heme is crucial in multiple industrial sectors, including pharmaceuticals and food additives [234, 248]. In the medical field, synthetic heme analogs are used for treating conditions like anemia and are pivotal in the development of diagnostics and therapeutic agents [249]. Additionally, heme's catalytic properties have garnered attention for its application in industrial chemical processes, such as environmental remediation and synthetic organic chemistry [250, 251].

Recent advances have focused on microbial heme biosynthesis, leveraging metabolic engineering in organisms such as *Escherichia coli* (*E. coli*) and *Saccharomyces cerevisiae* (*S. cerevisiae*). These biotechnological approaches aim to improve yields, reduce costs, and mitigate the environmental impacts of traditional chemical synthesis methods [82, 81]. The biosynthesis of heme involves three primary pathways: the porphyrin-dependent (PPD)

pathway, the coproporphyrin-dependent (CPD) pathway, and the siroheme-dependent (SHD) pathway. Initially, all three pathways share similar routes up to the formation of uroporphyrinogen III (UPG-III) [252] (Figure 6.1).

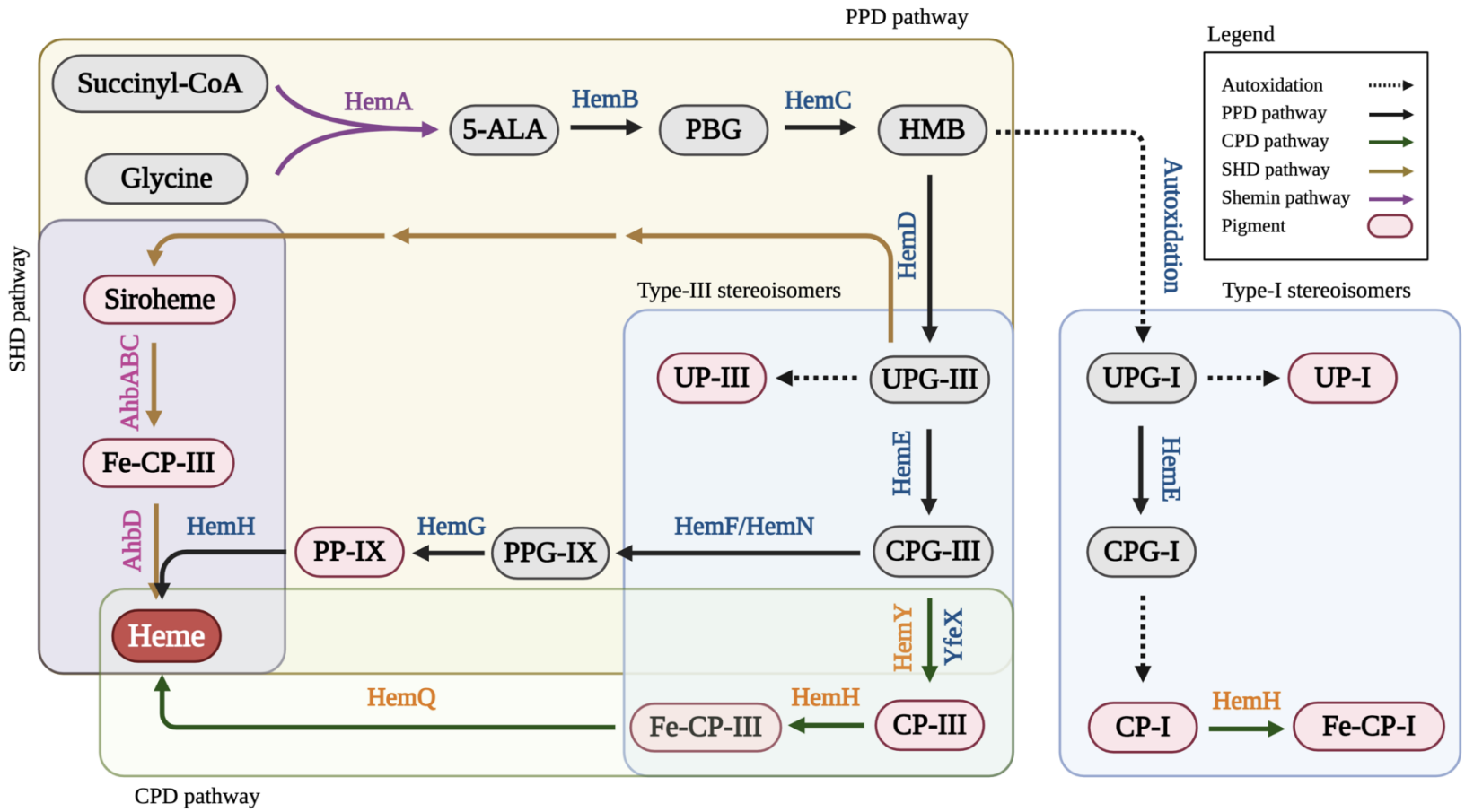


Figure 6.1: The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. 5-ALA, 5-aminolevulinic acid; AhbA, siroheme decarboxylase α subunit; AhbB, siroheme decarboxylase β subunit; AhbC, Fe-coproporphyrin III synthase; AhbD, SAM-dependent heme synthase; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; Fe-CP-I, Fe-coproporphyrin I; Fe-CP-III, coproheme; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin/coproporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HemQ, coproheme decarboxylase; HemY, coproporphyrinogen III oxidase; HMB, Hydroxymethylbilane; PBG, porphobilinogen; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase.

The SHD pathway is primarily found in archaea and sulfate-reducing bacteria occuring under anaerobic condition [253]. In this pathway, following the formation of UPG-III, the pathway diverges and involves multiple enzymatic steps where siroheme is produced as an intermediate. Siroheme is subsequently converted to coproheme (Fe-CP-III), and eventually to heme [39, 89]. Meanwhile, in the PPD and CPD pathways, the next step after the formation of UPG-III involves its conversion to coproporphyrinogen III (CPG-III) by the enzyme uroporphyrinogen III decarboxylase (HemE) [254]. From this enzymatic transformation, the pathways diverge into distinct routes for further heme synthesis [228].

In the PPD pathway found in gram negative bacteria, the conversion of CPG-III to heme involves several enzymatic transformations, each mediated by specific enzymes. Initially, CPG-III is converted to protoporphyrinogen IX (PPG-IX) by coproporphyrinogen III oxidase (HemF). This is followed by further enzymatic reactions catalyzed by protoporphyrinogen oxidase (HemG), leading to the formation of protoporphyrin IX (PP-IX). Finally, ferrochelatase (HemH) catalyzes the insertion of an iron ion into PP-IX to produce heme [255].

In contrast, the CPD pathway found in gram positive bacteria, offers a more streamlined route to heme production. After the formation of CPG-III, it is enzymatically converted to coproporphyrin III (CP-III) by coproporphyrinogen III dehydrogenase (HemY from gram positive bacteria or YfeX from gram negative bacteria). CP-III is then converted to Fe-CP-III by coproheme decarboxylase (ChdC/HemH), and eventually to heme by heme synthase (HemQ) [10, 227]. Unlike the SHD pathway, the CPD pathway involves fewer enzymatic steps, which reduces production costs and enhances potential yields [252]. This streamlined approach makes the CPD pathway a highly attractive option for sustainable heme production. Furthermore, the PPD pathway encounters significant challenges, such as cytotoxicity from PP-IX accumulation and negative feedback regulation on the hemB gene, both of which hinder its efficiency [214, 256, 219]. Consequently, the CPD pathway stands out as the promising choice for efficient and sustainable heme biosynthesis.

Despite the promise of bio-based heme production, significant challenges remain, primarily related to the toxicity associated with intracellular heme and porphyrin accumulation in microbial hosts. The buildup of these intermediates can lead to cytotoxicity, causing cell damage and limited yields [38, 257]. A lot of studies have been conducted to address these challenges by improving and enhancing the production of key intermediates including 5-aminolevulinic acid (5-ALA)[216], porphobilinogen (PBG)[218], UP-III (Chapter 4), and CP-III (Chapter 5). In addition to refining metabolic pathways, addressing issues such as pathway bottlenecks, substrate availability, and regulatory feedback mechanisms is essential. Innovative approaches are needed to overcome these challenges, one of which is cell-free metabolic engineering. This strategy offers a promising solution for heme pro-

duction by allowing metabolic pathways to be reconstituted outside living cells. Cell-free systems provide precise control over enzymatic reactions and circumvent the toxicity issues associated with intracellular heme accumulation [258]. They leverage the modularity and scalability of cell-free systems to enhance biochemical pathways for maximum efficiency and yield. Decoupling metabolic processes from cellular constraints enables the precise design of reaction environments tailored to specific production requirements [259, 260]. Furthermore, cell-free systems offer advantages such as reduced complexity and enhanced stability compared to traditional in vivo approaches [261].

In this manuscript, we propose a comprehensive two-step approach for bio-based heme production that integrates in vivo and in vitro strategies. In the first stage, we engineer *E. coli* strains to express the CPD pathway, yielding coproheme in vivo. Subsequently, we employ cell-free metabolic engineering techniques to convert coproheme to heme in vitro (Figure 6.2), circumventing intracellular toxicity issues and maximizing product yield. This synergistic approach not only addresses the inherent challenges of heme production but also provides a scalable and sustainable framework for industrial applications.

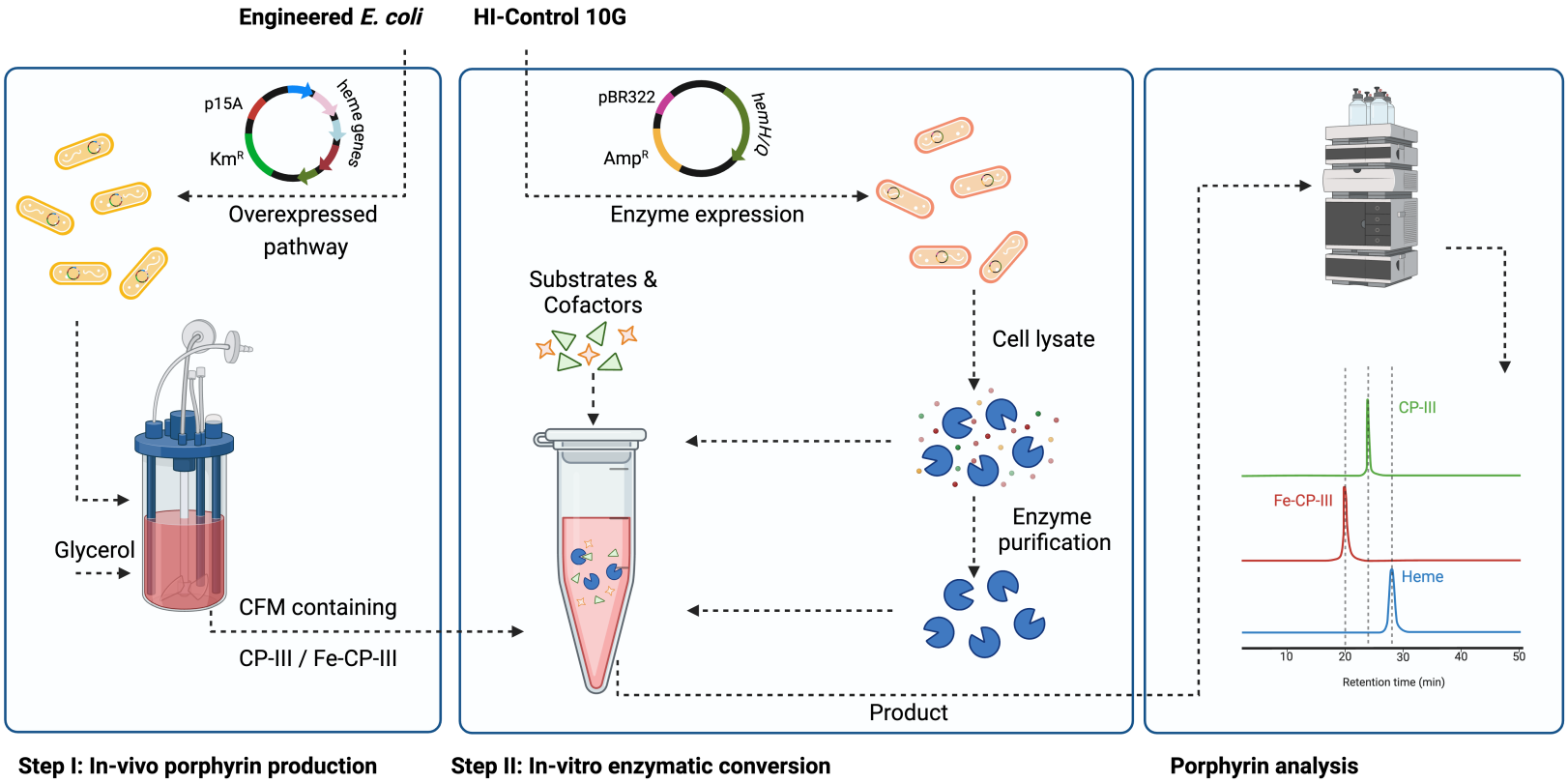


Figure 6.2: Two step heme production using in vivo porphyrin production and in vitro cell-free enzymatic conversion.

6.2 Materials and methods

6.2.1 Bacterial strains and plasmids

The bacterial strains and plasmids used in this study are provided in Table 6.1, and the primer sequences are listed in Table S6.1. Taq DNA polymerase was obtained from New England Biolabs (Ipswich, MA, USA). Genomic DNA of the bacterial cells was extracted using the Qiagen Blood & Tissue DNA Isolation Kit (Hilden, Germany). Plasmids were purified using the Qiagen Miniprep kit. All host cells in this study were derived from CPC-Sbm through the introduction of mutation(s) in the *sdhA* and/or *iclR* genes [200]. It is noteworthy that CPC-Sbm originated from BW25113 with the *ldhA* gene inactivated [184]. For molecular cloning, *E. coli* HI-Control 10G (Lucigen, Middleton, WI, USA) was used. DNA sequencing was performed by Plasmidsaurus (Eugene, Oregon, USA). All plasmids were constructed through Gibson assembly [181]. All oligonucleotides were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA).

Plasmids pK-hemABCD, pK-hemABD-E, and pK-hemAB-E were used from Chapters 4 and 5 to serve as the template for the remaining plasmids.

pK-hemABD-EHQ was constructed by amplifying *hemA* along with the backbone, and amplifying *hemB*, *hemD* and *hemE*, using primer sets P1/P2 and P3/P4 and pK-hemABD-E as template. The *hemH* and *hemQ* genes were amplified from the genomic DNA of *Bacillus Subtilis* (*B. Subtilis*) strain 168 using primer sets P5/P6 and P7/P8. These four fragments were Gibson assembled to form pK-hemABD-EHQ. For effective coexpression, *hemA*, *hemB* and *hemD* were aligned to form an operon *hemABD* regulated by a common strong *trc* promoter and *hemEHQ* were on the second operon regulated by the *grac*max promoter with an individual strong RBS for each gene.

Similarly, pK-hemAB-EHQ was constructed by amplifying *hemA* along with the backbone, and amplifying *hemB*, *hemD* and *hemE*, using primer sets P1/P2 and P3/P4 and pK-hemAB-E as template. The *hemH* and *hemQ* genes were amplified from the genomic DNA of *B. Subtilis* strain 168 using primer sets P5/P6 and P7/P8. These four fragments were Gibson assembled to form pK-hemAB-EHQ. *hemAB* on the first operon were regulated by *trc* promoter and *hemEHQ* were regulated by the *grac*max promoter.

pK-hemABD-EH was constructed by amplifying *hemA* along with the backbone, and amplifying *hemB*, *hemD*, *hemE* and *hemH*, using primer sets P9/P10 and P3/P11 and pK-hemABD-EHQ as template. These two fragments were Gibson assembled to form pK-hemABD-EH. For effective coexpression, *hemABD* were expressed on the first operon regulated by *trc* promoter and *hemEH* were expressed on the second operon regulated by the *grac*max promoter.

Table 6.1: Strains and plasmids used in this study.

Name	Description or Relevant Genotype	Source
Host strains		
HI-Control 10G	*mcrA*, Δ(*mrr-hsdRMS-mcrBC*), *endA1*, *recA1*, ϕ80*dlacZ*ΔM15, Δ*lacX74*, *araD139*, Δ(*ara leu*)*7697*, *galU*, *galK*, *rpsL* (Str^R), *nupG*, λ^- , *tonA*, Mini-F $lacI^{q1}$ ($Gent^R$)	Lucigen
MG1655	K-12; F^-, λ^-, *rph-1*	Lab stock
Bacillus Subtilis 168	Wild type	Lab stock
Staphylococcus aureus	Wild type	Lab stock
CPC-SbmΔ*iclR*Δ*sdhA*	F^-, Δ*(araD-araB)567*, Δ*lacZ4787(::rrnB-3)*, λ^-, *rph-1*, Δ*(rhaD-rhaB)568*, *hsdR514*, Δ*ldhA*, P_{trc}::*sbm* (i.e., with the FRT-P_{trc} cassette replacing the 204-bp upstream of the Sbm operon),Δ*iclR*, Δ*sdhA*	[200]
FECP1	CPC-SbmΔ*iclR*Δ*sdhA*/ pK-hemABD-EH	This study
FECP2	CPC-SbmΔ*iclR*Δ*sdhA*/ pK-hemAB-EH	This study
FECP3	CPC-SbmΔ*iclR*Δ*sdhA*/ pK-hemABE-H	This study
HEM1	CPC-SbmΔ*iclR*Δ*sdhA*/ pK-hemABD-EHQ	This study
HEM2	CPC-SbmΔ*iclR*Δ*sdhA*/ pK-hemAB-EHQ	This study
HEM3	CPC-SbmΔ*iclR*Δ*sdhA*/ pK-hemABE-$HQ_{S.a}$	This study
Plasmids		
pK-hemABCD-E	p15A ori, Km^R, P_{trc}::*hemABCD*-$P_{gracmax}$::*hemE*	Chapter 5
pK-hemABD-EH	p15A ori, Km^R, P_{trc}::*hemABD*-$P_{gracmax}$::*hemEH* -*hemH* from *B. subtilis* 168	This study
pK-hemAB-EH	p15A ori, Km^R, P_{trc}::*hemAB*-$P_{gracmax}$::*hemEH* -*hemH* from *B. subtilis* 168	This study
pK-hemABE-H	p15A ori, Km^R, P_{trc}::*hemABE*-$P_{gracmax}$::*hemH* -*hemH* from *B. subtilis* 168	This study
pK-hemABD-EHQ	p15A ori, Km^R, P_{trc}::*hemABD*-$P_{gracmax}$::*hemEHQ* -*hemH* and *hemQ* from *B. subtilis 168*	This study
pK-hemAB-EHQ	p15A ori, Km^R, P_{trc}::*hemAB*-$P_{gracmax}$::*hemEHQ* -*hemH* and *hemQ* from *B. subtilis 168*	This study
pK-hemABE-$HQ_{S.a}$	p15A ori, Km^R, P_{trc}::*hemABE*-$P_{gracmax}$::*hemHQ* -*hemH* from *B. subtilis* 168 and *hemQ* from *S. aureus*	This study
pK-hemH-His_6	pBR322 ori, Amp^R, P_{trc}::*hemH*:His_6 -*hemH* from *B. subtilis* 168	This study
pK-His_6-hemH	pBR322 ori, Amp^R, P_{trc}::His_6::*hemH* -*hemH* from *B. subtilis* 168	This study

pK-hemQ-His_6	pBR322 ori, Amp^R, P_{trc}::*hemQ*:His_6 -*hemQ* from *B. subtilis* 168	This study
pK-His_6-hemQ	pBR322 ori, Amp^R, P_{trc}::His_6::*hemQ* -*hemQ* from *B. subtilis* 168	This study
pK-$hemQ_{S.a}$-His_6	pBR322 ori, Amp^R, P_{trc}::*hemQ*:His_6 -*hemQ* from *S. aureus*	This study
pK-His_6-$hemQ_{S.a}$	pBR322 ori, Amp^R, P_{trc}::His_6::*hemQ* -*hemQ* from *S. aureus*	This study

pK-hemAB-EH was constructed by amplifying *hemA*, *hemB* and *hemE*, and amplifying the backbone, using primer sets P12/P13 and P9/P14 and pK-hemAB-E as template. The *hemH* gene was amplified from the genomic DNA of *B. Subtilis* strain 168 using primers P5/P15. These three fragments were Gibson assembled to form pK-hemAB-EH. For effective coexpression, *hemAB* were expressed on the first operon regulated by *trc* promoter and *hemEH* were expressed on the second operon regulated by the *grac*max promoter.

pK-hemABE-H was constructed by amplifying *hemA* and *hemB* using primers P16/P17, and amplifying *hemE* and *grac*max promoter separately, using primers sets P18/P19 and P20/P21 and pK-hemABD-EH as template. The *hemH* was amplified along with the backbone, using primers P22/P23 and pK-hemABD-EH as template. These four fragments were Gibson assembled to form pK-hemABE-H. *hemA*, *hemB* and *hemE* were aligned to form an operon *hemABE* regulated by *trc* promoter and *hemH* was included on the second operon regulated by the *grac*max promoter.

pK-hemABE-$HQ_{S.a}$ was constructed by amplifying *hemA* and *hemB* using primers P16/P17, and amplifying *hemE*, *grac*max promoter and *hemH*, using primers P18/P6 and pK-hemABD-EH as template. The $hemQ_{S.a}$ was amplified from the genomic DNA of *Staphylococcus aureus* (S. *aureus*) using primers P24/P25. The backbone was amplified from pK-hemABD-EH using primers P26/P23. These four fragments were Gibson assembled to form pK-hemABE-$HQ_{S.a}$. *hemABE* were regulated by *trc* promoter and $hemHQ_{S.a}$ were regulated by the *grac*max promoter.

pK-hemH-His_6 and pK- His_6-hemH were constructed by amplifying *hemH* from *B. Subtilis* strain 168 using primers P27/P28 and P29/P30, respectively. The fragment was then Gibson assembled with the backbone amplified from the pTrc99a plasmid using primer sets P31/P32 and P33/P34, respectively. There primers were designed to incorporate a 6x His-tag on either the C-terminus or N-terminus. A similar approach was applied for construction of the plasmids pK-hemQ- His_6 , pK-His_6-hemQ, pK-$hemQ_{S.a}$-His_6 and pK-His_6-$hemQ_{S.a}$. Briefly, *hemQ* was amplified using primer sets P35/P36, P37/P38, P39/P40, and

P41/P42, respectively, from the genomic DNA of *B. Subtilis* strain 168 and S. *aureus*. The backbone was amplified from the pTrc99a plasmid using primer sets P43/P44, P45/P46, P43/P47, and P48/P46, respectively. Each amplified gene was then Gibson assembled with its respective backbone to generate these four plasmids.

6.2.2 Media and bacterial cell cultivation

All medium components were obtained from Sigma-Aldrich Co. (St. Louis, MO, USA), with the exception of yeast extract and tryptone, which were obtained from BD Diagnostic Systems (Franklin Lakes). *E. coli* strains were preserved as glycerol stocks at -80 °C and streaked on lysogeny broth (LB; 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl) agar plates, then incubated at 37 °C for 14-16 h. For bioreactor cultivation, individual single colonies were picked from LB plates to inoculate 12 mL of super broth (SB) medium (32 g/L tryptone, 20 g/L yeast extract, and 5 g/L NaCl) in a 125 mL conical flask. The culture was incubated at 37 °C and 280 revolutions per min (rpm) using a rotary shaker (New Brunswick Scientific) for 4-6 h and subsequently used as a starter culture to inoculate 220 mL of SB medium at a 2% (vol/vol) concentration in a 1 L conical flask. The seed culture was incubated at 37 °C and 280 rpm for 14-16 h. Cells were harvested by centrifugation at 4,500 g and 20 °C for 8 min and resuspended in 40 mL of fresh SB medium. The resuspended culture was used to inoculate a stirred tank bioreactor (CelliGen 115, Eppendorf AG) with a working volume of 0.8 L at 37 °C and 430 rpm. The semi-defined production medium in the batch bioreactor contained 30 g/L glycerol, 0.23 g/L K_2HPO_4, 0.51 g/L NH_4Cl, 49.8 mg/L $MgCl_2$, 48.1 mg/L K_2SO_4, 1.52 mg/L $FeSO_4$, 0.055 mg/L $CaCl_2$, 2.93 g/L NaCl, 0.72 g/L tricine, 10 g/L yeast extract, 10 mM $NaHCO_3$, and 1 mL/L trace elements (2.86 g/L H_3BO_3, 1.81 g/L $MnCl_2{\cdot}4H_2O$, 0.222 g/L $ZnSO_4{\cdot}7H_2O$, 0.39 g/L $Na_2MoO_4{\cdot}2H_2O$, 79 μg/L $CuSO_4{\cdot}5H_2O$, 49.4 μg/L $Co(NO_3)_2{\cdot}6H_2O$) [224], supplemented with 0.05 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). To avoid glycine and Fe^{2+} limitation during the cultivation, 2 g of glycine and 0.45 g of $FeSO_4.7H_2O$ was supplemented into the bioreactor 30 h post inoculation. Aerobic condition was maintained by continuously purging the air into the bulk culture at 1 volume of air per volume of liquid per min (vvm). The pH of the bioreactor culture was maintained at 7.0 $\pm$ 0.1 using 3 M NH_4OH and 3 M H_3PO_4.

6.2.3 Enzyme purification

To purify the HemH and HemQ enzymes, immobilized metal affinity chromatography (IMAC) was employed. Briefly, a bacterial strain expressing His-tagged HemH or HemQ

was used to inoculate 200 mL of LB medium, starting with an initial OD_{600} of approximately 0.1. The culture was grown to an OD_{600} of 1, at which point protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG). After 5 hours post-induction, cells were harvested by centrifugation at 4 °C. The cell pellet was resuspended in lysis buffer and subjected to sonication for 4 minutes to lyse the cells. Following sonication, the lysate was centrifuged at 12,000 x g for 20 minutes at 4 °C to remove the cells. The resulting supernatant was filtered and loaded onto a 1 mL IMAC column pre-equilibrated with lysis/wash buffer. The column was washed with a low-concentration imidazole wash buffer to remove non-specifically bound proteins. HemH and HemQ enzymes were then eluted using an imidazole-rich elution buffer. The buffer of the eluted enzyme fractions was exchanged with a non-imidazole buffer and the enzymes were concentrated using Amicon centrifugal filters. The concentration of the purified enzymes was determined using the Bradford assay.

6.2.4 In vitro experiments reaction set up

The HemH assays were performed by quantifying the production of Fe-CP-III in reaction mixtures containing varying concentrations of CP-III and ferrous ions. Specifically, assays were performed with X μM CP-III and 2-8X μM $FeSO_4.7H_2O$, supplemented with 2-4 μM of the enzyme. In some assays, the effect of different buffers on enzyme activity was also tested.

HemQ activity was assessed by measuring the formation of heme from Fe-CP-III in the presence of hydrogen peroxide (H_2O_2). Each reaction mixture included X μM Fe-CP-III, 2-50X μM H_2O_2, and approximately 0.5-1X μM of the enzyme, using either purified enzyme or cell lysate. In some assays, the effect of different buffers on enzyme activity was also tested. In certain assays, standard reagents were substituted with cultivation broth to test the enzyme activity under varied conditions.

6.2.5 Intracellular porphyrin extraction

Porphyrins were extracted from the cells using a modified method from Fyrestam and Östman [262]. Briefly, cells were harvested by centrifugation and washed twice with 0.15 M saline solution. The washed cells were resuspended in 50 mM Tris-HCl buffer, chilled on ice, and subjected to ultrasonication for 4 minutes with 4-second pulses. The lysate was then centrifuged at 10,000 $\times$ g for 10 minutes to separate the supernatant, which could be analyzed for porphyrin content.

To extract heme, an equal volume of acetonitrile:1.7 M HCl (8:2, v/v) was added to the pellet, and the mixture was shaken for 20 minutes to release heme from proteins into the acetonitrile. A two-phase system was created by adding 0.25 mL of saturated $MgSO_4$(aq) and 0.25 g of NaCl(s) for each mL of sample. The solution was vortexed for 5 minutes and centrifuged again at 10,000 $\times$ g for 5 minutes. The top organic layer was filtered and diluted with acetonitrile before analysis.

6.2.6 Analysis

To measure cell density at OD_{600}, all culture samples were washed once and then appropriately diluted with a 0.15 M saline solution and analyzed using a spectrophotometer (GENESYSTM 40/50 Vis/UV-Vis, Thermo Fisher Scientific Inc., Waltham, MA, USA). To prepare cell-free medium, culture samples were centrifuged at 17,000 g for 1 min, followed by filter sterilization using a 0.2 μm syringe filter. Extracellular metabolites and glycerol were analyzed by high-performance liquid chromatography (HPLC; LC-10AT; Shimadzu, Kyoto, Japan) with a refractive index detector (RID; RID-10A; Shimadzu) and a chromatographic column (Aminex HPX-87H; Bio-Rad Laboratories, Hercules, CA, USA). The HPLC column temperature was maintained at 35 °C and the mobile phase was 5 mM H2SO4 (pH 2) running at 0.6 mL/min. The RID signal was analyzed by data processing software (Clarity Lite v. 7.4.1.88 software, DataApex, Prague, Czech Republic). The levels of 5-ALA and PBG in the cell-free medium were assessed using a modified Ehrlich's reagent assay [225]. Porphyrins were analyzed using an HPLC (2690 separation module, Waters™, Milford, MA, USA) equipped with a photodiode array (2996 PDA detector, Waters™, Milford, MA, USA) detector and a chromatographic column (Chromolith® HighResolution RP-18 endcapped, Supelco, Darmstadt, Germany). The UV absorbance was detected at 400 nm and data were processed using Empower 3 software (Waters™, Milford, MA, USA). A mobile-phase system previously outlined [11] was used with minor modifications. The percentage yield was determined as the ratio of the produced product to the maximum theoretical yield, calculated based on the glycerol consumed in cultivations or the porphyrin substrate consumed in the in vitro experiments. For in vitro experiments, the conversion efficiency was calculated as the ratio of the porphyrin substrate consumed to its initial concentration.

6.3 Results

6.3.1 Effects of HemH on Fe-CP-III biosynthesis

For effective Fe-CP-III biosynthesis, we derived Fe-CP-III-producing strains by heterologous coexpression of *hemA* from *R. sphaeroides*, *hemB*, *hemD*, and *hemE* from *E. coli*, and *hemH* from *B. subtilis*. Three plasmids were constructed for this purpose, each containing two operons regulated by the strong *trc* and *gracmax* promoters, respectively, and each cloned gene had an individual strong RBS. For the plasmid pK-hemABD-EH in strain FECP1, *hemA*, *hemB*, and *hemD* were cloned into the first operon, whereas *hemE* and *hemH* were placed in the second operon. Plasmid pK-hemAB-EH in FECP2 was similar to pK-hemABD-EH except that *hemD* was excluded from the construct. In the plasmid pK-hemABE-H in FECP3, the genes from pK-hemAB-EH were rearranged by including *hemABE* in the first operon and *hemH* in the second operon. This design was made to enhance the expression of *hemH* by placing it as the first gene in the second operon, potentially improving Fe-CP-III biosynthesis.

Time profiles for cell growth, glycerol consumption, and by-product (acetate and succinate) formation for each strain are shown in Figure 6.3a, d, and g. The three strains showed similar glycerol consumption patterns, with complete depletion of glycerol by 48 h. Among the strains, FECP2 exhibited the highest cell growth, reaching a cell density of 18.4 OD_{600} at 24 h. All strains produced high levels of acetate as a by-product, while succinate accumulation remained minimal in all cultivations.

Time profiles of 5-ALA and PBG, two key intermediates in the Shemin/C4 pathway, are shown in Figure 6.3b, e, and h. Among the three strains, FECP3 produced the highest levels of PBG, with the peak concentration reaching 1537 mg/L (104 mg/OD_{600}/L) by 48 h. In contrast, 5-ALA production was highest in FECP2, peaking at 261 mg/L (15.2 mg/OD_{600}/L) at 48 h. The high PBG titer in FECP3 indicates a high metabolic flux toward porphyrin biosynthesis.

Porphyrin production at 144 h for each strain is shown in Figure 6.3c, f, and i. Among these strains, FECP3 exhibited the highest extracellular Fe-CP-III concentration, reaching 84.9 mg/L and intracellular accumulation of only 2.3 mg/L, indicating efficient secretion of Fe-CP-III. In contrast, both FECP1 and FECP2 showed lower extracellular Fe-CP-III titers of 18.8 mg/L and 44.4 mg/L, respectively, with intracellular Fe-CP-III titers of 5.1 mg/L and 9.8 mg/L, respectively. In addition to Fe-CP-III, various byproducts were detected in both intracellular and extracellular fractions, including CP-I, CP-III, Fe-CP-I, PP-IX, and heme. The accumulation of intracellular PP-IX and heme in these strains indicates that a small portion of the dissimilated carbon flux was side-directed toward the

PPD pathway, even though the flux was metabolically driven toward the CPD pathway for Fe-CP-III production.

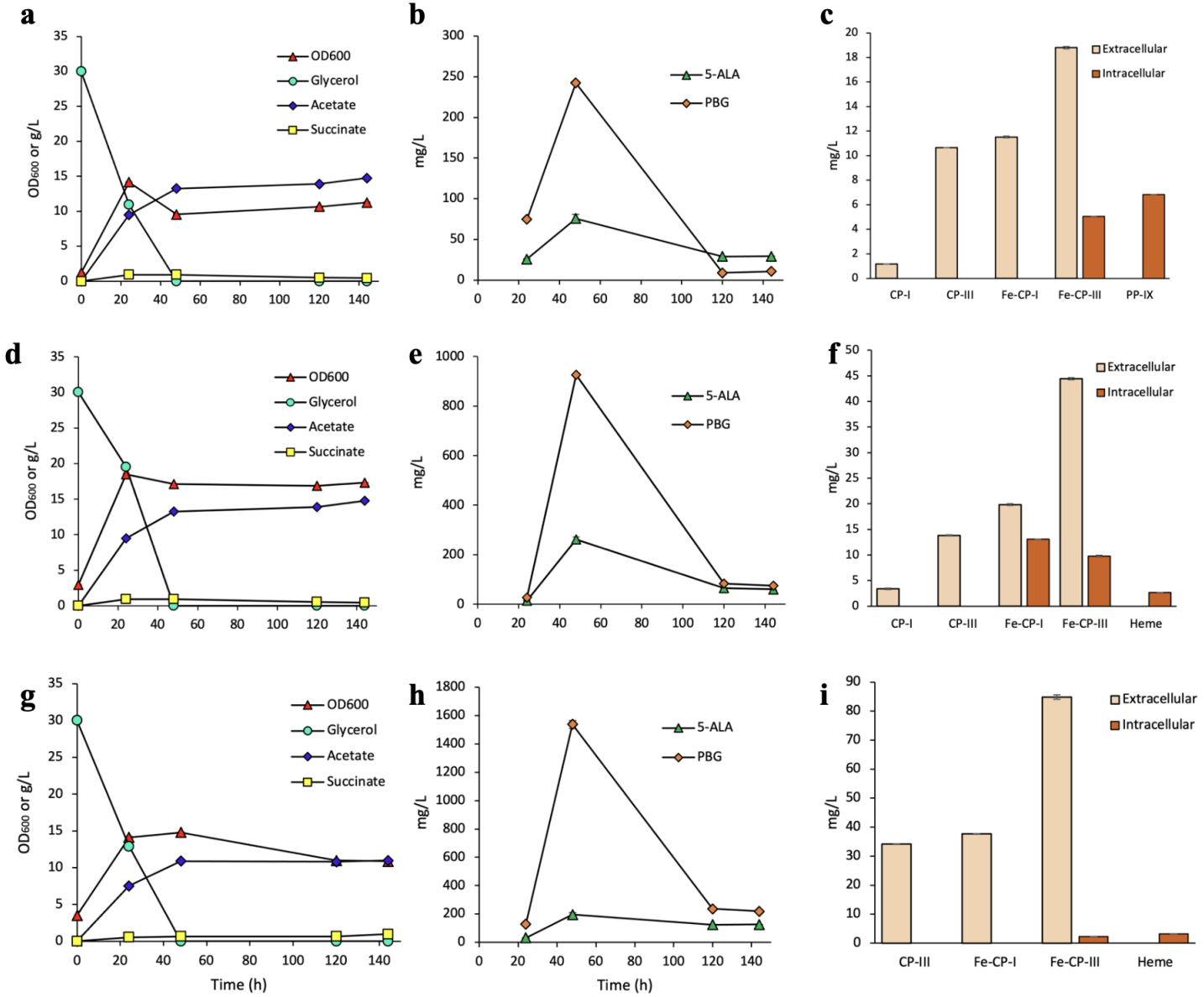


Figure 6.3: Aerobic bioreactor cultivation of FECP1, FECP2 and FECP3 for porphyrin biosynthesis. Time profiles of (**a**,**d** and **g**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**b**,**e** and **h**) 5-ALA and PBG biosynthesis, and (**c**,**f** and **i**) porphyrin biosynthesis. All values are reported as means±SD (n=2).

6.3.2 Effects of HemQ on heme biosynthesis

To investigate the effects of *hemQ* on heme biosynthesis, we derived heme-producing strains by heterologous coexpression of *hemA* from *R. sphaeroides*, *hemBDE* from *E. coli*, *hemH* from *B. subtilis*, and *hemQ* from either *B. subtilis* (referred as *hemQ*) or *S. aureus* (referred as $hemQ_{S.a}$). Three plasmids were constructed for this purpose, each containing two operons regulated by the strong *trc* and *gracmax* promoters, and each cloned gene had an individual strong RBS. In the first strain HEM1, *hemABD* and *hemEHQ* were cloned into the first and second operons, respectively, to form plasmid pK-hemABD-EHQ. In the second strain HEM2, the plasmid pK-hemAB-EHQ was similar to pK-hemABD-EHQ except that *hemD* was excluded. In the third strain HEM3, the harboring plasmid pK-hemABE-$HQ_{S.a}$ was constructed by including *hemABE* in the first operon and *hemH* and $hemQ_{S.a}$ in the second operon.

Time profiles for cell growth, glycerol consumption, and by-product formation for each strain are shown in Figure 6.4a, d, and g. All strains fully depleted glycerol by 48 h, with HEM3 showing slower consumption in the first 24 h. HEM3 exhibited the highest cell growth, reaching a peak density of 31.3 OD600 at 48 h. HEM1 had considerable formation of acetate and succinate, with both monotonically increasing, while HEM2 showed acetate depletion by the end of cultivation and minimal succinate formation. In contrast, HEM3 had negligible acetate and succinate production.

Time profiles of 5-ALA and PBG are shown in Figure 6.4b, e, and h. All strains showed similar patterns in 5-ALA and PBG formation, with both peaking at 48 h. 5-ALA titers reached 94.7, 143, and 36.6 mg/L (9.0, 6.8, and, 1.2 mg/OD_{600}/L), while PBG titers reached 232, 366, and 316 mg/L (22.1, 17.0, and 10.1 mg/OD_{600}/L), in HEM1, HEM2, and HEM3, respectively. Note that HEM2 produced the highest levels of both 5-ALA and PBG.

Heme production at 144 h for each strain is shown in Figure 6.4c, f, and i. Heme was produced intracellularly at 15.8, 17.8, and 4.0 mg/L in HEM1, HEM2, and HEM3, respectively. In addition to heme, various porphyrins were also detected in both intracellular and extracellular fractions, including UP-I, UP-III, CP-I, CP-III, Fe-CP-III and PP-IX.

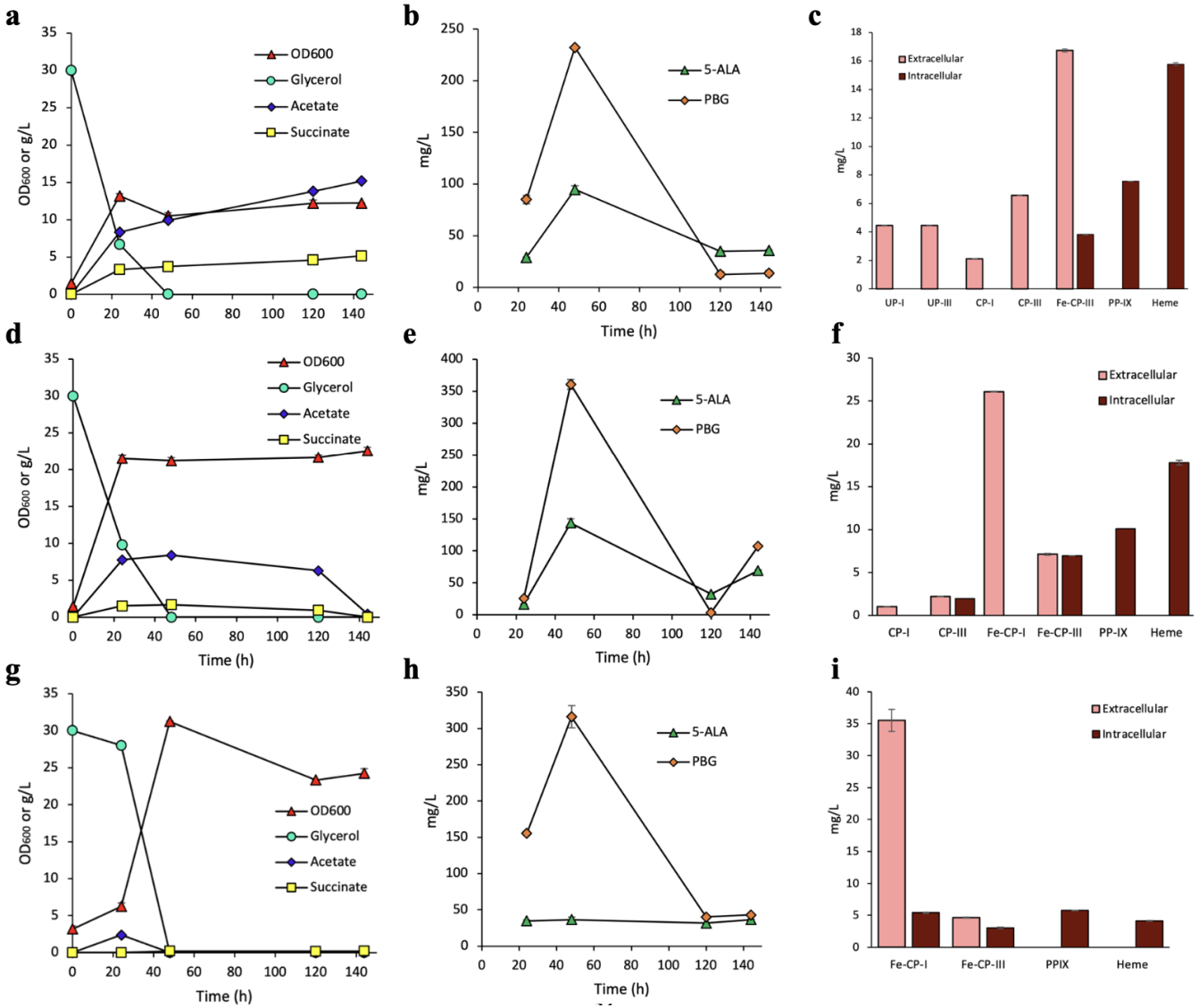


Figure 6.4: Aerobic bioreactor cultivation of HEM1, HEM2 and HEM3 for porphyrin biosynthesis. Time profiles of (**a**,**d** and **g**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**b**,**e** and **h**) 5-ALA and PBG biosynthesis, and (**c**,**f** and **i**) porphyrin biosynthesis. All values are reported as means±SD (n=2).

6.3.3 In vitro conversion of CP-III to Fe-CP-III via HemH

To explore the in vitro enzymatic conversion of CP-III to Fe-CP-III, *hemH* from *B. subtilis*, fused with a His_6 tag at either the C- or N-terminal, was PCR-amplified and cloned into plasmid pTrc99a for expression. The overexpressed HemH was purified using affinity chromatography. As illustrated in Figure 6.5, the insoluble fraction (lane 2), soluble fraction (lane 3), and purified protein (lane 4) were analyzed using SDS-PAGE. Clear bands corresponding to the expected molecular weight of 35 kDa were observed, confirming functional expression and purification [263]. The expressed HemH was predominantly located in the soluble fraction, consistent with a previous report [91].

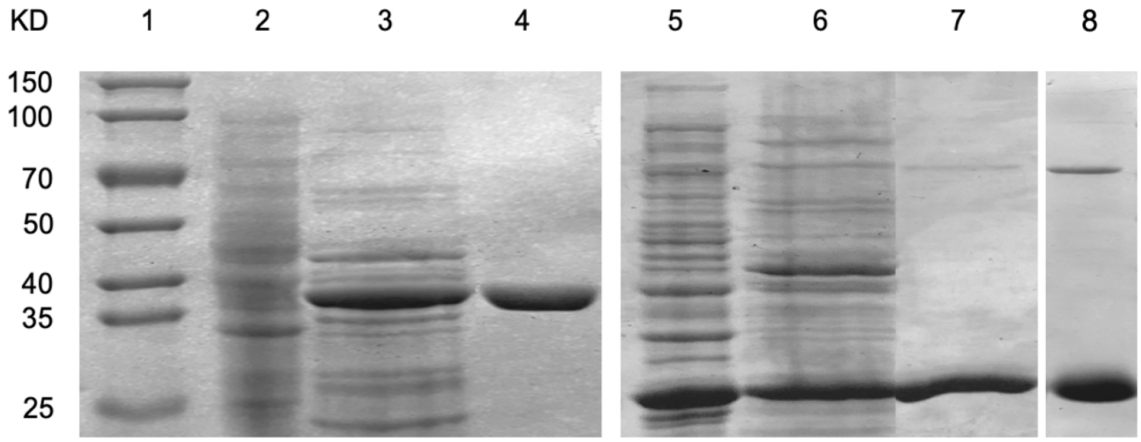


Figure 6.5: SDS-PAGE analysis of HemH and HemQ expression and purification. Lane 1: Protein ladder; Lane 2: HemH insoluble fraction; Lane 3: HemH soluble fraction; Lane 4: Purified HemH from *B. subtilis*; Lane 5: HemQ insoluble fraction; Lane 6: HemQ soluble fraction; Lane 7: Purified HemQ from *B. subtilis*; Lane 8: Purified HemQ from *S. aureus*.

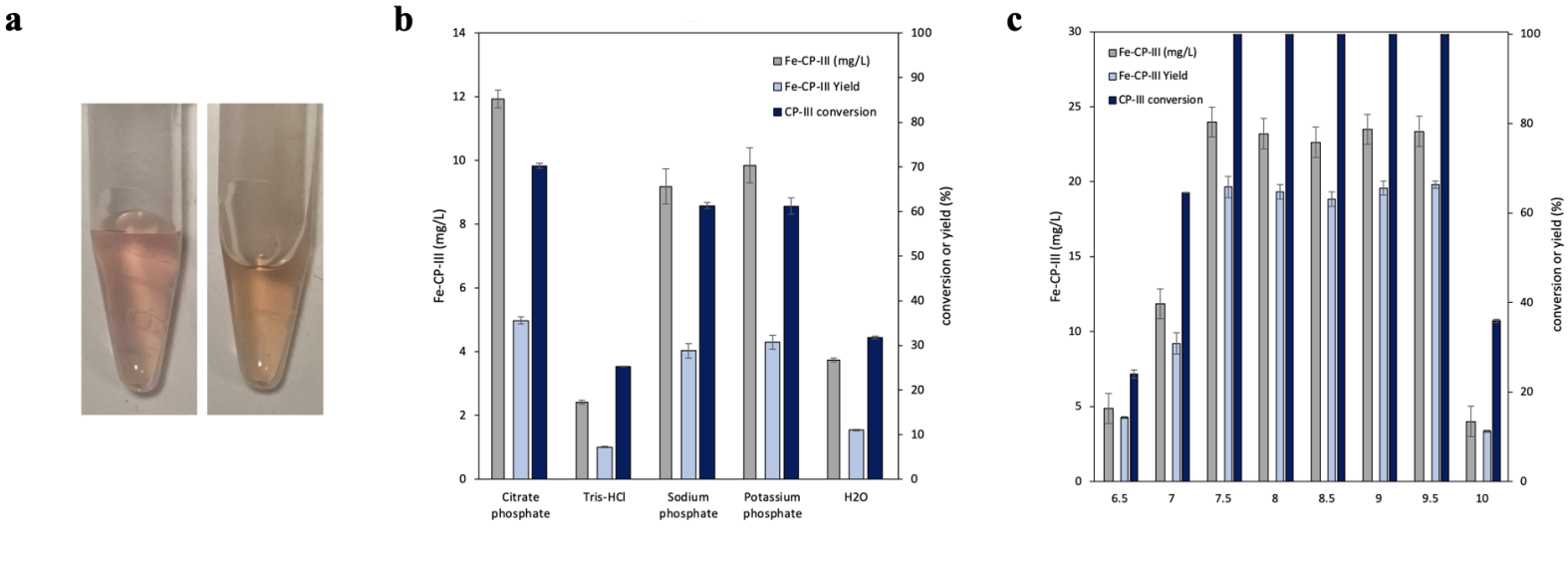
Purified HemH was applied to convert CP-III to Fe-CP-III in the presence of ferrous ions. The color change observed during the reaction from red to orange (Figure 6.6a) suggested successful conversion [82]. The reaction was enhanced by adjusting the buffer composition and pH. Four buffers based on citrate phosphate, Tris-HCl, sodium phosphate, and potassium phosphate at 100 mM and pH 7 were used for comparison. The buffer based on citrate phosphate had the best performance with a 35.6% yield (Figure 6.6b), whereas the other buffers had significantly lower yields. Next, the effects of pH was investigated and high yield up to 66.3% was achieved by adjusting pH to a range between 7.5 and 9.5

(Figure 6.6c). Based on these results, future in vitro conversions of CP-III to Fe-CP-III were conducted using citrate phosphate buffer at pH 8.

In the initial set of Experiments (A–D) (Figure 6.6d), the ratio among substrates and enzyme was examined. We began with the conversion of 50 µM CP-III, 2x ferrous ions (100 µM) as $FeSO_4.7H_2O$, and 2 µM HemH without the addition of buffer (Experiment A). This initial condition resulted in a 15.6% yield. Then, we increased the levels of ferrous ions and HemH in subsequent experiments. Raising the ferrous ion concentration to 200 µM while keeping the enzyme at 2 µM, led to an increased yield of 35.7%. (Experiment B) Further increasing the ferrous ion concentration to 400 µM decreased the yield to 11.7% (Experiment C). Similarly, doubling the enzyme concentration to 4 µM while maintaining 200 µM ferrous ions resulted in a 27.4% yield (Experiment D). These results indicate the importance of modulating the concentrations of substrates and enzyme for the in vitro conversion.

Further enhancement was made by the inclusion of citrate phosphate buffer at pH 8 to achieve 69.2% reaction yield (Experiment E). In Experiments F and G, the CP-III concentration was increased to 63 µM and 85 µM, respectively, while maintaining the substrate ratio (Fe^{2+}/CP-III) constant and the enzyme concentration at 4 µM. Despite the increased substrate concentration, high Fe-CP-III yields of 72.9% and 77.2% were obtained.

In addition to CP-III, we also observed that CP-I, the stereoisomer of CP-III could be another substrate for HemH with 68.0% yield of Fe-CP-I (Experiment H), suggesting that the enzyme HemD has a broad substrate specificity within the coproporphyrin family. Cell-free cultivation medium from CP-producing strains, even without purification or buffer exchange, could be directly used for HemD-conversion (Experiment I K) with Fe-CP-III yields ranging from 37% to 61.9%. Additionally, all CP-I in the three reactions was converted to Fe-CP-I with yields ranging from 47.6 to 64.0% (data not shown), indicating that HemH may have a higher specificity toward CP-I. Finally, note that similar results were obtained for both N-terminal- and C-terminal-His_6-tagged HemH enzymes (data not shown).



	A	B	C	D	E	F	G	H	I	J	K
CP-I/III (μM)	45	45	45	45	45	63	85	45	194	45	30
Fe^{2+} (μM)	90	180	360	180	180	252	340	180	776	180	120
HemH (μM)	2	2	2	4	4	4	4	4	4	4	4
buffer (100mM)	-	-	-	-	Citrate phosphate	Citrate phosphate	Citrate phosphate	Citrate phosphate	-	-	-

Figure 6.6: In vitro Fe-CP-III synthesis from CP-III. (**a**) Visible color change from red to orange, indicating the conversion of CP-III to Fe-CP-III. Reaction before adding enzyme (left) and after addition of enzyme (right) (**b**) Evaluation of buffer conditions for Fe-CP-III production. (**c**) Evaluation of pH conditions for Fe-CP-III production. (**d**) Tuning Fe^{2+} and HemH enzyme concentration, and buffer types for Fe-CP-III production.

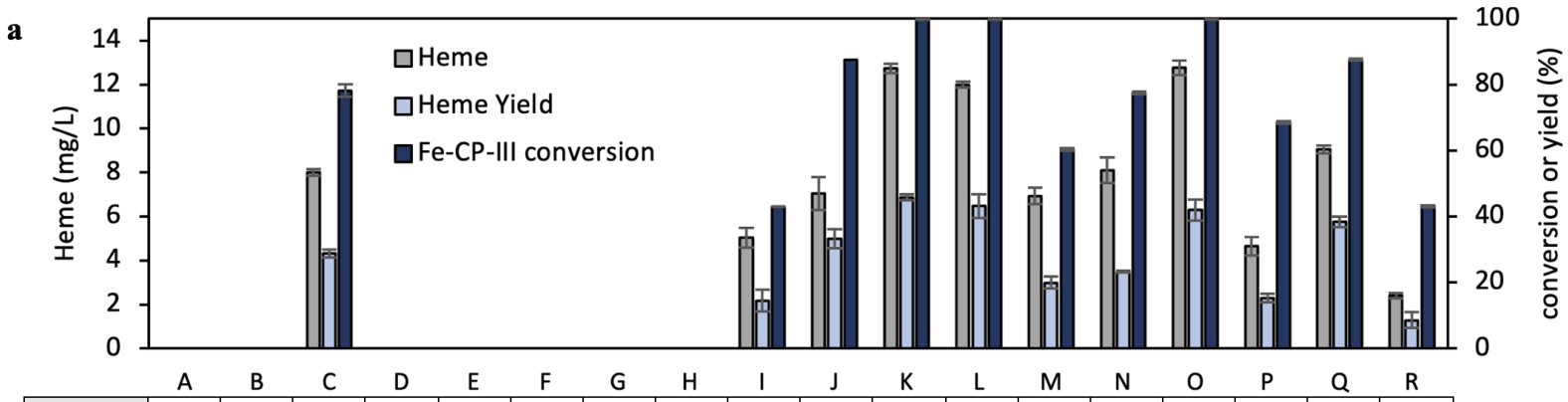
6.3.4 In vitro conversion of Fe-CP-III to heme via HemQ

For the in vitro enzymatic conversion of Fe-CP-III to heme, *hemQ* from *B. subtilis* was amplified, fused with a His_6 tag at either the C- or N-terminal, and cloned into the pTrc99a plasmid for expression and purification. The overexpressed HemQ enzyme was purified using affinity chromatography. As shown in Figure 6.5, the insoluble fraction (lane 5), soluble fraction (lane 6), purified HemQ (lane 7), and purified $HemQ_{s.a}$ (lane 8) were analyzed with SDS-PAGE. All lanes had clear bands corresponding to the expected molecular weight of 26 kDa, confirming successful expression and purification [264]. The enzyme was predominantly located in the soluble fraction.

Purified HemQ was used in reaction mixtures containing Fe-CP-III and H_2O_2 under various conditions (Figure 6.7a). We first investigated the effects of the ratio of Fe-CP-III/H_2O_2 by adjusting it from 1/2 to 1/50 (Experiment B–E). The best performance was observed in Experiment C with a 1/5 ratio, achieving 78.1% conversion and a heme yield of 28.7%. Increasing the ratio beyond 1/5 led to discoloration of the reaction mixture and impacted the conversion since hydrogen peroxide potentially acted as an oxidizing agent.

Among different buffers tested, including sodium phosphate, potassium phosphate, citrate phosphate, and Tris-HCl (Experiment C, F–H), sodium phosphate was the only effective one. Further increasing the amount of HemQ did not show positive effects (Experiment I). In addition, cell-free cultivation broth of Fe-CP-III-producing strains was directly used for the in vitro conversion, leading to a heme yield of 33.3% (Experiment J). Furthermore, the use of HemQ-overexpressed cell lysate could also achieve a high heme yield, either with pure Fe-CP-III or cell-free Fe-CP-III-containing medium as the substrate, achieving 45.8% (Experiment K) and 43.1% (Experiment L), respectively. Similar in vitro conversions were conducted using HemQ from *S. aureus*, (Experiment M–R) with different enzyme sources (i.e., purified HemQ or cell lysate) or the substrate type (i.e., pure Fe-CP-III or cell-free cultivation broth), achieving relatively heme yields with the highest one up to 41.9% (Experiment O).

To track the in vitro conversion of Fe-CP-III to heme, HPLC analysis of a reaction mixture was conducted at 1 h and 16 h post-reaction times (Figure 6.7b). Partial conversion of Fe-CP-III into an intermediate product with no detectable heme formation was observed at 1 h, whereas nearly all Fe-CP-III was converted to heme at 16 h. The results suggest the presence of a porphyrin intermediate for the in vitro conversion [265].



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Fe-CP-III (μM)	45	45	45	45	45	45	45	45	55	45*	45	45*	55	55	50*	50*	38*	45*
H_2O_2 (μM)	90	90	225	450	2250	225	225	225	275	225	225	225	275	275	250	250	190	225
HemQ (μM)	2	20	20	20	20	20	20	20	68	20	28**	28**	68	14**	68	3.5**	68	6.8
buffer (100mM)	Sodium phosphate	Sodium phosphate	Sodium phosphate	Sodium phosphate	Sodium phosphate	Potassium phosphate	Citrate phosphate	Tris-HCl	Sodium phosphate	-	-	-	Sodium phosphate	-	-	-	-	-

* Indicate the usage of cell free cultivation broth of Fe-CP-III producing strains.
** Indicate the use of HemQ-rich cell lysate

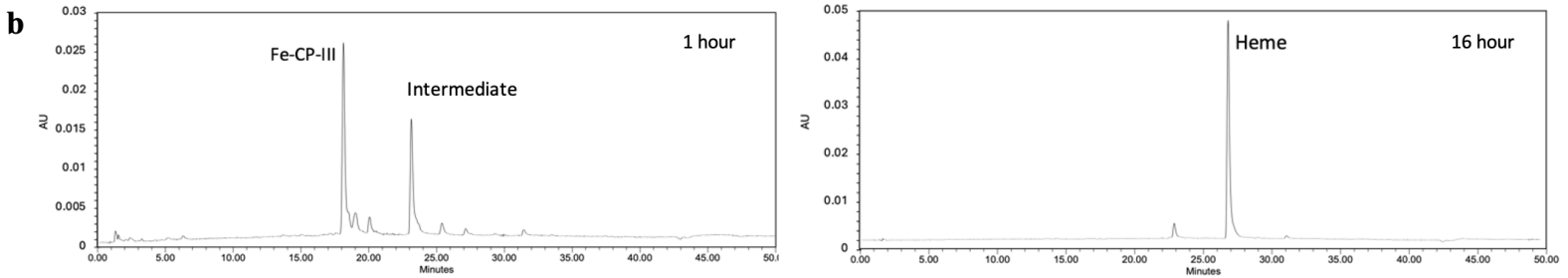


Figure 6.7: In vitro Heme synthesis from Fe-CP-III. (**a**) Tuning H_2O_2 and HemQ concentration, and buffer types types for heme production. (**b**) HPLC analysis of Fe-CP-III conversion to heme at different time points: 1-hour (top) and overnight (bottom).

6.4 Discussion

This study presents a novel, integrated two-step approach for the bio-based production of Fe-CP-III and heme using *E. coli*, combining in vivo biosynthesis with in vitro enzymatic conversion to tackle critical challenges associated with heme biosynthesis. The issues related to the intracellular accumulation of toxic metabolites, such as heme, are well-documented, highlighting the need for alternative production strategies [266]. Our strategy addresses this issue by decoupling the production into two distinct yet complementary phases. By engineering *E. coli* strains to express the CPD pathway, we achieved significant Fe-CP-III production, which was subsequently converted to heme through refined in vitro reactions. This two-step approach not only enhances heme yields but also addresses intracellular heme toxicity, offering a scalable pathway for industrial applications.

Comparative analysis of *E. coli* strains engineered for Fe-CP-III production revealed notable differences in cell growth, Fe-CP-III secretion, and metabolic flux toward heme intermediates. Among the strains tested, FECP3 performed best, with the highest extracellular Fe-CP-III concentration of 84.9 mg/L and minimal intracellular accumulation, indicating effective secretion and reduced metabolic stress. This efficiency results from the improved arrangement and expression of pathway genes, specifically positioning *hemH* at the start of the second operon under a strong promoter. This arrangement appears to enhance HemH expression and promotes effective CP-III conversion to Fe-CP-III, underscoring the importance of operon design and promoter strength in engineering high-flux pathways for targeted products. Additionally, the observed differences in secretion efficiency and metabolic stress among strains highlight the value of strain-specific engineering to address bottlenecks in porphyrin biosynthesis. A key observation in this study was the efficient extracellular production of coproheme by FECP3, which alleviated intracellular metabolic stress and facilitated downstream processing by simplifying product recovery. Such efficient secretion underscores the scalability of this system for industrial applications, where reducing intracellular toxicity and streamlining recovery processes are critical.

The production of intermediates in the Shemin/C4 pathway, including 5-ALA and PBG, provided insights into metabolic flux differences among strains. FECP3 showed the highest PBG levels, correlating with its superior Fe-CP-III production. The accumulation of by-products such as PP-IX and heme indicated minor flux diversion into PPD pathway. Addressing these imbalances could enhance Fe-CP-III yield and reduce by-product formation.

The role of *hemQ* in heme biosynthesis was further validated, with in vivo findings demonstrating heme production upon *hemQ* expression, aligning with previous reports [93]. Among engineered strains, HEM2, coexpressing hemAB-EHQ, exhibited the highest heme

production of 17.8 mg/L. HemQ from *S. aureus* showed moderate heme yields in HEM3 compared to the enzyme from *B. subtilis*, highlighting differences in enzymatic performance that merit further investigation. Intracellular heme accumulation limited overall yields, highlighting the advantage of decoupling biosynthesis and enzymatic conversion, as the in vitro phase offers a controlled environment to overcome these toxicity-related bottlenecks.

The integration of in vitro enzymatic reactions using purified HemH and HemQ enzymes allowed for precise control over reaction conditions, and also overcoming intracellular toxicity limitations, where heme production occurs solely within the cells [246, 261]. Buffer selection and pH control emerged as critical parameters in the in vitro conversion of CP-III to Fe-CP-III and Fe-CP-III to heme. The use of citrate phosphate buffer at pH 8 proved most effective for Fe-CP-III production. An alkaline pH is well-aligned with the reaction's chemistry, effectively neutralizing the two protons (H^+) released during the conversion of CP-III to Fe-CP-III [91]. This ensures that the buffer maintains enzyme stability and activity and supports the reaction's efficient progression without being disrupted by the increasing acidity. Similarly, sodium phosphate buffer was most suitable for HemQ-mediated heme synthesis. HemH exhibited broad substrate specificity, converting both CP-I and CP-III to their respective iron-chelated forms with high yields. This versatility highlights its potential application in synthesizing diverse porphyrin derivatives. In addition, tuning the substrate ratio also significantly enhanced the efficiency of the in vitro reactions. Adjusting the Fe^{2+}/CP-III and Fe-CP-III/H_2O_2 ratios and enzyme concentrations emphasized the critical need to balance substrate levels to avoid inhibition and achieve a higher efficiency. For instance, excessive Fe^{2+} concentrations resulted in reduced yields. Similarly, we observed heme degradation in the presence of excessive hydrogen peroxide, underscoring its dual role as both a substrate and an oxidizing agent. Note that during in vivo cultivations, H_2O_2 is not externally supplemented, as it is naturally produced under aerobic conditions [267], particularly during the enzymatic conversion of CPG-III to CP-III [91]. Future studies could focus on strategies to mitigate heme degradation, such as refining reaction kinetics or exploring alternative oxidants.

In our in vitro reactions, while we frequently achieved high conversion rates of up to 100%, the yield of porphyrins in the same reactions, particularly those involving HemQ, was notably lower. One possible explanation for this discrepancy is the degradation of heme in the presence of hydrogen peroxide, as previously reported [93, 268]. Another contributing factor could be the incomplete release of heme from the HemQ enzyme. HemQ may retain heme within its active site or form stable intermediates that are not easily released, resulting in lower detectable yields of free heme [265, 268]. This phenomenon has also been reported in other enzymatic reactions.

A notable advantage of this system is its scalability. The use of cell-free cultivation

broth as a substrate for HemH and HemQ minimized substrate purification requirements, significantly reducing production costs. This demonstrates the robustness of the system and its adaptability for synthesizing other high-value compounds, such as chlorophyll precursors or pigments. Decoupling in vivo and in vitro steps allows fine-tuning at each stage, offering a modular and versatile framework that serves as a sustainable alternative to traditional chemical synthesis, which typically involves harsh conditions and generates toxic intermediates.

Despite the promising results, some limitations were observed in the study. For instance, while the cell-free conversion of Fe-CP-III to heme was efficient, the overall yields were still dependent on the substrate concentration and enzyme loading. Enhancing enzyme stability through protein engineering or immobilization could improve yields and reduce enzyme consumption.

Among the highest reported heme titers, Choi et al. (2022) achieved 1.03 g/L of heme in *E. coli* through systematic fermentation optimization, including carbon source selection, iron supplementation, and controlled induction timing [40]. In contrast, Chen et al. (2024) implemented a simialr two-step strategy by first producing 16.5 g/L of coproporphyrin III (CPIII) in *Rhodobacter sphaeroides*, followed by in vitro enzymatic conversion to yield 10.8 g/L of heme [82], supporting the validity of our proposed strategy and highlighting the potential of integrating microbial biosynthesis with biocatalysis. As these studies differ in microbial hosts, cultivation modes (batch vs. fed-batch), and production methodologies, comparisons should be interpreted in the context of their distinct experimental frameworks.

This work advances synthetic biology and metabolic engineering by introducing a scalable and adaptable two-step production process that decouples intracellular biosynthesis from downstream enzymatic conversion. This modular framework, demonstrated for Fe-CP-III and heme production, enables fine-tuning at each stage, supports scalability for industrial applications, and offers a sustainable alternative to chemical synthesis. By integrating metabolic engineering, synthetic biology, and cell-free systems, this approach sets a precedent for sustainable biomanufacturing and highlights the potential of microbial hosts for high-value compound production [269].

Chapter 7

Manipulating Intracellular Oxidative Conditions to Enhance Porphyrin Production in *Escherichia coli*

7.1 Introduction

Porphyrins are essential tetrapyrrolic compounds that serve as precursors for the biosynthesis of heme, chlorophyll, and other biologically critical molecules [270, 271, 272]. In the Shemin/C4 pathway, which is primarily found in mammals, yeast, fungi, and α-proteobacteria [273], porphyrin biosynthesis begins with the condensation of glycine and succinyl-CoA by 5-aminolevulinic acid synthase (HemA) [274, 275, 276] (Figure 7.1). Then, two molecules of 5-aminolevulinic acid (5-ALA) are fused by porphobilinogen synthase (HemB) to form porphobilinogen (PBG) [84]. Four PBG units are subsequently polymerized by porphobilinogen deaminase (HemC) to generate hydroxymethylbilane (HMB), a linear tetrapyrrole [11]. At the HMB branch point, the porphyrin biosynthetic pathway diverges into two distinct routes, leading to the formation of two stereoisomers of uroporphyrinogen I (UPG-I) and uroporphyrinogen III (UPG-III). UPG-I is formed via the spontaneous and non-enzymatic cyclization of HMB [91, 277], whereas UPG-III is synthesized enzymatically via uroporphyrinogen III synthase (HemD) [278, 252]. Both uroporphyrinogen stereoisomers can undergo further oxidation to form their respective uroporphyrins, i.e., uroporphyrin I (UP-I) and uroporphyrin III (UP-III) [279]. However, the metabolic fates of these stereoisomers differ significantly. In the non-enzymatic branch, UPG-I is converted to coproporphyrinogen I (CPG-I) in a reaction involving uroporphyrinogen decarboxylase (HemE) [280] and is subsequently oxidized to coproporphyrin I (CP-I). CP-I

is a metabolic dead-end and cannot be further processed, making it the final product of type-I sterioisomers [279]. In contrast, UPG-III enters the enzymatic branch and undergoes a series of reactions catalyzed by HemE, HemF, HemG, and HemH, ultimately leading to the formation of heme [10, 281] (Figure 7.1).

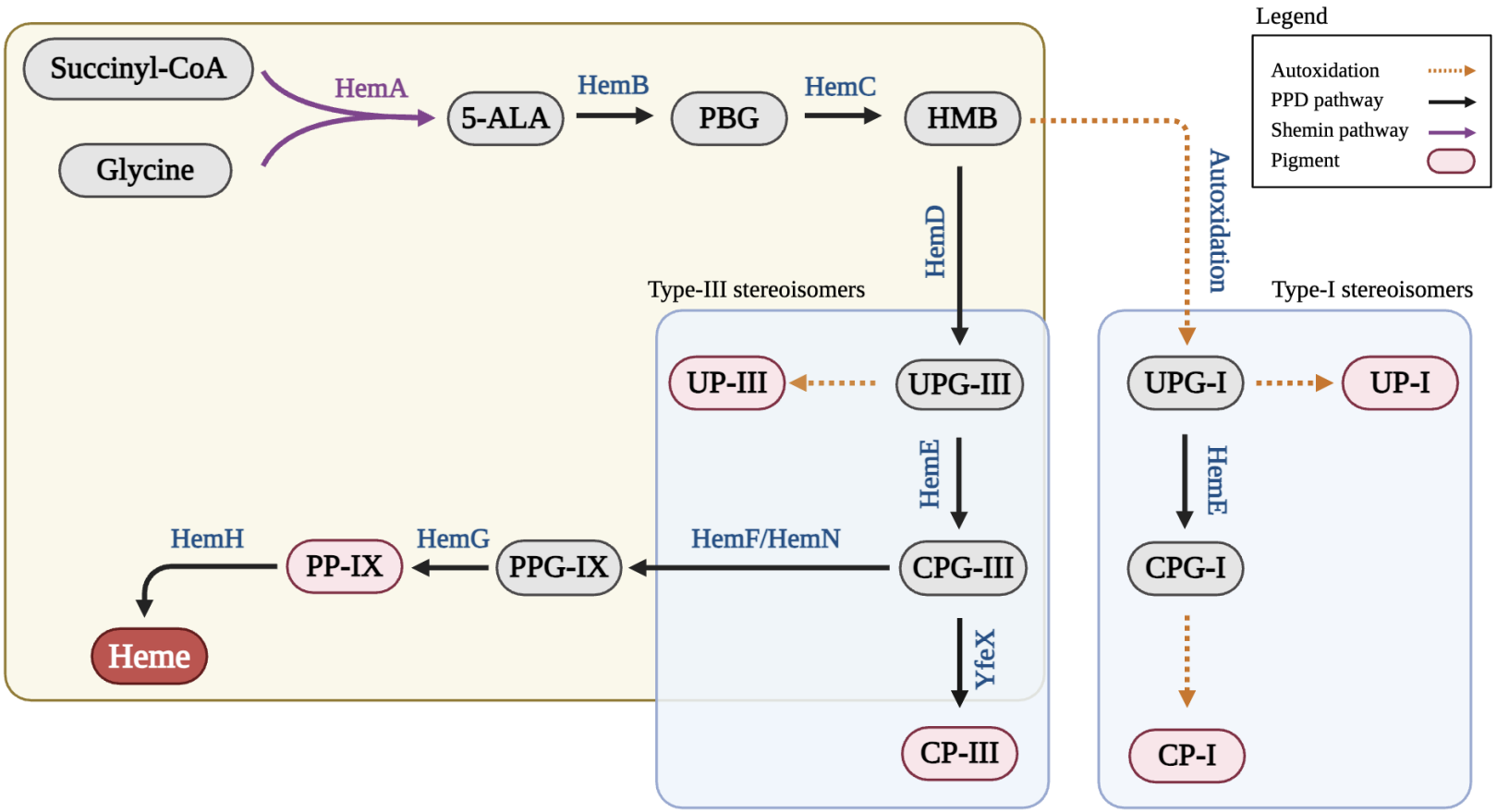


Figure 7.1: The Shemin/C4 pathway for porphyrin biosynthesis from succinyl-CoA and glycine. 5-ALA, 5-aminolevulinic acid; CPG-I, coproporphyrinogen I; CPG-III, coproporphyrinogen III; CP-I, coproporphyrin I; CP-III, coproporphyrin III; HemA, 5-aminolevulinate synthase; HemB, porphobilinogen synthase; HemC, porphobilinogen deaminase; HemD, uroporphyrinogen III synthase; HemE, uroporphyrinogen decarboxylase; HemF, coproporphyrinogen III oxidase; HemG, protoporphyrinogen oxidase; HemH, protoporphyrin ferrochelatase; HemN, oxygen-independent coproporphyrinogen III oxidase; HMB, Hydroxymethylbilane; PBG, porphobilinogen; PP-IX, protoporphyrin IX; PPG-IX, protoporphyrinogen IX; UP-I, uroporphyrin I; UP-III, uroporphyrin III; UPG-I, uroporphyrinogen I; UPG-III, uroporphyrinogen III; YfeX, porphyrinogen peroxidase.

In Chapters 4 and 5, we reported on strain engineering of bacterial *Escherichia coli* for porphyrin production via the Shemin/C4 pathway. A metabolic strategy adapted for strain engineering in these studies was the redirection of dissimilated carbon flux toward succinyl-CoA in the citric acid cycle by introducing two mutations of *iclR* and *sdhA*, such that engineered strains with these two mutations can be cultivated aerobically for biosynthesis of succinyl-CoA-derived compounds. However, bacterial cultivation under aerobic conditions might somehow favor the auto-oxidation branch at the metabolic node of HMB, limiting the flux into the HemD-enzymatic branch and the subsequent biosynthesis of type-III stereoisomeric species [82]. In fact, cultivation of the engineered *E. coli* strain for UP biosynthesis led to the production of a stereoisomeric mixture with UP-I being dominant. Therefore, redirecting more dissimilated carbon flux toward the HemD-enzymatic branch is highly desirable to enhance the biosynthesis of type-III porphyrins (i.e., UP-III and CP-III) and even heme.

The auto-oxidation at the HMB node is a non-enzymatic process potentially intensified by the presence of reactive oxygen species (ROS) such as superoxide anions, hydrogen peroxide, and hydroxyl radicals [282] which are commonly produced by cells during cellular metabolism under aerobic conditions [283]. Since ROS favor the oxidation of pathway intermediates, their formation can potentially shift more dissimilated carbon flux toward the type I branch and, therefore, the production of UPG-I over UPG-III. This imbalance caused by ROS highlights the critical need to regulate intracellular oxidative stress and, therefore, ensure a more favorable porphyrin metabolism and flux channeling into the pathway toward type-III stereoisomeric species.

ROS are an unavoidable byproduct of aerobic metabolism, particularly under the conditions of high agitation and/or aerobiosis during cultivation [284]. Lowering the aeration in microbial cultures can reduce dissolved oxygen levels and, consequently, the generation of ROS. An alternative strategy to mitigate oxidative stress involves the use of antioxidants, either enzymatic or non-enzymatic ones. Antioxidants are compounds that can delay or inhibit oxidative processes by neutralizing ROS or preventing their formation [285]. Non-enzymatic antioxidants, such as ascorbic acid (vitamin C), are particularly effective due to their dual role as ROS scavengers and stimulators of enzymatic antioxidants [286]. Ascorbic acid, a water-soluble and cost-effective compound, neutralizes ROS by forming ascorbyl radicals, preventing the formation of secondary free-radical byproducts [287]. On the other hand, enzymatic antioxidants, including superoxide dismutases (SODs) and catalases (KATs), play a pivotal role in maintaining redox balance in *E. coli* [288]. SODs act as the first line of defense by converting superoxide anions ($O_2^{\bullet -}$) into hydrogen peroxide (H_2O_2) and molecular oxygen (O_2), Equation 7.1, reducing the potential for oxidative damage caused by superoxide radicals [289]. In *E. coli*, three distinct SOD enzymes are

encoded, i.e., SodA (Mn-SOD), SodB (Fe-SOD), and SodC (Cu,Zn-SOD) [290]. SodA and SodB function in the cytoplasm, where they neutralize intracellular superoxide [291, 292], whereas SodC operates in the periplasm to manage oxidative stress in that compartment [293]. Among these, SodA has been shown to be particularly effective in preventing DNA damage [294].

$$2O_2^{\bullet -} + 2H^+ \rightarrow H_2O_2 + O_2 \tag{7.1}$$

Catalases complement the action of SODs by decomposing hydrogen peroxide, a byproduct of SOD activity, into water and oxygen, thereby preventing its accumulation (Equation 7.2) [295]. *E. coli* has two cytoplasmic catalases: KatE, a monofunctional catalase expressed during the stationary phase, and KatG, a bifunctional catalase–peroxidase active during exponential growth [296, 297, 298, 299]. The coordinated activity of these enzymes ensures that ROS levels remain tightly controlled, with superoxide anions concentrations maintained below 0.1 nM and hydrogen peroxide levels below 20 nM [300, 301]. Overexpression of both SODs and catalases has been shown to enhance oxidative stress resistance, providing a robust defense system that protects cellular components and stabilizes key biosynthetic processes.

$$2H_2O_2 \rightarrow 2H_2O + O_2 \tag{7.2}$$

In this study, we explored both antioxidant supplementation and overexpression of ROS-scavenging enzymes to mitigate the effects of oxidative stress and enhance the production of type III porphyrins in *E. coli*. By implementing these bioprocessing strategies, we aim to develop a more efficient and sustainable platform for the bio-based production of porphyrins and their derivatives, as well as address key challenges associated with oxidative stress and metabolic pathway manipulation.

7.2 Materials and Methods

7.2.1 Bacterial Strains and Plasmids

Details of the bacterial strains and plasmids used in this study are provided in Table 7.1, while the primer sequences are listed in Table S7.1. Taq DNA polymerase was sourced from New England Biolabs (Ipswich, MA, USA). Genomic DNA extraction from bacterial cells was carried out using the Qiagen Blood & Tissue DNA Isolation Kit (Hilden, Germany),

and plasmid purification was performed using the Qiagen Miniprep kit. The CPC-Sbm-derived host cells used in this study were developed from the CPC-Sbm strain, which carries targeted mutations in the *sdhA* and/or *iclR* genes [200]. This strain originates from the *E. coli* BW25113 background, where the ldhA gene had been previously inactivated [184]. For molecular cloning purposes, *E. coli* HI-Control 10G (Lucigen, Middleton, WI, USA) was employed. Plasmid construction was carried out using the Gibson assembly method [181], and DNA sequencing services were provided by Plasmidsaurus (Eugene, OR, USA). Oligonucleotide synthesis was conducted by Integrated DNA Technologies (IDT) (Coralville, IA, USA).

Plasmids pK-hemA, pK-hemAB, pK-hemABC, pK-hemABCD, and pK-hemAB-E were used from Chapters 4 and 5.

pK-hemABsodABC was constructed by amplifying *hemAB* and amplifying the backbone using primer sets P1/P2 and P3/P4 with pK-hemABCD as the template. The *sodA*, *sodB*, and *sodC* genes were amplified from the genomic DNA of *E. coli* MG1655 using primer sets P5/P6, P7/P8, and P9/P10. These five fragments were Gibson assembled to form pK-hemABsodABC. For effective coexpression, all genes were aligned to form an operon *hemABsodABC* regulated by a common strong *trc* promoter with an individual, strong RBS for each gene.

pK-hemABsodA was constructed by amplifying *hemAB* and amplifying the backbone using primer sets P1/P2 and P11/P4 with pK-hemABCD as the template. The *sodA* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P5/P12. These three fragments were Gibson assembled to form pK-hemABsodA with all genes arranged on a single operon regulated by a common strong *trc* promoter.

Similarly, pK-hemABsodB was constructed by amplifying *hemAB* and amplifying the backbone using primer sets P1/P2 and P13/P4 with pK-hemABCD as the template. The *sodB* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P14/P15. These three fragments were Gibson assembled to form pK-hemABsodB with all genes arranged on a single operon regulated by a common strong *trc* promoter.

pK-hemABsodC was constructed by amplifying *hemAB* and amplifying the backbone using primer sets P1/P2 and P3/P4 with pK-hemABCD as the template. The *sodC* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P16/P10. These three fragments were Gibson assembled to form pK-hemABsodB with all genes arranged on a single operon regulated by a common strong *trc* promoter.

pK-hemABsodABC-katEG was constructed by amplifying *hemAB* and amplifying the backbone using primer sets P1/P2 and P17/P4 with pK-hemABCD as the template. The *sodABC* genes were amplified from pK-hemABsodABC using the primers P18/P19. The

*grac*max promoter was amplified from pK-hemAB-E using the primers P20/P21. The *katE* and *katG* genes were amplified from the genomic DNA of *E. coli* MG1655 using primer sets P22/P23 and P24/P25. These six fragments were Gibson assembled to form pK-hemABsodABC-katEG, and for effective expression, *hemABsodABC* were arranged on the first operon regulated by a common strong *trc* promoter, and *katEG* was arranged on the second operon regulated by the strong *grac*max promoter.

pK-hemAB-katE was constructed by amplifying *hemAB* along with the *grac*max promoter and amplifying the backbone using primer sets P26/P21 and P27/P4 and pK-hemAB-E as the template. The *katE* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P22/P28. These three fragments were Gibson assembled to form pK-hemAB-katE with *hemAB* arranged on the first operon regulated by a common strong *trc* promoter and *katE* on the second operon regulated by the strong *grac*max promoter.

Similarly, pK-hemAB-katG was constructed by amplifying *hemAB* along with the *grac*max promoter and amplifying the backbone using primer sets P26/P29 and P17/P4 with pK-hemAB-E as the template. The *katG* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P30/P25. These three fragments were Gibson assembled to form pK-hemAB-katG with *hemAB* arranged on the first operon regulated by a common strong *trc* promoter and *katG* on the second operon regulated by the strong *grac*max promoter.

pK-hemABsodA-katE was constructed by amplifying *hemABsodA* and amplifying the backbone using primer sets P26/P31 and P27/P4 with pK-hemABsodA as the template. The *grac*max promoter along with the *katE* gene were amplified from pK-hemABsodABC-katEG using the primers P32/P28. These three fragments were Gibson assembled to form pK-hemABsodA-katE with *hemABsodA* arranged on the first operon regulated by a common strong *trc* promoter and *katE* on the second operon regulated by the strong *grac*max promoter.

Similarly, pK-hemABsodA-katG was constructed by amplifying *hemABsodA* and amplifying the backbone using primer sets P26/P31 and P17/P4 with pK-hemABsodA as the template. The *grac*max promoter was amplified from pK-hemABsodABC-katEG using the primers P32/P29. The *katG* gene was amplified from the genomic DNA of *E. coli* MG1655 using primers P30/P25. These four fragments were Gibson assembled to form pK-hemABsodA-katG with *hemABsodA* arranged on the first operon regulated by a common strong *trc* promoter and *katG* on the second operon regulated by the strong *grac*max promoter.

Table 7.1: Strains and plasmids used in this study.

Name	Description or Relevant Genotype	Source
Host strains		
HI-Control 10G	*mcrA*, Δ(*mrr-hsdRMS-mcrBC*),*endA1*, *recA1*, ϕ80*dlacZ*ΔM15, Δ*lacX74*, *araD139*,Δ(*ara leu*)*7697*, *galU*, *galK*, *rpsL* (Str^R), *nupG*, λ^-, *tonA*, Mini-F $lacI^{q1}$ ($Gent^R$)	Lucigen
MG1655	K-12; F^-, λ^-, *rph-1*	Lab stock
CPC-SbmΔ*iclR*Δ*sdhA*	F^-, Δ(*araD-araB*)567, Δ*lacZ4787(::rrnB-3)*, λ^-, *rph-1*, Δ(*rhaD-rhaB*)568, *hsdR514*, Δ*ldhA*, P_{trc}::*sbm*, Δ*iclR*, Δ*sdhA*	[200]
DMB	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemAB	Chapter 4
SOD1	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodABC	This study
SOD2	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodA	This study
SOD3	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodB	This study
SOD4	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodC	This study
KAT1	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodABC-katEG	This study
KAT2	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemAB-katE	This study
KAT3	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemAB-katG	This study
KAT4	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodA-katE	This study
KAT5	CPC-SbmΔ*iclR*Δ*sdhA*/pK-hemABsodA-katG	This study
Plasmids		
pK-hemA	p15A ori, Km^R, P_{trc}::*hemA*	Chapter 4
pK-hemAB	p15A ori, Km^R, P_{trc}::*hemAB*	Chapter 4
pK-hemABC	p15A ori, Km^R, P_{trc}::*hemABC*	Chapter 4
pK-hemAB-E	p15A ori, Km^R, P_{trc}:: *hemAB* -$P_{gracmax}$::*hemE*	Chapter 5
pK-hemABsodABC	p15A ori, Km^R, P_{trc}::*hemABsodABC*	This study
pK-hemABsodA	p15A ori, Km^R, P_{trc}::*hemABsodA*	This study
pK-hemABsodB	p15A ori, Km^R, P_{trc}::*hemABsodB*	This study
pK-hemABsodC	p15A ori, Km^R, P_{trc}::*hemABsodC*	This study
pK-hemABsodABC-katEG	p15A ori, Km^R, P_{trc}:: *hemABsodABC*-$P_{gracmax}$::*katEG*	This study
pK-hemAB-katE	p15A ori, Km^R, P_{trc}:: *hemAB*-$P_{gracmax}$::*katE*	This study
pK-hemAB-katG	p15A ori, Km^R, P_{trc}:: *hemAB*-$P_{gracmax}$::*katG*	This study
pK-hemABsodA-katE	p15A ori, Km^R, P_{trc}:: *hemABsodA*-$P_{gracmax}$::*katE*	This study
pK-hemABsodA-katG	p15A ori, Km^R, P_{trc}:: *hemABsodA*-$P_{gracmax}$::*katG*	This study

7.2.2 Media and Bacterial Cell Cultivation

All media components were obtained from Sigma-Aldrich Co. (St. Louis, MO, USA), except for yeast extract and tryptone, which were obtained from BD Diagnostic Systems (Franklin Lakes, NJ, USA). *E. coli* strains were stored as glycerol stocks at -80 °C and streaked onto lysogeny broth (LB) agar plates, prepared with 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl. Plates were incubated at 37 °C for 14–16 h. For bioreactor experiments, single colonies were picked from LB agar plates and inoculated into 12 mL of super broth (SB) medium (32 g/L tryptone, 20 g/L yeast extract, and 5 g/L NaCl) in a 125 mL Erlenmeyer flask. This culture was grown at 37 °C and 280 rpm on a rotary shaker (New Brunswick Scientific, Edison Township, NJ, USA) for 4–6 h and was then used to inoculate 220 mL of SB medium at 2% (v/v) in a 1 L conical flask. The seed culture was incubated overnight (14–16 h) under the same conditions. Cells were harvested by centrifugation at 4500× g and 20 °C for 8 min and then resuspended in 40 mL of fresh SB medium. The resuspended culture was used to inoculate a stirred tank bioreactor (CelliGen 115, Eppendorf AG) containing 0.8 L of working volume, set at 37 °C and stirred at 430 rpm. The semi-defined medium used in the batch bioreactor consisted of 30 g/L glycerol, 0.23 g/L K_2HPO_4, 0.51 g/L NH_4Cl, 49.8 mg/L $MgCl_2$, 48.1 mg/L K_2SO_4, 1.52 mg/L $FeSO_4$, 0.055 mg/L $CaCl_2$, 2.93 g/L NaCl, 0.72 g/L tricine, 10 g/L yeast extract, 10 mM $NaHCO_3$, and 1 mL/L trace elements (2.86 g/L H_3BO_3, 1.81 g/L $MnCl_2{\cdot}4H_2O$, 0.222 g/L $ZnSO_4{\cdot}7H_2O$, 0.39 g/L $Na_2MoO_4{\cdot}2H_2O$, 79 μg/L $CuSO_4{\cdot}5H_2O$, and 49.4 μg/L $Co(NO_3)_2{\cdot}6H_2O$) [224]. To induce protein expression, 0.05 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added. To prevent glycine limitation during cultivation, 2 g of glycine was supplemented into the bioreactor approximately 30 h after inoculation. For certain batches, ascorbic acid was included in the medium at concentrations between 0.5 and 2 g/L. Aerobic conditions were maintained by purging air into the bioreactor at a rate of 1 volume of air per volume of liquid per minute (vvm). The pH of the culture was controlled at 7.0 ± 0.1 by automatic addition of 3 M NH_4OH or 3 M H_3PO_4.

7.2.3 Analysis

Cell density at OD_{600} was measured by washing all culture samples once with a 0.15 M saline solution, diluting them appropriately, and analyzing them using a GENESYSTM 40/50 Vis/UV-Vis spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). For the preparation of cell-free medium, culture samples were centrifuged at 17,000× g for 1 min and then filtered through a 0.2 μm syringe filter to ensure sterility. High-performance liquid chromatography (HPLC) was used to analyze extracellular metabolites and glycerol, employing a Shimadzu LC-10AT system (Shimadzu, Kyoto, Japan) with a refractive index

detector (RID-10A, Shimadzu) and an Aminex HPX-87H chromatographic column (Bio-Rad Laboratories, Hercules, CA, USA). The column temperature was maintained at 35 °C, and the mobile phase consisted of 5 mM H2SO4 (pH 2), running at a flow rate of 0.6 mL/min. Data from the RID were processed using Clarity Lite software (v. 7.4.1.88, DataApex, Prague, Czech Republic).

Levels of 5-ALA and PBG in the cell-free medium were quantified using a modified Ehrlich's reagent assay [87]. Porphyrins were analyzed using a Waters™ 2690 separation module equipped with a photodiode array (PDA) detector (2996 PDA detector, Waters™, Milford, MA, USA) and a Chromolith® HighResolution RP-18 endcapped column (Supelco, Darmstadt, Germany). UV absorbance was detected at 400 nm, and data were processed using Empower 3 software (Waters™, Milford, MA, USA). The mobile phase system used for porphyrin analysis followed a previously published method [11] with minor modifications.

The percentage yield of UP was determined by calculating the molar ratio of the produced UP to the theoretical maximum production, based on the amount of glycerol consumed. Statistical analyses were performed to compare the UP-III/UP-I ratio under different conditions, and the detailed results are provided in the Supplementary Materials.

7.3 Results

7.3.1 Effects of Ascorbic Acid Supplementation on UP Biosynthesis

To evaluate the effects of antioxidant supplementation on limiting flux toward the auto-oxidation pathway, we utilized DMB, a UP-producing strain heterologously co-expressing *hemA* from *R. sphaeroides* and *hemB* from *E. coli*, as described in Chapter 4. Cultivation of DMB was conducted with ascorbic acid at 0, 0.5, 1, and 2 g/L, respectively.

Enhanced glycerol consumption was observed with ascorbic acid supplementation (Figure 7.2a–d). Maximum cell density (14.4 OD_{600}) was observed with ascorbic acid concentrations ranging from 0 to 1 g/L, whereas increasing the concentration to 2 g/L slightly reduced the maximum cell density to 12.2 OD_{600}. Acetate profile followed a similar increasing trend across all culture conditions, with final titers reaching 14.5 g/L by the end of cultivation. In contrast, compared to the control, the succinate accumulation profile was significantly reduced in cultivations supplemented with ascorbic acid.

Time profiles of 5-ALA and PBG showed a similar trend with peaks occurring at 48 h across all conditions (Figure 7.2e–h). However, cultivations supplemented with ascorbic acid yielded slightly lower PBG titers of 1600 mg/L, compared to 1995 mg/L (193 mg/OD_{600}/L) for the control cultivation without ascorbic acid.

While the total UP production was slightly reduced with the addition of ascorbic acid, the UP distribution was more favorable toward UP-III with the UP-III/UP-I ratios of 0.6, 1.4, 2.6, and 1.5 for cultivations with 0, 0.5, 1, and 2 g/L of ascorbic acid, respectively (Figure 7.2i). Notably, cultivation with 1 g/L ascorbic acid had 215 mg/L (1.3% yield) UP-I and 552 mg/L (3.3% yield) UP-III, indicating a significant shift in dissimilated carbon flux toward the enzymatic (HemD) pathway. However, increasing the ascorbic acid concentration beyond 1 g/L did not further increase the UP-III/UP-I ratio but led to reduced cell growth.

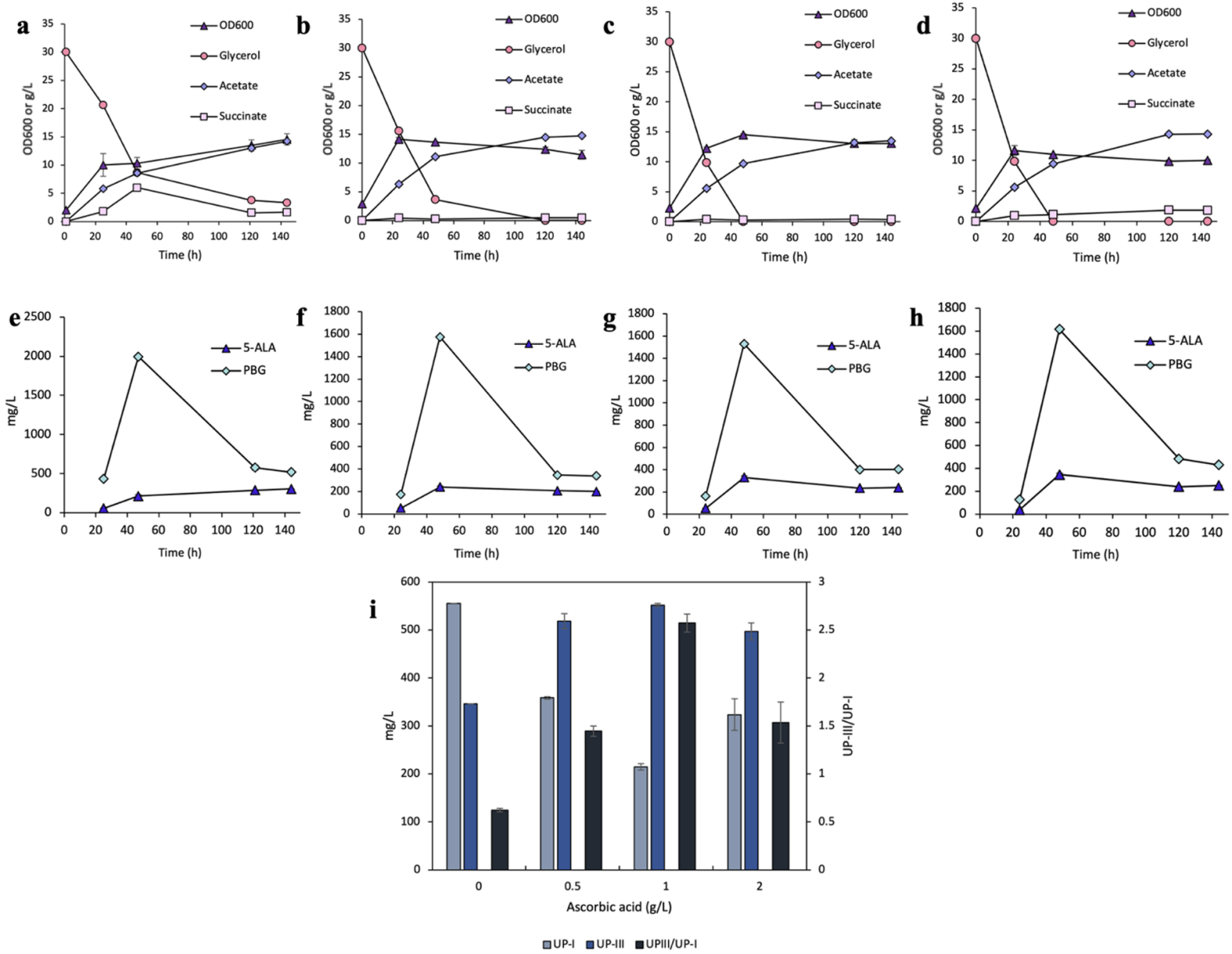


Figure 7.2: Aerobic bioreactor cultivation of DMB for UP biosynthesis with ascorbic acid supplementation. Time profiles of (**a**–**d**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**e**–**h**) 5-ALA and PBG biosynthesis, and (**i**) UP biosynthesis at 144 h. All values are reported as means ± SD (n = 2).

7.3.2 Effects of SODs on UP Biosynthesis

To assess the effects of SODs in redirecting dissimilated carbon flux toward the enzymatic (HemD) pathway, we engineered a UP-producing strain, SOD1, by heterologously co-expressing *hemA* from *R. sphaeroides* along with *hemB* and *sodABC* from *E. coli*. These genes were arranged in a single operon under the control of a strong *trc* promoter, with each gene translationally regulated by a strong RBS. Unexpectedly, this strain did not exhibit any pigmentation, suggesting that the overexpression of *sodABC* might somehow interfere with porphyrin production. To identify the specific gene(s) responsible for the interference, we constructed three additional strains, each co-expressing *hemAB* and an individual *sod* gene, i.e., *sodA* (SOD2), *sodB* (SOD3), or *sodC* (SOD4). Notably, SOD4 showed no pigmentation, indicating that the overexpression of *sodC* was likely responsible for interfering with porphyrin production.

Both strains of SOD2 and SOD3 exhibited a similar and slow glycerol consumption pattern, and glycerol was completely depleted by 120 h (Figure 7.3a,b). Cell density for SOD2 was 14.4 OD_{600} at 120 h, while SOD3 reached 13.4 OD_{600} at 48 h. SOD2 produced 10.4 g/L acetate within the first 48 h, followed by diauxic growth with the accumulated acetate being depleted by the end of the cultivation. On the other hand, SOD3 exhibited consistent acetate accumulation throughout the cultivation, reaching 17.7 g/L by 120 h. Both strains showed minimal succinate production.

The time profiles of 5-ALA and PBG for SOD2 and SOD3 displayed a similar trend, with titers peaking at 48 h (Figure 7.3c,d). 5-ALA titers reached 469 and 217 mg/L (33.9 and 16.1 mg/OD_{600}/L), while PBG titers reached 2215 mg/L and 316 mg/L (160 and 86.4 mg/OD_{600}/L) in SOD2 and SOD3, respectively.

The final UP production at 144 h for SOD2 was 531 mg/L (3.1% yield) UP-I and 1029 mg/L (6.1% yield) UP-III, respectively, achieving a UP-III/UP-I ratio of 1.9 (Figure 7.3g). This represented a significant shift in dissimilated carbon flux from the auto-oxidation pathway to the enzymatic (HemD) pathway compared to the control DMB strain. Additionally, SOD2 achieved a high total UP titer of 1.56 g/L (9.2% yield), 72.9% higher than the control DMB strain. On the other hand, SOD3 produced 219.9 mg/L (1.3% yield) UP-I and 385.3 mg/L (2.3% yield) of UP-III, with a UP-III/UP-I ratio of 1.8 (Figure 7.3g). While SOD3 had a high UP-IIII/UP-I ratio similar to SOD2, the total UP titer was significantly lower than SOD2.

In another experiment, we combined the above non-enzymatic and enzymatic strategies by supplementing 1 g/L ascorbic acid upon cultivating SOD2. The resulting cultivation had a slightly reduced cell density and slower glycerol consumption (Figure 7.3c) compared to the control SOD2 cultivation without ascorbic acid. Additionally, this strain exhibited

significant acetogenesis with the acetate titer reaching 14.4 g/L by the end of cultivation, suggesting that ascorbic acid suppressed the diauxic growth observed in the control SOD2 cultivation. Time profiles for 5-ALA and PBG had peaks of 361 mg/L (27.2 mg/OD_{600}/L) and 1814 mg/L (137 mg/OD_{600}/L), respectively (Figure 7.3d), both lower than those in the control SOD2 cultivation. This cultivation produced 477 mg/L (2.8% yield) UP-I and 674 mg/L (4.0% yield) UP-III, with a UP-III/UP-I ratio of 1.4 (Figure 7.3g). These results suggest that there were no combined or synergistic effects for the two enzymatic or non-enzymatic strategies, although the possibility of such interactions cannot be ruled out without further analysis.

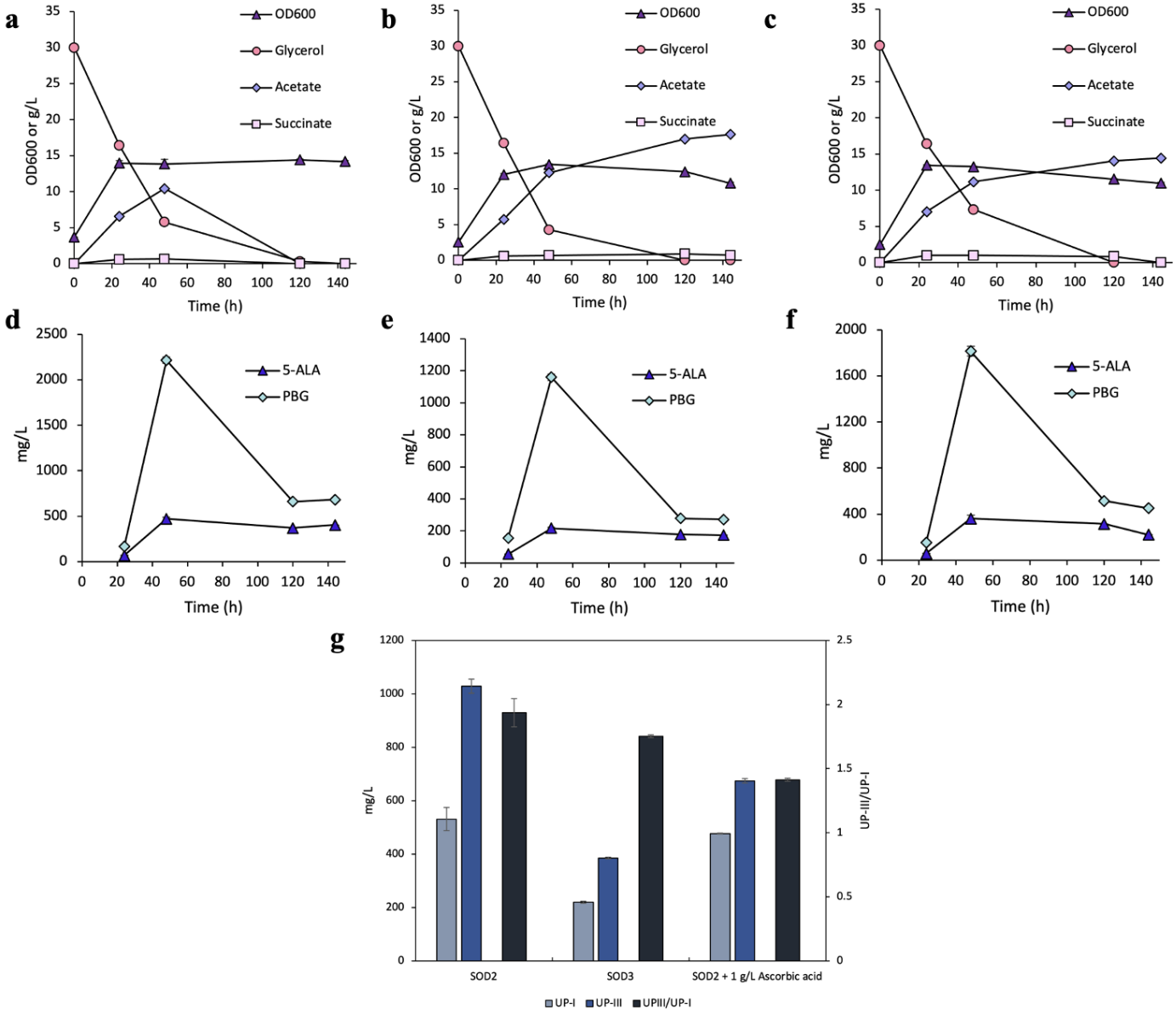


Figure 7.3: Aerobic bioreactor cultivation of SOD2 and SOD3 for UP biosynthesis with ascorbic acid supplementation. Time profiles of (**a**–**c**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**d**–**f**) 5-ALA and PBG biosynthesis, and (**g**) UP biosynthesis at 144 h. All values are reported as means ± SD (n = 2).

7.3.3 Effects of Catalases on UP Biosynthesis

To investigate the effects on UP biosynthesis, we initially constructed a KAT1 strain, which contains the first operon for co-expressing *hemAB* and *sodABC* under the control of the *trc* promoter and the second operon for co-expressing *katEG* under the control of the *gracmax* promoter. This expression strategy was designed to integrate the superoxide-neutralizing activity of SOD-ABC with the H_2O_2-scavenging capacity of KAT-EG. However, KAT1 only produced a limited amount of porphyrin with minimal pigmentation, prompting further investigation into the effects from individual *kat* genes. Hence, we constructed the strains KAT2 and KAT3, which expressed *katE* and *katG* in the plasmids pK-HemAB-KatE and pK-HemAB-KatG, respectively. To investigate the combined effects of *kat* and *sodA* genes, we generated additional strains, KAT4 and KAT5, harboring the plasmids pK-HemABSodA-KatE and pK-HemABSodA-KatG, respectively.

All KAT strains, except KAT4, had enhanced cell growth, achieving maximum cell densities between 20 and 25 OD_{600}, with distinct diauxic growth (Figure 7.4a–e). KAT4 had a lower cell density of 17.2 OD_{600}, with notable acetogenesis. Minimal 5-ALA and PBG were produced across all strains except KAT4 (Figure 7.4f–j). Specifically, KAT4 produced 393 mg/L (22.8 mg/OD_{600}/L) 5-ALA at 144 h and 1591 mg/L (121 mg/OD_{600}/L) PBG at 48 h (Figure 7.4i).

Unexpectedly, KAT2, KAT3, and KAT5 produced negligible UP, with minimal UP-I being detected (Figure 7.4k). While KAT1 had a remarkable UP-III/UP-I ratio of 6.2, its total UP production was less than 40 mg/L. On the other hand, KAT4 had substantial UP production comparable to SOD2, with a high total UP titer of 1.58 g/L, consisting of 693 mg/L (4.1% yield) UP-I and 890 mg/L (5.3% yield) UP-III, and a decent UP-III/UP-I ratio of 1.3.

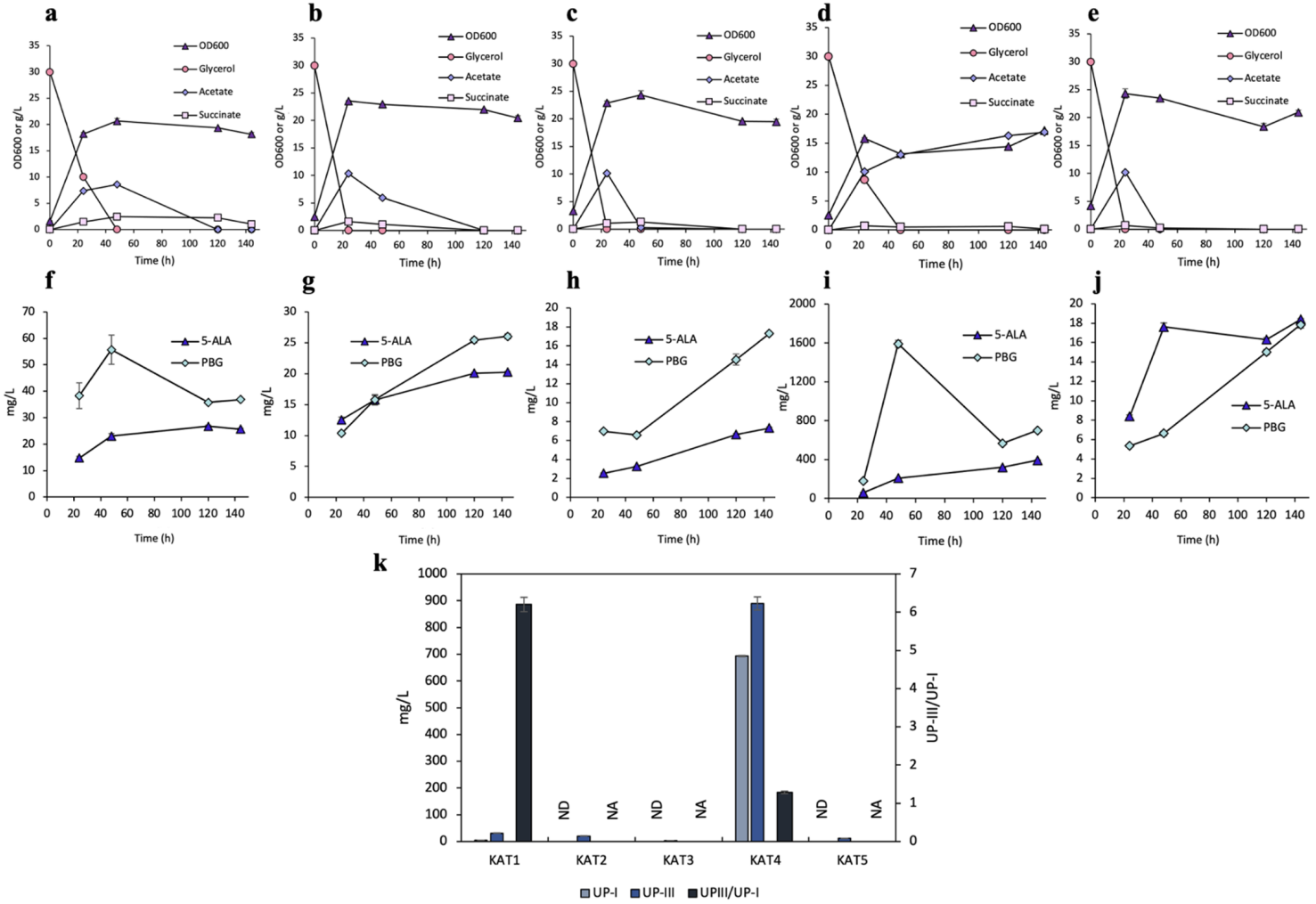


Figure 7.4: Aerobic bioreactor cultivation of KAT1, KAT2, KAT3, KAT4, and KAT5 for UP biosynthesis. Time profiles of (**a**–**e**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation, (**f**–**j**) 5-ALA and PBG biosynthesis, and (**k**) UP biosynthesis at 144 h. All values are reported as means ± SD (n = 2).

7.4 Discussion

In Chapter 4, we demonstrated that while HemABCD are the enzymes involved in UP biosynthesis, overexpression of *hemAB* alone was sufficient for effective UP biosynthesis. However, the majority of the produced UP was the stereoisomer UP-I rather than UP-III, primarily due to the more effective auto-oxidation pathway rather than the HemD-enzymatic pathway at the metabolic node of HMB (Figure 7.1) under the aerobic cultivation conditions for porphyrin biosynthesis [282]. In this study, we explored redirecting the dissimilated carbon flux more toward the HemD-enzymatic pathway to favor UP-III production while limiting the auto-oxidation pathway through a combination of bioprocessing and genetic strategies, specifically including antioxidant supplementation and strain engineering for ROS-scavenging enzyme overexpression.

We began by investigating the effects of antioxidant supplementation using ascorbic acid, which is a natural, water-soluble, and cost-effective compound known for its ROS-scavenging properties [302]. A range of concentrations (0–2 g/L) of ascorbic acid was supplemented upon cultivation to evaluate its influence on UP biosynthesis in *E. coli*. Our results revealed that supplementing ascorbic acid up to 1 g/L significantly enhanced the UP-III/UP-I ratio while maintaining approximately the same level of total UP biosynthesis, indicating a successful redirection of the dissimilated carbon flux toward the HemD-enzymatic pathway. However, increasing the concentration to 2 g/L did not yield further improvement in flux redirection, but, in fact, led to a reduction in total UP biosynthesis and a noticeable retardation of cell growth. This inhibitory effect at higher ascorbic acid concentrations aligns with previous studies, which reported that ascorbic acid could exhibit antibacterial properties in the range of 2–20 g/L [303, 304]. Ascorbic acid facilitates intracellular redox homeostasis by directly scavenging ROS, thereby reducing oxidative stress and protecting key biosynthetic enzymes from damage [305]. By lowering ROS levels, ascorbic acid minimizes the diversion of intermediates toward the non-enzymatic auto-oxidation pathway, allowing metabolic flux to be redirected toward the HemD-enzymatic pathway. This redirection enhances type-III porphyrin production while reducing the accumulation of type-I porphyrins. These observations highlight the importance of adjusting antioxidant levels to achieve a physiological balance between ROS mitigation and maintaining cellular viability and metabolic activity during cell cultivation.

Next, we explored the use of enzymatic antioxidants by overexpressing various *sod* genes in UP-producing *E. coli* strains. Among these *sod* genes, *sodA* appeared to be the most effective one in redirecting the dissimilated carbon flux toward the HemD-enzymatic pathway. Notably, the UP-producing strain with *sodA* overexpression not only had a high UP-III/UP-I ratio of 1.9 but also enhanced total UP production to 1.56 g/L, which was

72.9% more than the control cultivation. The observed improvement in UP biosynthesis could be attributed to the role of SodA in mitigating oxidative stress, which protects key enzymes in the biosynthetic pathway from damage caused by ROS. Overexpression of *sodA* has been shown in multiple studies to enhance product titers such as 5-ALA [306, 81]. Notably, one study reported that the improved redox balance achieved by *sod* gene overexpression enhanced the activity of key enzymes directly involved in product formation, a mechanism that may also explain the improved yields observed in our study [225]. In addition, the ability of SodA to balance intracellular redox conditions may indirectly shift more dissimilated carbon flux away from auto-oxidation for UP-III biosynthesis. The function of SodA in protecting enzymatic activities under oxidative stress has been reported, as its inactivation led to elevated oxidation of specific enzymes and reduced ATP/NADH levels [307].

Additionally, while overexpression of *sodB* also resulted in a relatively high UP-III/UP-I ratio, the total UP biosynthesis was significantly reduced compared to overexpression of *sodA*. This may be due to differences in the catalytic efficiency and metal-ion-dependency of Sod enzymes, with Mn-dependent SodA likely providing more effective oxidative stress mitigation under the physiological conditions for UP biosynthesis. Conversely, *sodC* overexpression somehow disrupted the cell's metabolic activity for porphyrin production, and the cultivation had no pigmentation. This effect may be associated with the periplasmic localization of SodC, limiting its activity to the periplasm as superoxide cannot cross cellular membranes [308]. The confined activity of SodC could result in an imbalance in redox homeostasis between the periplasm and cytoplasm, potentially interfering with select intracellular processes critical for porphyrin biosynthesis. Additionally, the activity of SodC may contribute to hydrogen peroxide accumulation in the periplasm, a phenomenon reported in studies associating *sodC* overexpression with altered redox dynamics and dye degradation [309].

Combining the two strategies of ascorbic acid supplementation and *sodA* overexpression did not show clear evidence of synergistic effects. Instead, the results suggest that relying on either enzymatic (i.e., SodA) or non-enzymatic (i.e., ascorbic acid) antioxidants alone may be more beneficial for redirecting dissimilated carbon flux upon UP biosynthesis. This could be attributed to potential redundancy or competition in their antioxidative mechanisms, indicating the need for a balanced approach tailored to specific oxidative conditions. However, further studies are needed to fully understand the relationship between their antioxidative mechanisms.

Overexpression of SOD genes can effectively scavenge superoxide radicals and convert them into hydrogen peroxide, which plays a dual role in cellular metabolism. While low-level hydrogen peroxide acts as a signaling molecule essential for various cellular processes,

including heme biosynthesis [227, 310] and catalase gene (*katG*) induction [311, 295], excessive concentration of hydrogen peroxide might be toxic to the cells by exhibiting adverse effects due to DNA damage [312, 313]. Therefore, regulating the hydrogen peroxide level is critical for maintaining cellular homeostasis for optimal physiological and metabolic activities. To achieve this balance in the intracellular level of hydrogen peroxide, we explored co-overexpression of the *sod* genes with various catalase genes by deriving strains KAT1-KAT5. KAT1, expressing both *sodABC* and *katEG* along with *hemAB*, appears to disrupt porphyrin biosynthesis by exhibiting minimal pigmentation, prompting us for a closer examination of individual *kat* genes. Strains KAT2 and KAT3, overexpressing *katE* and *katG*, respectively, along with *hemAB*, also exhibited negligible porphyrin production. This could be attributed to excessive degradation of the minimal-level hydrogen peroxide required for triggering porphyrin biosynthesis and other cellular processes. As catalase overexpression alone was proved ineffective, individual *katE* and *katG* were subsequently co-expressed with *sodA*, generating the strains KAT4 and KAT5. Among them, KAT4 had superior culture performance, producing 1.58 g/L UP with a UP-III/UP-I ratio of 1.3. These results suggest that the combination of *sodA* and *katE* could not only effectively eliminate superoxides but also balance ROS detoxification, successfully directing more dissimilated carbon flux toward the HemD-enzymatic pathway for enhanced UP-III biosynthesis. The results highlight the importance of the intricate balance required to manage ROS and minimize cellular stress, both contributing toward UP-III biosynthesis. Hence, identification of alternative antioxidant systems and their interactions for genetic modification may provide new avenues to enhance porphyrin biosynthesis.

Recent advancements in porphyrin synthesis have demonstrated the applicability of metabolic engineering strategies and pathway regulation in enhancing biosynthetic efficiency [13, 84], whereas integrated optimization approaches further highlight the potential for improving biosynthesis [314]. Aligned with these advancements, our study demonstrates that cost-effective antioxidant supplementation, such as ascorbic acid, combined with the overexpression of ROS-scavenging enzymes (e.g., SodA and KatE), significantly enhances porphyrin biosynthesis and highlights its potential for industrial-scale applications.

Chapter 8

Original Contributions, Recommendations, and Future Prospects

8.1 Original Contributions

This thesis has made significant advancements in metabolic engineering and synthetic biology, particularly in the development of genome engineering tools and innovative strategies for porphyrin biosynthesis in *E. coli*. These contributions are detailed below:

8.1.1 Establishing a versatile genome engineering toolkit for manipulating *E. coli*

In Chapter 3 of this thesis, a highly efficient genome engineering toolkit was developed by integrating CRISPR-Cas9 and transposon-based technologies. This toolkit combines the CRISPR-Cas9 system with λ Red recombineering and the Tn6677 transposon-associated CRISPR-Cas system to enable precise, efficient, and iterative genomic modifications tailored for metabolic engineering applications. The genomic incorporation of the *cas9* gene and λ Red recombineering machinery enhances the efficiency of targeted gene inactivation, while the Tn6677 transposon system allows for insertion of large DNA fragments, up to 5.5 kb, at predefined genomic loci. This approach eliminates the reliance on extrachromosomal plasmids, significantly reducing the physiological burden on the host strain and mitigating adverse effects associated with plasmid maintenance and antibiotic selection pressures. By

integrating these systems, the toolkit streamlines both gene knockouts and the insertion of complex pathways, providing the ability to perform successive genetic modifications with high precision and reproducibility. The development of this toolkit provides a versatile platform that enables the engineering of strains for enhanced production of target metabolites, as demonstrated through the research objectives and outcomes presented in this chapter.

8.1.2 Developing efficient production of uroporphyrin in *E. coli*

In Chapter 4 of this thesis, a novel bio-based production platform for uroporphyrin (UP) was developed by engineering *E. coli* strains. UP, a porphyrin compound with significant medical applications and a key precursor in heme biosynthesis, has lacked an efficient biosynthetic production strategy. To address this challenge, the Shemin/C4 pathway was heterologously implemented in *E. coli*, and metabolic engineering strategies were previously employed to enhance the intracellular pool of succinyl-CoA, a critical precursor for 5-aminolevulinic acid (5-ALA) synthesis. A synthetic operon regulated by the trc promoter was constructed to express key pathway genes, including *hemA*, *hemB*, *hemC*, and *hemD*. The effects of these genes on UP biosynthesis were systematically characterized. Cultivation of the engineered *E. coli* strains in a batch bioreactor with glycerol as the carbon source under aerobic conditions resulted in the production of UP with titers exceeding 900 mg/L. Notably, the majority of the synthesized UP was secreted extracellularly with a high purity, facilitating downstream purification processes.

8.1.3 Bioprocess engineering toward high-level coproporphyrin production in *E. coli*

In Chapter 5, building upon the knowledge and strategies established in the previous chapter for uroporphyrin production, we expanded the biosynthetic pathway in *E. coli* to enable the efficient production of coproporphyrin (CP). CP is a porphyrin pigment with various applications, including clinical diagnostics and as a precursor for heme biosynthesis. Similarly, the Shemin/C4 pathway was heterologously expressed in *E. coli* strains engineered to enhance the intracellular availability of succinyl-CoA. To regulate the expression of the key pathway genes (*hemA*, *hemB*, *hemD*, *hemE*, and *hemY*), we employed a plasmid system comprising two operons controlled by strong trc and grac max promoters, respectively. This differential expression strategy allowed for fine-tuning of the associated metabolic flux in the CP biosynthetic pathway, optimizing substrate channeling and minimizing the accumulation of intermediates or byproducts. Using the engineered *E. coli* strains for bioreactor

cultivation under aerobic conditions with glycerol as the carbon source, we achieved CP production titers of up to 353 mg/L with minimal byproduct formation. Notably, the overproduced CP was secreted extracellularly, posing minimal physiological toxicity and impact on the producing cells. This secretion simplifies downstream purification processes and enhances the overall feasibility of large-scale CP production.

8.1.4 Developing a two-step strategy for coproheme and heme production via heterologous coproporphyrin-dependent (CPD) pathway

In Chapter 6, a novel two-step approach was developed to enable efficient bio-based production of coproheme and heme using the heterologous coproporphyrin-dependent (CPD) pathway in *E. coli*. This strategy addresses the inherent challenges associated with intracellular heme accumulation and its associated toxicity by decoupling the production process into distinct in-vivo and in-vitro phases. In the in-vivo phase, *E. coli* strains were engineered to express the CPD pathway, producing up to 85 mg/L of coproheme and 18 mg/L of heme under batch cultivations. A plasmid-based system was constructed to co-express critical pathway genes (*hemA*, *hemB*, *hemD*, *hemE*, *hemH*, and *hemQ*), with each gene under the control of strong trc and grac max promoters. This design facilitated efficient metabolic flux toward coproheme biosynthesis while minimizing the accumulation of toxic intermediates. Among the strains tested, the best-performing strain demonstrated high extracellular secretion of coproheme, reducing intracellular toxicity and streamlining downstream processing. In the in-vitro phase, purified HemH and HemQ enzymes were applied to convert coproporphyrin III (CP-III) to coproheme and subsequently to heme. This cell-free enzymatic system enabled precise control over reaction conditions, achieving a conversion efficiency of 77.2% for CP-III to coproheme and 45.8% for coproheme to heme. The modularity of this approach allows for refined scalability and overcomes the bottlenecks of intracellular biosynthesis.

8.1.5 Mitigating oxidative stress to enhance porphyrin biosynthetic efficiency

In Chapter 7, strategies to mitigate oxidative stress were implemented to enhance the efficiency of porphyrin biosynthesis in *E. coli*. Oxidative stress, caused by the accumulation of reactive oxygen species (ROS), disrupts cellular homeostasis, adversely impacts metabolic pathways, and reduces the efficiency of biosynthetic processes, including those

for porphyrin production. One approach involved supplementing the culture medium with ascorbic acid (vitamin C) at concentrations up to 1 g/L. This supplementation significantly shifted the uroporphyrin III (UP-III) to uroporphyrin I (UP-I) ratio from 0.62 to 2.57, favoring the production of the desired type-III isomer. In addition, genetic engineering strategies were employed to strengthen the endogenous ROS-scavenging mechanisms. Overexpression of the *sodA* gene, encoding superoxide dismutase, led to a 72.9% increase in total porphyrin production, reaching 1.56 g/L. The UP-III/UP-I ratio improved to 1.94, indicating a significant enhancement in pathway efficiency. Superoxide dismutase converts harmful superoxide radicals into hydrogen peroxide, which is further decomposed into water and oxygen by catalases, reducing oxidative damage and improving cell viability. These combined strategies demonstrate the critical role of managing oxidative stress in microbial systems to optimize biosynthesis.

8.2 Recommendations and Future Prospects

The findings of this thesis provide a promising foundation for future research and industrial applications. Future work could focus on advancing the strategies and methodologies established in this thesis to enhance the biosynthesis and industrial applicability of porphyrins. Scaling up porphyrin production to industrial bioreactors presents significant opportunities. Optimization of medium compositions and supplementation strategies, such as controlled iron delivery systems, could enhance metabolic yields while minimizing oxidative stress from free iron ions. Techniques like controlled-release formulations might balance iron availability and mitigate associated toxicity. Addressing challenges related to aeration, agitation, and shear stress in large-scale bioreactors could also improve cell viability and production efficiency under intensified conditions. Expanding the substrate base by incorporating renewable feedstocks, such as lignocellulosic hydrolysates or agricultural residues, could be explored to reduce production costs and align the process with sustainability goals. Recycling byproducts into the growth medium might further enhance efficiency and minimize waste, contributing to a circular economy.

Enzyme immobilization and stabilization could be investigated as a means to improve cell-free biosynthetic processes. Stabilizing enzymes like HemH and HemQ on solid supports could allow for their reuse and enhance their catalytic efficiency. Immobilized enzymes might also enable more precise control of reaction conditions and simplify downstream processing, paving the way for scalable heme production. Protein engineering and directed evolution could be employed to enhance the stability and activity of key enzymes in the porphyrin pathway. Improved enzyme performance under diverse conditions could further increase yields and enable industrial-scale production.

The effects of antioxidants on other porphyrins could be further examined to extend their role in mitigating oxidative stress during porphyrin synthesis. Our ongoing investigations aim to evaluate how antioxidants influence downstream products such as coproporphyrin, coproheme, and heme. This knowledge could refine strategies to optimize pathway efficiency and reduce byproduct formation. Exploring other oxidative stress regulation mechanisms could also be beneficial. This includes engineering the redox balance of the cell through overexpression of NADH-dependent oxidoreductases or enhancing the production of glutathione, a key intracellular antioxidant. Additionally, tuning cellular respiration rates or introducing alternative electron acceptors could be investigated to minimize the generation of ROS under aerobic conditions. These mechanisms could complement antioxidant supplementation and endogenous ROS-scavenging systems to further improve oxidative stress management.

Another promising direction is the production of linear tetrapyrroles, including biliverdin, phycocyanobilin (PCB), and phycoerythrobilin (PEB), which have diverse applications in medicine, diagnostics, and food industries. These extensions of the porphyrin biosynthesis platform could be explored by introducing alternative enzymatic pathways or modifying downstream reactions. Current work in our lab is investigating biliverdin synthesis to assess the feasibility of such adaptations.

References

[1] Hal Alper and Gregory Stephanopoulos. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology*, 7(10):715–723, 2009.

[2] Martin Gustavsson and Sang Yup Lee. Prospects of microbial cell factories developed through systems metabolic engineering. *Microbial Biotechnology*, 9(5):610–617, 2016.

[3] Cong T Trinh and Brian Mendoza. Modular cell design for rapid, efficient strain engineering toward industrialization of biology. *Current Opinion in Chemical Engineering*, 14:18–25, 2016.

[4] Stanley N Cohen, Annie CY Chang, Herbert W Boyer, and Robert B Helling. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences*, 70(11):3240–3244, 1973.

[5] Roy Curtiss and Frederick Carl Neidhardt. *Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press, 1996.

[6] Parayil Kumaran Ajikumar, Keith Tyo, Simon Carlsen, Oliver Mucha, Too Heng Phon, and Gregory Stephanopoulos. Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Molecular pharmaceutics*, 5(2):167–190, 2008.

[7] Leanne B Josefsen and Ross W Boyle. Unique diagnostic and therapeutic roles of porphyrins and phthalocyanines in photodynamic therapy, imaging and theranostics. *Theranostics*, 2(9):916, 2012.

[8] Yiming Zhou, Xiaolong Liang, and Zhifei Dai. Porphyrin-loaded nanoparticles for cancer theranostics. *Nanoscale*, 8(25):12394–12405, 2016.

[9] Ncediwe Tsolekile, Simphiwe Nelana, and Oluwatobi Samuel Oluwafemi. Porphyrin as diagnostic and therapeutic agent. *Molecules*, 24(14):2669, 2019.

[10] Hadia Aftab and Rebecca K Donegan. Regulation of heme biosynthesis via the coproporphyrin dependent pathway in bacteria. *Frontiers in Microbiology*, 15:1345389, 2024.

[11] Seok Joon Kwon, Arjo L De Boer, Ralf Petri, and Claudia Schmidt-Dannert. High-level production of porphyrins in metabolically engineered escherichia coli: systematic extension of a pathway assembled from overexpressed genes involved in heme biosynthesis. *Applied and environmental microbiology*, 69(8):4875–4883, 2003.

[12] Yoshiko Kubota, Kazumi Nomura, Yasutake Katoh, Rina Yamashita, Kiriko Kaneko, and Kazumichi Furuyama. Novel mechanisms for heme-dependent degradation of alas1 protein as a component of negative feedback regulation of heme biosynthesis. *Journal of Biological Chemistry*, 291(39):20516–20529, 2016.

[13] Ying-Chen Yi, I-Tai Shih, Tzu-Hsuan Yu, Yen-Ju Lee, and I-Son Ng. Challenges and opportunities of bioprocessing 5-aminolevulinic acid using genetic and metabolic engineering: a critical review. *Bioresources and Bioprocessing*, 8:1–18, 2021.

[14] Zhen Kang, Wenwen Ding, Xu Gong, Qingtao Liu, Guocheng Du, and Jian Chen. Recent advances in production of 5-aminolevulinic acid using biological strategies. *World Journal of Microbiology and Biotechnology*, 33:1–7, 2017.

[15] Sandra M Lelli, Leonor C San Martín de Viale, and Marta B Mazzetti. Response of glucose metabolism enzymes in an acute porphyria model: Role of reactive oxygen species. *Toxicology*, 216(1):49–58, 2005.

[16] Karthik P Jayapal, Katie F Wlaschin, Wei Hu, and Miranda GS Yap. Recombinant protein therapeutics from cho cells-20 years and counting. *Chemical engineering progress*, 103(10):40, 2007.

[17] Gary Walsh. Biopharmaceutical benchmarks 2014. *Nature biotechnology*, 32(10):992–1000, 2014.

[18] Jens Nielsen and Jay D Keasling. Engineering cellular metabolism. *Cell*, 164(6):1185–1197, 2016.

[19] Victor Chubukov, Aindrila Mukhopadhyay, Christopher J Petzold, Jay D Keasling, and Héctor García Martín. Synthetic and systems biology for microbial production of commodity chemicals. *NPJ systems biology and applications*, 2(1):1–11, 2016.

[20] Arnold L Demain and Preeti Vaishnav. Production of recombinant proteins by microbes and higher organisms. *Biotechnology advances*, 27(3):297–306, 2009.

[21] Guo-Qiang Chen and Xiao-Ran Jiang. Engineering microorganisms for improving polyhydroxyalkanoate biosynthesis. *Current opinion in biotechnology*, 53:20–25, 2018.

[22] Sang Yup Lee, Hyun Uk Kim, Tong Un Chae, Jae Sung Cho, Je Woong Kim, Jae Ho Shin, Dong In Kim, Yoo-Sung Ko, Woo Dae Jang, and Yu-Sin Jang. A comprehensive metabolic map for production of bio-based chemicals. *Nature catalysis*, 2(1):18–33, 2019.

[23] Jay D Keasling. Manufacturing molecules through metabolic engineering. *Science*, 330(6009):1355–1358, 2010.

[24] Judith Becker and Christoph Wittmann. Advanced biotechnology: Metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angewandte Chemie International Edition*, 54(11):3328–3350, 2015.

[25] Laura R Jarboe, Xueli Zhang, Xuan Wang, Jonathan C Moore, KT Shanmugam, and Lonnie O Ingram. Metabolic engineering for production of biorenewable fuels and chemicals: contributions of synthetic biology. *BioMed Research International*, 2010(1):761042, 2010.

[26] Cheng Li, Keaton L Lesnik, and Hong Liu. Microbial conversion of waste glycerol from biodiesel production into value-added products. *Energies*, 6(9):4739–4768, 2013.

[27] Furkan H Isikgor and C Remzi Becer. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer chemistry*, 6(25):4497–4559, 2015.

[28] Jonas Kühlborn, Jonathan Groß, and Till Opatz. Making natural products from renewable feedstocks: back to the roots? *Natural Product Reports*, 37(3):380–424, 2020.

[29] John M Woodley. Biocatalysis for future sustainable manufacturing. *The Biochemist*, 44(3):6–8, 2022.

[30] James M Clomburg, Anna M Crumbley, and Ramon Gonzalez. Industrial biomanufacturing: the future of chemical production. *Science*, 355(6320):aag0804, 2017.

[31] Tom A Ewing, Niels Nouse, Matthijs van Lint, Jacco van Haveren, Jeroen Hugenholtz, and Daan S van Es. Fermentation for the production of biobased chemicals in a circular economy: a perspective for the period 2022–2050. *Green Chemistry*, 24(17):6373–6405, 2022.

[32] Devendra S Wagare, Sagar E Shirsath, Mujahed Shaikh, and Prashant Netankar. Sustainable solvents in chemical synthesis: a review. *Environmental Chemistry Letters*, 19(4):3263–3282, 2021.

[33] Gregory Stephanopoulos. Synthetic biology and metabolic engineering. *ACS synthetic biology*, 1(11):514–525, 2012.

[34] Yi-Heng Percival Zhang, Jibin Sun, and Yanhe Ma. Biomanufacturing: history and perspective. *Journal of Industrial Microbiology and Biotechnology*, 44(4-5):773–784, 2017.

[35] Roger A Sheldon. Metrics of green chemistry and sustainability: past, present, and future. *ACS Sustainable Chemistry & Engineering*, 6(1):32–48, 2018.

[36] James E Bailey. Toward a science of metabolic engineering. *Science*, 252(5013):1668–1675, 1991.

[37] Lauren B Pickens, Yi Tang, and Yit-Heng Chooi. Metabolic engineering for the production of natural products. *Annual review of chemical and biomolecular engineering*, 2(1):211–236, 2011.

[38] Laura L Anzaldi and Eric P Skaar. Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. *Infection and immunity*, 78(12):4977–4989, 2010.

[39] Shilpa Bali, Andrew D Lawrence, Susana A Lobo, Lígia M Saraiva, Bernard T Golding, David J Palmer, Mark J Howard, Stuart J Ferguson, and Martin J Warren. Molecular hijacking of siroheme for the synthesis of heme and d 1 heme. *Proceedings of the National Academy of Sciences*, 108(45):18260–18265, 2011.

[40] Kyeong Rok Choi, Hye Eun Yu, Hoseong Lee, and Sang Yup Lee. Improved production of heme using metabolically engineered escherichia coli. *Biotechnology and Bioengineering*, 119(11):3178–3193, 2022.

[41] D Ewen Cameron, Caleb J Bashor, and James J Collins. A brief history of synthetic biology. *Nature Reviews Microbiology*, 12(5):381–390, 2014.

[42] Weiwen Zhang, Feng Li, and Lei Nie. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. *Microbiology*, 156(2):287–301, 2010.

[43] Zachary D Blount. The unexhausted potential of e. coli. *elife*, 4:e05826, 2015.

[44] David V Goeddel, Dennis G Kleid, Francisco Bolivar, Herbert L Heyneker, Daniel G Yansura, Roberto Crea, Tadaaki Hirose, Adam Kraszewski, Keiichi Itakura, and Arthur D Riggs. Expression in escherichia coli of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences*, 76(1):106–110, 1979.

[45] Sang Yup Lee and Hyun Uk Kim. Systems strategies for developing industrial microbial strains. *Nature biotechnology*, 33(10):1061–1072, 2015.

[46] Jessica C Zweers, Imrich Barák, Dörte Becher, Arnold JM Driessen, Michael Hecker, Vesa P Kontinen, Manfred J Saller, L'udmila Vavrová, and Jan Maarten van Dijl. Towards the development of bacillus subtilis as a cell factory for membrane proteins and protein complexes. *Microbial cell factories*, 7:1–20, 2008.

[47] Elise K Phillips, Jordan A Cannon, Yue Zhou, Kyle S Bonifer, and Todd B Reynolds. Conjugation-mediated plasmid transfer enables genetic modification of diverse bacillus species. *Microbiology Spectrum*, 11(2):e03700–22, 2023.

[48] Meirong Zhao, Jianfan Ma, Lei Zhang, and Haishan Qi. Engineering strategies for enhanced heterologous protein production by saccharomyces cerevisiae. *Microbial cell factories*, 23(1):32, 2024.

[49] Yangdanyu Long, Xiao Han, Xuanlin Meng, Ping Xu, and Fei Tao. A robust yeast chassis: comprehensive characterization of a fast-growing saccharomyces cerevisiae. *Mbio*, 15(2):e03196–23, 2024.

[50] Jee Yon Kim, Yeon-Gu Kim, and Gyun Min Lee. Cho cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied microbiology and biotechnology*, 93:917–930, 2012.

[51] Klaudia Meskova, Katarina Martonova, Patricia Hrasnova, Kristina Sinska, Michaela Skrabanova, Lubica Fialova, Stefana Njemoga, Ondrej Cehlar, Olga Parmar, and Petr Kolenko. Cost-effective protein production in cho cells following polyethylenimine-mediated gene delivery showcased by the production and crystallization of antibody fabs. *Antibodies*, 12(3):51, 2023.

[52] Sarfaraz K Niazi and Matthias Magoola. Advances in escherichia coli-based therapeutic protein expression: mammalian conversion, continuous manufacturing, and cell-free production. *Biologics*, 3(4):380–401, 2023.

[53] Kenan C Murphy. Use of bacteriophage λ recombination functions to promote gene replacement in escherichia coli. *Journal of bacteriology*, 180(8):2063–2071, 1998.

[54] Shyam K Sharan, Lynn C Thomason, Sergey G Kuznetsov, and Donald L Court. Recombineering: a homologous recombination-based method of genetic engineering. *Nature protocols*, 4(2):206–223, 2009.

[55] Jeffrey L Bose. Chemical and uv mutagenesis. *The Genetic Manipulation of Staphylococci: Methods and Protocols*, pages 111–115, 2016.

[56] Thomas Gaj, Charles A Gersbach, and Carlos F Barbas. Zfn, talen, and crispr/cas-based methods for genome engineering. *Trends in biotechnology*, 31(7):397–405, 2013.

[57] Peter J Chen and David R Liu. Prime editing for precise and highly versatile genome manipulation. *Nature Reviews Genetics*, 24(3):161–177, 2023.

[58] Holly A Rees and David R Liu. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature reviews genetics*, 19(12):770–788, 2018.

[59] Harris H Wang, Farren J Isaacs, Peter A Carr, Zachary Z Sun, George Xu, Craig R Forest, and George M Church. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, 460(7257):894–898, 2009.

[60] Christine L Schneider. Bacteriophage-mediated horizontal gene transfer: transduction. *Bacteriophages: biology, technology, therapy*, pages 151–192, 2021.

[61] Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. A programmable dual-rna–guided dna endonuclease in adaptive bacterial immunity. *science*, 337(6096):816–821, 2012.

[62] Kyoung-Hee Choi and Kang-Ju Kim. Applications of transposon-based gene delivery system in bacteria. *Journal of microbiology and biotechnology*, 19(3):217–228, 2009.

[63] Le Cong, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, Xuebing Wu, Wenyan Jiang, and Luciano A Marraffini. Multiplex genome engineering using crispr/cas systems. *Science*, 339(6121):819–823, 2013.

[64] Patrick D Hsu, Eric S Lander, and Feng Zhang. Development and applications of crispr-cas9 for genome engineering. *Cell*, 157(6):1262–1278, 2014.

[65] Giedrius Gasiunas, Rodolphe Barrangou, Philippe Horvath, and Virginijus Siksnys. Cas9–crrna ribonucleoprotein complex mediates specific dna cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, 109(39):E2579–E2586, 2012.

[66] FAFA Ran, Patrick D Hsu, Jason Wright, Vineeta Agarwala, David A Scott, and Feng Zhang. Genome engineering using the crispr-cas9 system. *Nature protocols*, 8(11):2281–2308, 2013.

[67] Fuguo Jiang and Jennifer A Doudna. Crispr–cas9 structures and mechanisms. *Annual review of biophysics*, 46(1):505–529, 2017.

[68] Addison V Wright, James K Nuñez, and Jennifer A Doudna. Biology and applications of crispr systems: harnessing nature's toolbox for genome engineering. *Cell*, 164(1):29–44, 2016.

[69] Mazhar Adli. The crispr tool kit for genome editing and beyond. *Nature communications*, 9(1):1911, 2018.

[70] Ian M Slaymaker, Linyi Gao, Bernd Zetsche, David A Scott, Winston X Yan, and Feng Zhang. Rationally engineered cas9 nucleases with improved specificity. *Science*, 351(6268):84–88, 2016.

[71] Bernd Zetsche, Jonathan S Gootenberg, Omar O Abudayyeh, Ian M Slaymaker, Kira S Makarova, Patrick Essletzbichler, Sara E Volz, Julia Joung, John Van Der Oost, and Aviv Regev. Cpf1 is a single rna-guided endonuclease of a class 2 crispr-cas system. *cell*, 163(3):759–771, 2015.

[72] Seung Woo Cho, Sojung Kim, Jong Min Kim, and Jin-Soo Kim. Targeted genome engineering in human cells with the cas9 rna-guided endonuclease. *Nature biotechnology*, 31(3):230–232, 2013.

[73] Alison B Hickman and Fred Dyda. Mechanisms of dna transposition. *Mobile DNA III*, pages 529–553, 2015.

[74] Zoltán Ivics and Zsuzsanna Izsvák. The expanding universe of transposon technologies for gene and cell engineering. *Mobile DNA*, 1:1–15, 2010.

[75] Shengrui Wu, Pingfang Tian, and Tianwei Tan. Genomic landscapes of bacterial transposons and their applications in strain improvement. *Applied Microbiology and Biotechnology*, 106(19):6383–6396, 2022.

[76] Hai-yue Bao, Hui-jie Li, Yong-yong Zhang, Andreas Bechthold, Xiao-ping Yu, and Zheng Ma. Transposon-based identification of genes involved in the rimocidin biosynthesis in streptomyces rimosus m527. *World Journal of Microbiology and Biotechnology*, 39(12):359, 2023.

[77] Solenne Bire, Déborah Ley, Sophie Casteret, Nicolas Mermod, Yves Bigot, and Florence Rouleux-Bonnin. Optimization of the piggybac transposon using mrna and insulators: toward a more reliable gene delivery system. *PLoS One*, 8(12):e82559, 2013.

[78] Sway P Chen and Harris H Wang. An engineered cas-transposon system for programmable and site-directed dna transpositions. *The CRISPR journal*, 2(6):376–394, 2019.

[79] Anirban Sen Gupta. Bio-inspired nanomedicine strategies for artificial blood components. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 9(6):e1464, 2017.

[80] Lu-Lin Li and Eric Wei-Guang Diau. Porphyrin-sensitized solar cells. *Chemical society reviews*, 42(1):291–304, 2013.

[81] Shaomei Yang, Zihao Guo, Jiuyu Sun, Jingxuan Wei, Qinyuan Ma, and Xiuzhen Gao. Recent advances in microbial synthesis of free heme. *Applied Microbiology and Biotechnology*, 108(1):68, 2024.

[82] Haihong Chen, Yaohong Wang, Weishan Wang, Ting Cao, Lu Zhang, Zhengduo Wang, Xuran Chi, Tong Shi, Huangwei Wang, and Xinwei He. High-yield porphyrin production through metabolic engineering and biocatalysis. *Nature Biotechnology*, pages 1–11, 2024.

[83] Alan R Battersby. Tetrapyrroles: the pigments of life. *Natural product reports*, 17(6):507–526, 2000.

[84] Gunhild Layer, Joachim Reichelt, Dieter Jahn, and Dirk W Heinz. Structure and function of enzymes in heme biosynthesis. *Protein Science*, 19(6):1137–1161, 2010.

[85] Joseph Shiloach and Rephael Fass. Growing e. coli to high cell density—a historical perspective on method development. *Biotechnology advances*, 23(5):345–357, 2005.

[86] Sang Yup Lee. High cell-density culture of escherichia coli. *Trends in biotechnology*, 14(3):98–105, 1996.

[87] Martin J Warren, Alison G Smith, Dieter Jahn, and Dirk W Heinz. Biosynthesis of 5-aminolevulinic acid. *Tetrapyrroles: birth, life and death*, pages 29–42, 2009.

[88] Harry A Dailey and Amy E Medlock. A primer on heme biosynthesis. *Biological Chemistry*, 403(11-12):985–1003, 2022.

[89] Katrin Müller, Toni Mingers, V Haskamp, Dieter Jahn, and Martina Jahn. Biosynthesis and insertion of heme. *Aerobic utilization of hydrocarbons, oils, and lipids*, pages 1–28, 2019.

[90] Richard S Ajioka, John D Phillips, and James P Kushner. Biosynthesis of heme in mammals. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1763(7):723–736, 2006.

[91] Harry A Dailey, Tamara A Dailey, Svetlana Gerdes, Dieter Jahn, Martina Jahn, Mark R O'Brian, and Martin J Warren. Prokaryotic heme biosynthesis: multiple pathways to a common essential product. *Microbiology and Molecular Biology Reviews*, 81(1):10.1128/mmbr. 00048–16, 2017.

[92] Peter M Jordan. *Biosynthesis of tetrapyrroles*. Elsevier, 1991.

[93] Harry A Dailey, Svetlana Gerdes, Tamara A Dailey, Joseph S Burch, and John D Phillips. Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. *Proceedings of the National Academy of Sciences*, 112(7):2210–2215, 2015.

[94] Charlie Hobbs, Harry A Dailey, and Mark Shepherd. The hemq coprohaem decarboxylase generates reactive oxygen species: implications for the evolution of classical haem biosynthesis. *Biochemical Journal*, 473(21):3997–4009, 2016.

[95] Lorenz Heidinger, Isabelle Fix, Thorsten Friedrich, and Gunhild Layer. Trapping the substrate radical of heme synthase ahbd. *Frontiers in Chemistry*, 12:1430796, 2024.

[96] Melanie Kühner, Kristin Haufschildt, Alexander Neumann, Sonja Storbeck, Judith Streif, and Gunhild Layer. The alternative route to heme in the methanogenic archaeon methanosarcina barkeri. *Archaea*, 2014(1):327637, 2014.

[97] M Elizabeth Stroupe, Helen K Leech, Douglas S Daniels, Martin J Warren, and Elizabeth D Getzoff. Cysg structure reveals tetrapyrrole-binding features and novel regulation of siroheme biosynthesis. *Nature Structural & Molecular Biology*, 10(12):1064–1073, 2003.

[98] Chibuike David Obi, Tawhid Bhuiyan, Harry A Dailey, and Amy E Medlock. Ferrochelatase: mapping the intersection of iron and porphyrin metabolism in the mitochondria. *Frontiers in Cell and Developmental Biology*, 10:894591, 2022.

[99] Prem Ponka. Tissue-specific regulation of iron metabolism and heme synthesis: distinct control mechanisms in erythroid cells. *Blood, The Journal of the American Society of Hematology*, 89(1):1–25, 1997.

[100] Fernando G Princ, Adela Ana Juknat, Adrian A Amitrano, and Alcira Batlle. Effect of reactive oxygen species promoted by δ-aminolevulinic acid on porphyrin biosynthesis and glucose uptake in rat cerebellum. *General Pharmacology: The Vascular System*, 31(1):143–148, 1998.

[101] Nobuyuki Tanimura, Eli Miller, Kazuhiko Igarashi, David Yang, Judith N Burstyn, Colin N Dewey, and Emery H Bresnick. Mechanism governing heme synthesis reveals a gata factor/heme circuit that controls differentiation. *EMBO reports*, 17(2):249–265, 2016.

[102] Ana Paula Oliveira and Uwe Sauer. The importance of post-translational modifications in regulating saccharomyces cerevisiae metabolism. *FEMS yeast research*, 12(2):104–117, 2012.

[103] Jawahar Sudhamsu, Mariam Kabir, Michael V Airola, Bhumit A Patel, Syun-Ru Yeh, Denis L Rousseau, and Brian R Crane. Co-expression of ferrochelatase allows for complete heme incorporation into recombinant proteins produced in e. coli. *Protein expression and purification*, 73(1):78–82, 2010.

[104] Katerina Krumova and Gonzalo Cosa. Overview of reactive oxygen species. 2016.

[105] Yuhang Hong, Alessandra Boiti, Daniela Vallone, and Nicholas S Foulkes. Reactive oxygen species signaling and oxidative stress: transcriptional regulation and evolution. *Antioxidants*, 13(3):312, 2024.

[106] Aleksandra Rapacka-Zdończyk, Agata Woźniak, Klaudia Michalska, Michał Pierański, Patrycja Ogonowska, Mariusz Grinholc, and Joanna Nakonieczna. Factors determining the susceptibility of bacteria to antibacterial photodynamic inactivation. *Frontiers in Medicine*, 8:642609, 2021.

[107] Celia Andrés Juan, José Manuel Pérez de la Lastra, Francisco J Plou, and Eduardo Pérez-Lebeña. The chemistry of reactive oxygen species (ros) revisited: outlining their role in biological macromolecules (dna, lipids and proteins) and induced pathologies. *International journal of molecular sciences*, 22(9):4642, 2021.

[108] Wafa Jallouli, Nabil Zouari, and Samir Jaoua. Involvement of oxidative stress and growth at high cell density in the viable but nonculturable state of photorhabdus temperata ssp. temperata strain k122. *Process Biochemistry*, 45(5):706–713, 2010.

[109] Antonino Baez and Joseph Shiloach. Escherichia coli avoids high dissolved oxygen stress by activation of soxrs and manganese-superoxide dismutase. *Microbial cell factories*, 12:1–9, 2013.

[110] Jose M Bruno-Bárcena, M Andrea Azcárate-Peril, and Hosni M Hassan. Role of antioxidant enzymes in bacterial resistance to organic acids. *Applied and environmental microbiology*, 76(9):2747–2753, 2010.

[111] Weijin Wu, Xiyan Li, Xu Zhang, Tingyue Gu, Yongqiu Qiu, Minglong Zhu, and Wensong Tan. Characteristics of oxidative stress and antioxidant defenses by a mixed culture of acidophilic bacteria in response to co 2+ exposure. *Extremophiles*, 24:485–499, 2020.

[112] Rima Siauciunaite, Nicholas S Foulkes, Viola Calabrò, and Daniela Vallone. Evolution shapes the gene expression response to oxidative stress. *International journal of molecular sciences*, 20(12):3040, 2019.

[113] Dongxu Shen, Xiaoli He, Peifang Weng, Yanan Liu, and Zufang Wu. A review of yeast: High cell-density culture, molecular mechanisms of stress response and tolerance during fermentation. *FEMS Yeast Research*, 22(1):foac050, 2022.

[114] Eleni Vasilakou, Mark CM Van Loosdrecht, and S Aljoscha Wahl. Escherichia coli metabolism under short-term repetitive substrate dynamics: adaptation and trade-offs. *Microbial Cell Factories*, 19:1–19, 2020.

[115] Martin Cárcamo, Pedro Saa, Jorge Torres, Samuel Torres, Patricio Mandujano, Jose Ricardo Perez Correa, and Eduardo Agosin. Effective dissolved oxygen control strategy for high-cell-density cultures. *IEEE Latin America Transactions*, 12(3):389–394, 2014.

[116] João RM Almeida, Léia CL Fávaro, and Betania F Quirino. Biodiesel biorefinery: opportunities and challenges for microbial production of fuels and chemicals from glycerol waste. *Biotechnology for biofuels*, 5(1):48, 2012.

[117] Zhen Chen and Dehua Liu. Toward glycerol biorefinery: metabolic engineering for the production of biofuels and chemicals from glycerol. *Biotechnology for biofuels*, 9:1–15, 2016.

[118] Syed Shams Yazdani and Ramon Gonzalez. Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. *Current opinion in biotechnology*, 18(3):213–219, 2007.

[119] Seraphim Papanikolaou and George Aggelis. Lipids of oleaginous yeasts. part ii: Technology and potential applications. *European Journal of Lipid Science and Technology*, 113(8):1052–1073, 2011.

[120] Afroditi Chatzifragkou and Seraphim Papanikolaou. Effect of impurities in biodiesel-derived waste glycerol on the performance and feasibility of biotechnological processes. *Applied microbiology and biotechnology*, 95:13–27, 2012.

[121] James M Clomburg and Ramon Gonzalez. Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals. *Trends in biotechnology*, 31(1):20–28, 2013.

[122] Duane T Johnson and Katherine A Taconi. The glycerin glut: Options for the value-added conversion of crude glycerol resulting from biodiesel production. *Environmental Progress*, 26(4):338–348, 2007.

[123] Syed Shams Yazdani and Ramon Gonzalez. Engineering escherichia coli for the efficient conversion of glycerol to ethanol and co-products. *Metabolic engineering*, 10(6):340–351, 2008.

[124] Kyung Ok Yu, Seung Wook Kim, and Sung Ok Han. Engineering of glycerol utilization pathway for ethanol production by saccharomyces cerevisiae. *Bioresource Technology*, 101(11):4157–4161, 2010.

[125] Matthew D Blankschien, James M Clomburg, and Ramon Gonzalez. Metabolic engineering of escherichia coli for the production of succinate from glycerol. *Metabolic engineering*, 12(5):409–419, 2010.

[126] Margarida Carvalho, Mariana Matos, Christophe Roca, and Maria AM Reis. Succinic acid production from glycerol by actinobacillus succinogenes using dimethylsulfoxide as electron acceptor. *New biotechnology*, 31(1):133–139, 2014.

[127] Erik Wilkens, Anne Katrin Ringel, Diana Hortig, Thomas Willke, and Klaus-Dieter Vorlop. High-level production of 1, 3-propanediol from crude glycerol by clostridium butyricum akr102a. *Applied microbiology and biotechnology*, 93:1057–1063, 2012.

[128] Ya-Nan Zhao, Guo Chen, and Shan-Jing Yao. Microbial production of 1, 3-propanediol from glycerol by encapsulated klebsiella pneumoniae. *Biochemical Engineering Journal*, 32(2):93–99, 2006.

[129] EJ Bormann and M Roth. The production of polyhydroxybutyrate by methylobacterium rhodesianum and ralstonia eutropha in media containing glycerol and casein hydrolysates. *Biotechnology Letters*, 21:1059–1063, 1999.

[130] Martin Koller and Stanislav Obruča. Biotechnological production of polyhydroxyalkanoates from glycerol: A review. *Biocatalysis and Agricultural Biotechnology*, 42:102333, 2022.

[131] SNRL Silva, CBB Farias, RD Rufino, JM Luna, and LA Sarubbo. Glycerol as substrate for the production of biosurfactant by pseudomonas aeruginosa ucp0992. *Colloids and Surfaces B: Biointerfaces*, 79(1):174–183, 2010.

[132] Seraphim Papanikolaou, Lionel Muniglia, Isabelle Chevalot, George Aggelis, and I Marc. Yarrowia lipolytica as a potential producer of citric acid from raw glycerol. *Journal of applied microbiology*, 92(4):737–744, 2002.

[133] M Paiders, A Gruduls, L Kalnina, S Valucka, I Dimanta, J Kleperis, and V Nikolajeva. Biogas and hydrogen production from glycerol by enterobacter aerogenes and anaerobic microbial communities. *Agronomy Research*, 15(1), 2017.

[134] Yuki Doi. Glycerol metabolism and its regulation in lactic acid bacteria. *Applied microbiology and biotechnology*, 103(13):5079–5093, 2019.

[135] Suman Mazumdar, James M Clomburg, and Ramon Gonzalez. Escherichia coli strains engineered for homofermentative production of d-lactic acid from glycerol. *Applied and environmental microbiology*, 76(13):4327–4336, 2010.

[136] An Zhang and Shang-Tian Yang. Propionic acid production from glycerol by metabolically engineered propionibacterium acidipropionici. *Process Biochemistry*, 44(12):1346–1351, 2009.

[137] Sweta Yadav, Garima Rawat, Priyanka Tripathi, and RK Saxena. A novel approach for biobutanol production by clostridium acetobutylicum using glycerol: a low cost substrate. *Renewable Energy*, 71:37–42, 2014.

[138] Bakht Zada, Chonglong Wang, Ji-Bin Park, Seong-Hee Jeong, Ju-Eon Park, Hawaibam Birla Singh, and Seon-Won Kim. Metabolic engineering of escherichia

coli for production of mixed isoprenoid alcohols and their derivatives. *Biotechnology for biofuels*, 11:1–12, 2018.

[139] Jaspreet Kaur, Anil Kumar Sarma, Mithilesh Kumar Jha, and Poonam Gera. Valorisation of crude glycerol to value-added products: Perspectives of process technology, economics and environmental issues. *Biotechnology Reports*, 27:e00487, 2020.

[140] Rosemary Dobson, Vincent Gray, and Karl Rumbold. Microbial utilization of crude glycerol for the production of value-added products. *Journal of Industrial Microbiology and Biotechnology*, 39(2):217–226, 2012.

[141] Fangxia Yang, Milford A Hanna, and Runcang Sun. Value-added uses for crude glycerol–a byproduct of biodiesel production. *Biotechnology for biofuels*, 5:1–10, 2012.

[142] Abhishek Murarka, Yandi Dharmadi, Syed Shams Yazdani, and Ramon Gonzalez. Fermentative utilization of glycerol by escherichia coli and its implications for the production of fuels and chemicals. *Applied and environmental microbiology*, 74(4):1124–1135, 2008.

[143] Matthew Theisen and James C Liao. Industrial biotechnology: Escherichia coli as a host. *Industrial biotechnology: microorganisms*, 1:149–181, 2017.

[144] Alaksh Choudhury, Jacob A Fenster, Reilly G Fankhauser, Joel L Kaar, Olivier Tenaillon, and Ryan T Gill. Crispr/cas9 recombineering-mediated deep mutational scanning of essential genes in escherichia coli. *Molecular systems biology*, 16(3):e9265, 2020.

[145] Yiwen Zhang, Xiaoman Sun, Qingzhuo Wang, Jiaqi Xu, Feng Dong, Siqi Yang, Jiawei Yang, Zixu Zhang, Yuan Qian, and Jun Chen. Multicopy chromosomal integration using crispr-associated transposases. *ACS Synthetic Biology*, 9(8):1998–2008, 2020.

[146] Marcelo C Bassalo, Andrew D Garst, Andrea L Halweg-Edwards, William C Grau, Dylan W Domaille, Vivek K Mutalik, Adam P Arkin, and Ryan T Gill. Rapid and efficient one-step metabolic pathway integration in e. coli. *ACS synthetic biology*, 5(7):561–568, 2016.

[147] Yong Chen, Xin-Yu Chen, He-Tong Du, Xu Zhang, Yi-Ming Ma, Jin-Chun Chen, Jian-Wen Ye, Xiao-Ran Jiang, and Guo-Qiang Chen. Chromosome engineering of the tca cycle in halomonas bluephagenesis for production of copolymers of 3-hydroxybutyrate and 3-hydroxyvalerate (phbv). *Metabolic engineering*, 54:69–82, 2019.

[148] Virginia C Dunkel, Errol Zeiger, David Brusick, Elena McCoy, Douglas McGregor, Kristien Mortelmans, Herbert S Rosenkranz, and Vincent F Simmon. Reproducibility of microbial mutagenicity assays: Ii. testing of carcinogens and noncarcinogens in salmonella typhimurium and escherichia coli. *Environmental mutagenesis*, 7(S5):1–19, 1985.

[149] Evelyn M Witkin. Ultraviolet mutagenesis and inducible dna repair in escherichia coli. *Bacteriological reviews*, 40(4):869–907, 1976.

[150] Jeff Smith, Marina Bibikova, Frank G Whitby, AR Reddy, Srinivasan Chandrasegaran, and Dana Carroll. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger dna-recognition domains. *Nucleic acids research*, 28(17):3361–3369, 2000.

[151] Jeffrey C Miller, Siyuan Tan, Guijuan Qiao, Kyle A Barlow, Jianbin Wang, Danny F Xia, Xiangdong Meng, David E Paschon, Elo Leung, and Sarah J Hinkley. A tale nuclease architecture for efficient genome editing. *Nature biotechnology*, 29(2):143–148, 2011.

[152] Alexis C Komor, Yongjoo B Kim, Michael S Packer, John A Zuris, and David R Liu. Programmable editing of a target base in genomic dna without double-stranded dna cleavage. *Nature*, 533(7603):420–424, 2016.

[153] Andrew V Anzalone, Peyton B Randolph, Jessie R Davis, Alexander A Sousa, Luke W Koblan, Jonathan M Levy, Peter J Chen, Christopher Wilson, Gregory A Newby, and Aditya Raguram. Search-and-replace genome editing without double-strand breaks or donor dna. *Nature*, 576(7785):149–157, 2019.

[154] Zen-Ichiro Horii and Alvin J Clark. Genetic analysis of the recf pathway to genetic recombination in escherichia coli k12: isolation and characterization of mutants. *Journal of molecular biology*, 80(2):327–344, 1973.

[155] RUPAP Simon, U Priefer, and Alfred Pühler. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/technology*, 1(9):784–791, 1983.

[156] ML Morse, Esther M Lederberg, and Joshua Lederberg. Transduction in escherichia coli k-12. *Genetics*, 41(1):142, 1956.

[157] Matxalen Llosa and Fernando de la Cruz. Bacterial conjugation: a potential tool for genomic engineering. *Research in Microbiology*, 156(1):1–6, 2005.

[158] Timothy M Wannier, Akos Nyerges, Helene M Kuchwara, Márton Czikkely, Dávid Balogh, Gabriel T Filsinger, Nathaniel C Borders, Christopher J Gregg, Marc J Lajoie, and Xavier Rios. Improved bacterial recombineering by parallelized protein discovery. *Proceedings of the National Academy of Sciences*, 117(24):13689–13698, 2020.

[159] JA Mosberg, MJ Lajoie, and GM Church. Lambda red recombineering in escherichia coli occurs through a fully single-stranded intermediate. *Genetics*, 186(3):791–799, 2010.

[160] Sadhana K Kulkarni and Franklin W Stahl. Interaction between the sbcc gene of escherichia coli and the gam gene of phage lambda. *Genetics*, 123(2):249–253, 1989.

[161] James A Sawitzke, Lynn C Thomason, Nina Costantino, Mikhail Bubunenko, Simanti Datta, and Donald L Court. Recombineering: in vivo genetic engineering in e. coli, s. enterica, and beyond. *Methods in enzymology*, 421:171–199, 2007.

[162] Tomoya Baba, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A Datsenko, Masaru Tomita, Barry L Wanner, and Hirotada Mori. Construction of escherichia coli k-12 in-frame, single-gene knockout mutants: the keio collection. *Molecular systems biology*, 2(1):2006.0008, 2006.

[163] Phuc Leo H Vo, Carlotta Ronda, Sanne E Klompe, Ethan E Chen, Christopher Acree, Harris H Wang, and Samuel H Sternberg. Crispr rna-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nature biotechnology*, 39(4):480–489, 2021.

[164] Devaki Bhaya, Michelle Davison, and Rodolphe Barrangou. Crispr-cas systems in bacteria and archaea: versatile small rnas for adaptive defense and regulation. *Annual review of genetics*, 45(1):273–297, 2011.

[165] Ho Joung Lee, Hyun Ju Kim, and Sang Jun Lee. Crispr-cas9-mediated pinpoint microbial genome editing aided by target-mismatched sgrnas. *Genome research*, 30(5):768–775, 2020.

[166] Yu Jiang, Biao Chen, Chunlan Duan, Bingbing Sun, Junjie Yang, and Sheng Yang. Multigene editing in the escherichia coli genome via the crispr-cas9 system. *Applied and environmental microbiology*, 81(7):2506–2514, 2015.

[167] Michael T Petassi, Shan-Chi Hsieh, and Joseph E Peters. Guide rna categorization enables target site choice in tn7-crispr-cas transposons. *Cell*, 183(7):1757–1771. e18, 2020.

[168] Ernest Fokum, Hossain M Zabed, Qi Guo, Junhua Yun, Miaomiao Yang, Hao Pang, Yingfeng An, Wen Li, and Xianghui Qi. Metabolic engineering of bacterial strains using crispr/cas9 systems for biosynthesis of value-added products. *Food bioscience*, 28:125–132, 2019.

[169] Yamin Li, Zachary Glass, Mingqian Huang, Zheng-Yi Chen, and Qiaobing Xu. Ex vivo cell-based crispr/cas9 genome editing for therapeutic applications. *Biomaterials*, 234:119711, 2020.

[170] Choun-Sea Lin, Chen-Tran Hsu, Ling-Hung Yang, Lan-Ying Lee, Jin-Yuan Fu, Qiao-Wei Cheng, Fu-Hui Wu, Han C-W Hsiao, Yesheng Zhang, and Ru Zhang. Application of protoplast technology to crispr/cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant biotechnology journal*, 16(7):1295–1310, 2018.

[171] Sanne E Klompe, Phuc LH Vo, Tyler S Halpin-Healy, and Samuel H Sternberg. Transposon-encoded crispr–cas systems direct rna-guided dna integration. *Nature*, 571(7764):219–225, 2019.

[172] Makoto Saito, Alim Ladha, Jonathan Strecker, Guilhem Faure, Edwin Neumann, Han Altae-Tran, Rhiannon K Macrae, and Feng Zhang. Dual modes of crispr-associated transposon homing. *Cell*, 184(9):2441–2453. e18, 2021.

[173] Jonathan Strecker, Alim Ladha, Zachary Gardner, Jonathan L Schmid-Burgk, Kira S Makarova, Eugene V Koonin, and Feng Zhang. Rna-guided dna insertion with crispr-associated transposases. *Science*, 365(6448):48–53, 2019.

[174] Tyler S Halpin-Healy, Sanne E Klompe, Samuel H Sternberg, and Israel S Fernández. Structural basis of dna targeting by a transposon-encoded crispr–cas system. *Nature*, 577(7789):271–274, 2020.

[175] John van der Oost and Ioannis Mougiakos. First structural insights into crispr-cas-guided dna transposition. *Cell Research*, 30(3):193–194, 2020.

[176] Beibei Wang, Wenhao Xu, and Hui Yang. Structural basis of a tn7-like transposase recruitment and dna loading to crispr-cas surveillance complex. *Cell research*, 30(2):185–187, 2020.

[177] Phuc Leo H Vo, Christopher Acree, Melissa L Smith, and Samuel H Sternberg. Unbiased profiling of crispr rna-guided transposition products by long-read sequencing. *Mobile DNA*, 12(1):1–8, 2021.

[178] Lidimarie Trujillo Rodríguez, Adam J Ellington, and Christopher R Reisch. Broad-host-range mutagenesis with crispr-associated transposase. *bioRxiv*, 2022.

[179] Junjie Yang, Jiawei Yang, Yiwen Zhang, Siqi Yang, Jieze Zhang, Yu Jiang, and Sheng Yang. Crispr-associated transposase system can insert multiple copies of donor dna into the same target locus. *The CRISPR Journal*, 4(6):789–798, 2021.

[180] Siqi Yang, Yiwen Zhang, Jiaqi Xu, Jiao Zhang, Jieze Zhang, Junjie Yang, Yu Jiang, and Sheng Yang. Orthogonal crispr-associated transposases for parallel and multiplexed chromosomal integration. *Nucleic acids research*, 49(17):10192–10202, 2021.

[181] Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison, and Hamilton O Smith. Enzymatic assembly of dna molecules up to several hundred kilobases. *Nature methods*, 6(5):343–345, 2009.

[182] Egon Amann, Birgit Ochs, and Karl-Josef Abel. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in escherichia coli. *Gene*, 69(2):301–315, 1988.

[183] David William Russell and Joseph Sambrook. *Molecular cloning: a laboratory manual*, volume 1. Cold Spring Harbor Laboratory Cold Spring Harbor, NY, 2001.

[184] Kajan Srirangan, Xuejia Liu, Adam Westbrook, Lamees Akawi, Michael E Pyne, Murray Moo-Young, and C Perry Chou. Biochemical, genetic, and metabolic engineering strategies to enhance coproduction of 1-propanol and ethanol in engineered escherichia coli. *Applied microbiology and biotechnology*, 98(22):9499–9515, 2014.

[185] Adam W Westbrook, Murray Moo-Young, and C Perry Chou. Development of a crispr-cas9 tool kit for comprehensive engineering of bacillus subtilis. *Applied and environmental microbiology*, 82(16):4876–4895, 2016.

[186] Peter P Cherepanov and Wilfried Wackernagel. Gene disruption in escherichia coli: Tcr and kmr cassettes with the option of flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, 158(1):9–14, 1995.

[187] Dragan Miscevic, Ju-Yi Mao, Bradley Mozell, Kajan Srirangan, Daryoush Abedi, Murray Moo-Young, and C Perry Chou. Bio-based production of poly (3-hydroxybutyrate-co-3-hydroxyvalerate) with modulated monomeric fraction in escherichia coli. *Applied Microbiology and Biotechnology*, 105(4):1435–1446, 2021.

[188] Ayaka Hiroe, Kenji Tsuge, Christopher T Nomura, Mitsuhiro Itaya, and Takeharu Tsuge. Rearrangement of gene order in the phacab operon leads to effective production of ultrahigh-molecular-weight poly [(r)-3-hydroxybutyrate] in genetically engineered escherichia coli. *Applied and environmental microbiology*, 78(9):3177–3184, 2012.

[189] Yohei Katoh, Saki Michisaka, Shohei Nozaki, Teruki Funabashi, Tomoaki Hirano, Ryota Takei, and Kazuhisa Nakayama. Practical method for targeted disruption of cilia-related genes by using crispr/cas9-mediated, homology-independent knock-in system. *Molecular biology of the cell*, 28(7):898–906, 2017.

[190] Karan Sukhija, Michael Pyne, Saad Ali, Valerie Orr, Daryoush Abedi, Murray Moo-Young, and C Perry Chou. Developing an extended genomic engineering approach based on recombineering to knock-in heterologous genes to escherichia coli genome. *Molecular biotechnology*, 51(2):109–118, 2012.

[191] Adolfo Rivero-Müller, Svetlana Lajić, and Ilpo Huhtaniemi. Assisted large fragment insertion by red/et-recombination (alfire)—an alternative and enhanced method for large fragment recombineering. *Nucleic acids research*, 35(10):e78, 2007.

[192] Michael E Pyne, Murray Moo-Young, Duane A Chung, and C Perry Chou. Coupling the crispr/cas9 system with lambda red recombineering enables simplified chromosomal gene replacement in escherichia coli. *Applied and environmental microbiology*, 81(15):5103–5114, 2015.

[193] Marcello Maresca, Axel Erler, Jun Fu, Anne Friedrich, Youming Zhang, and Afrancis Stewart. Single-stranded heteroduplex intermediates in λ red homologous recombination. *BMC molecular biology*, 11(1):1–15, 2010.

[194] Yifan Li, Zhenquan Lin, Can Huang, Yan Zhang, Zhiwen Wang, Ya-jie Tang, Tao Chen, and Xueming Zhao. Metabolic engineering of escherichia coli using crispr–cas9 meditated genome editing. *Metabolic engineering*, 31:13–21, 2015.

[195] Lun Cui and David Bikard. Consequences of cas9 cleavage in the chromosome of escherichia coli. *Nucleic acids research*, 44(9):4243–4251, 2016.

[196] Kenji Okano, Yu Sato, Tatsuya Hizume, and Kohsuke Honda. Genome editing by miniature crispr/cas12f1 enzyme in escherichia coli. *Journal of bioscience and bioengineering*, 132(2):120–124, 2021.

[197] Kajan Srirangan, Xuejia Liu, Tam T Tran, Trevor C Charles, Murray Moo-Young, and C Perry Chou. Engineering of escherichia coli for direct and modulated biosynthesis of poly (3-hydroxybutyrate-co-3-hydroxyvalerate) copolymer using unrelated carbon sources. *Scientific reports*, 6(1):1–11, 2016.

[198] Y-T Horng, C-C Chien, C-T Huang, Y-H Wei, S-Y Chen, JC-W Lan, and P-C Soo. Biosynthesis of poly (3-hydroxybutyrate-co-3-hydroxyvalerate) with co-expressed propionate permease (prpp), beta-ketothiolase b (bktb), and propionate-coa synthase (prpe) in escherichia coli. *Biochemical engineering journal*, 78:73–79, 2013.

[199] Jung Eun Yang, Yong Jun Choi, Se Jin Lee, Kyoung-Hee Kang, Hyuk Lee, Young Hoon Oh, Seung Hwan Lee, Si Jae Park, and Sang Yup Lee. Metabolic engineering of escherichia coli for biosynthesis of poly (3-hydroxybutyrate-co-3-hydroxyvalerate) from glucose. *Applied microbiology and biotechnology*, 98(1):95–104, 2014.

[200] Dragan Miscevic, Ju-Yi Mao, Murray Moo-Young, and Chih-Hsiung Perry Chou. High-level heterologous production of propionate in engineered escherichia coli. *Biotechnology and Bioengineering*, 117(5):1304–1315, 2020.

[201] Anu Jose Mattam, James M Clomburg, Ramon Gonzalez, and Syed Shams Yazdani. Fermentation of glycerol and production of valuable chemical and biofuel molecules. *Biotechnology letters*, 35(6):831–842, 2013.

[202] Lawrence P Cook, Greg Brewer, and Winnie Wong-Ng. Structural aspects of porphyrins for functional materials applications. *Crystals*, 7(7):223, 2017.

[203] Yingyan Shi, Fuming Zhang, and Robert J Linhardt. Porphyrin-based compounds and their applications in materials and medicine. *Dyes and Pigments*, 188:109136, 2021.

[204] Qi Zhou, Atsushi Yamada, Qingguo Feng, Austin Hoskins, Barry D Dunietz, and Kim M Lewis. Modification of molecular conductance by in situ deprotection of thiol-based porphyrin. *ACS Applied Materials & Interfaces*, 9(19):15901–15906, 2017.

[205] Sruti Mondal, Tanmoy Pain, Kasturi Sahu, and Sanjib Kar. Large-scale green synthesis of porphyrins. *ACS omega*, 6(35):22922–22936, 2021.

[206] Satoru Hiroto, Yoshihiro Miyake, and Hiroshi Shinokubo. Synthesis and functionalization of porphyrins through organometallic methodologies. *Chemical reviews*, 117(4):2910–3043, 2017.

[207] SF MacDonald and RJ Stedman. The synthesis of uroporphyrin i. *Journal of the American Chemical Society*, 75(12):3040–3041, 1953.

[208] Abraham Goldberg and Claude Rimington. Experimentally produced porphyria in animals. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, 143(911):257–279, 1955.

[209] Constantin A Rebeiz, Mounir Abou Haidar, Moustapha Yaghi, and Paul A Castelfranco. Porphyrin biosynthesis in cell-free homogenates from higher plants. *Plant Physiology*, 46(4):543–549, 1970.

[210] Alfred Hassner and Irishi Namboothiri. *Organic syntheses based on name reactions: a practical guide to 750 transformations*. Elsevier, 2012.

[211] M Moulin and AG Smith. Regulation of tetrapyrrole biosynthesis in higher plants. *Biochemical Society Transactions*, 33(4):737–742, 2005.

[212] Lipeng Wu, Takahiko Moteki, Amit A Gokhale, David W Flaherty, and F Dean Toste. Production of fuels and chemicals from biomass: condensation reactions and beyond. *Chem*, 1(1):32–58, 2016.

[213] Luděk Kořený, Miroslav Oborník, Eva Horáková, Ross F Waller, and Julius Lukeš. The convoluted history of haem biosynthesis. *Biological Reviews*, 97(1):141–162, 2022.

[214] Meiru Jiang, Kunqiang Hong, Yufeng Mao, Hongwu Ma, Tao Chen, and Zhiwen Wang. Natural 5-aminolevulinic acid: Sources, biosynthesis, detection and applications. *Frontiers in bioengineering and biotechnology*, 10:841443, 2022.

[215] Rohit Sharma, Sayonara M Viana, Dennis KP Ng, Bala K Kolli, Kwang Poo Chang, and Camila I de Oliveira. Photodynamic inactivation of leishmania braziliensis doubly sensitized with uroporphyrin and diamino-phthalocyanine activates effector functions of macrophages in vitro. *Scientific Reports*, 10(1):17065, 2020.

[216] Dragan Miscevic, Ju-Yi Mao, Teshager Kefale, Daryoush Abedi, Murray Moo-Young, and C Perry Chou. Strain engineering for high-level 5-aminolevulinic acid production in escherichia coli. *Biotechnology and Bioengineering*, 118(1):30–42, 2021.

[217] Dragan Miscevic, Ju-Yi Mao, Teshager Kefale, Daryoush Abedi, Chih-Ching Huang, Murray Moo-Young, and C Perry Chou. Integrated strain engineering and bioprocessing strategies for high-level bio-based production of 3-hydroxyvalerate in escherichia coli. *Applied Microbiology and Biotechnology*, 104:5259–5272, 2020.

[218] Davinder Lall, Dragan Miscevic, Mark Bruder, Adam Westbrook, Marc Aucoin, Murray Moo-Young, and C Perry Chou. Strain engineering and bioprocessing strategies for biobased production of porphobilinogen in escherichia coli. *Bioresources and Bioprocessing*, 8:1–16, 2021.

[219] Junli Zhang, Zhen Kang, Jian Chen, and Guocheng Du. Optimization of the heme biosynthesis pathway for the production of 5-aminolevulinic acid in escherichia coli. *Scientific Reports*, 5(1):1–7, 2015.

[220] Sébastien Zappa, Keran Li, and Carl E Bauer. The tetrapyrrole biosynthetic pathway and its regulation in rhodobacter capsulatus. *Recent advances in phototrophic prokaryotes*, pages 229–250, 2010.

[221] Anne Mathilde Davy, Helene Faustrup Kildegaard, and Mikael Rørdam Andersen. Cell factory engineering. *Cell systems*, 4(3):262–275, 2017.

[222] George H McArthur IV and Stephen S Fong. Toward engineering synthetic microbial metabolism. *BioMed Research International*, 2010(1):459760, 2010.

[223] Adam W Westbrook, Dragan Miscevic, Shane Kilpatrick, Mark R Bruder, Murray Moo-Young, and C Perry Chou. Strain engineering for microbial production of value-added chemicals and fuels from glycerol. *Biotechnology advances*, 37(4):538–568, 2019.

[224] Frederick C Neidhardt, Philip L Bloch, and David F Smith. Culture medium for enterobacteria. *Journal of bacteriology*, 119(3):736–747, 1974.

[225] David Mauzerall and Saul Granick. The occurrence and determination of δ-aminolevulinic acid and porphobilinogen in urine. *Journal of Biological Chemistry*, 219(1):435–446, 1956.

[226] Thomas L Poulos. Heme enzyme structure and function. *Chemical reviews*, 114(7):3919–3962, 2014.

[227] Jordi Zamarreno Beas, Marco AM Videira, and Lígia M Saraiva. Regulation of bacterial haem biosynthesis. *Coordination chemistry reviews*, 452:214286, 2022.

[228] Arianna I Celis and Jennifer L DuBois. Making and breaking heme. *Current opinion in structural biology*, 59:19–28, 2019.

[229] John D Phillips. Heme biosynthesis and the porphyrias. *Molecular genetics and metabolism*, 128(3):164–177, 2019.

[230] Masoud Aman Mohammadi, Hossein Ahangari, Saeed Mousazadeh, Seyede Marzieh Hosseini, and Laurent Dufossé. Microbial pigments as an alternative to synthetic dyes and food additives: A brief review of recent studies. *Bioprocess and biosystems engineering*, 45(1):1–12, 2022.

[231] Zhijie Qin, Xinglong Wang, Song Gao, Dong Li, and Jingwen Zhou. Production of natural pigments using microorganisms. *Journal of Agricultural and Food Chemistry*, 2023.

[232] Dongsoo Yang, Seon Young Park, Yae Seul Park, Hyunmin Eun, and Sang Yup Lee. Metabolic engineering of escherichia coli for natural product biosynthesis. *Trends in Biotechnology*, 38(7):745–765, 2020.

[233] N Frankenberg, J Moser, and D Jahn. Bacterial heme biosynthesis and its biotechnological application. *Applied microbiology and biotechnology*, 63:115–127, 2003.

[234] Hongfei Su, Xiaolin Chen, Shijing Chen, Mingzhang Guo, and Huilin Liu. Applications of the whole-cell system in the efficient biosynthesis of heme. *International Journal of Molecular Sciences*, 24(9):8384, 2023.

[235] Gunhild Layer. Heme biosynthesis in prokaryotes. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1868(1):118861, 2021.

[236] Elena Di Pierro, Michele De Canio, Rosa Mercadante, Maria Savino, Francesca Granata, Dario Tavazzi, Anna Maria Nicolli, Andrea Trevisan, Stefano Marchini, and Silvia Fustinoni. Laboratory diagnosis of porphyria. *Diagnostics*, 11(8):1343, 2021.

[237] Abdulrahman Al-Hussaini, Ali Asery, and Omar Alharbi. Urinary coproporphyrins as a diagnostic biomarker of dubin-johnson syndrome in neonates: A diagnostic pathway is proposed. *Saudi Journal of Gastroenterology*, page 10.4103, 2023.

[238] Mikko Neuvonen, Aleksi Tornio, Päivi Hirvensalo, Janne T Backman, and Mikko Niemi. Performance of plasma coproporphyrin i and iii as oatp1b1 biomarkers in humans. *Clinical Pharmacology & Therapeutics*, 110(6):1622–1632, 2021.

[239] IM Belousova, MV Dobrun, LV Galebskaya, SI Gorelov, IM Kislyakov, SE Kolbasov, AV Kris' Ko, TK Kris' ko, MA Malkov, and TD Murav'eva. New preparation based on coproporphyrin iii for photoluminescence diagnostics and photodynamic therapy. In *Laser Optics 2010*, volume 7822, pages 200–205. SPIE.

[240] Ravindranath Reddy Gilibili, Sagnik Chatterjee, Pravin Bagul, Kathleen W Mosure, Bokka Venkata Murali, T Thanga Mariappan, Sandhya Mandlekar, and Yurong Lai. Coproporphyrin-i: a fluorescent, endogenous optimal probe substrate for abcc2 (mrp2) suitable for vesicle-based mrp2 inhibition assay. *Drug Metabolism and Disposition*, 45(6):604–611, 2017.

[241] Alec B Walter, Jocelyn Simpson, J Logan Jenkins, Eric P Skaar, and E Duco Jansen. Optimization of optical parameters for improved photodynamic therapy of staphylococcus aureus using endogenous coproporphyrin iii. *Photodiagnosis and photodynamic therapy*, 29:101624, 2020.

[242] J Andrew Jones and Mattheos AG Koffas. *Optimizing metabolic pathways for the improved production of natural products*, volume 575, pages 179–193. Elsevier, 2016.

[243] AJ Makoff and AE Smallwood. The use of two-cistron constructions in improving the expression of a heterologous gene in e. coli. *Nucleic acids research*, 18(7):1711–1718, 1990.

[244] Sarah F Newbury, Noel H Smith, and Christopher F Higgins. Differential mrna stability controls relative gene expression within a polycistronic operon. *Cell*, 51(6):1131–1143, 1987.

[245] Heather Panek and Mark R O'Brian. A whole genome view of prokaryotic haem biosynthesis. *Microbiology*, 148(8):2273–2282, 2002.

[246] Harry A Dailey, Alecia N Septer, Lauren Daugherty, Daniel Thames, Svetlana Gerdes, Eric V Stabb, Anne K Dunn, Tamara A Dailey, and John D Phillips. The escherichia coli protein yfex functions as a porphyrinogen oxidase, not a heme dechelatase. *MBio*, 2(6):10.1128/mbio. 00248–11, 2011.

[247] Prem Ponka. Cell biology of heme. *The American journal of the medical sciences*, 318(4):241–256, 1999.

[248] Yilin Xing, Shanxing Gao, Xinyu Zhang, and Jiachen Zang. Dietary heme-containing proteins: Structures, applications, and challenges. *Foods*, 11(22):3594, 2022.

[249] Ibrahim A Abdelazim, Mohannad Abu-Faza, Svetlana Shikanova, Gulmira Zhurabekova, and Manal M Maghrabi. Heme-bound iron in treatment of pregnancy-associated iron deficiency anemia. *Journal of family medicine and primary care*, 7(6):1434–1438, 2018.

[250] Suzanne M Adam, Gayan B Wijeratne, Patrick J Rogler, Daniel E Diaz, David A Quist, Jeffrey J Liu, and Kenneth D Karlin. Synthetic fe/cu complexes: toward understanding heme-copper oxidase structure and function. *Chemical reviews*, 118(22):10840–11022, 2018.

[251] Namhyun Chung, Kapsung Park, David K Stevens, and Guyoung Kang. Verification of heme catalytic cycle with 5-aminosalicylic acid and its application to soil remediation of polycyclic aromatic hydrocarbons. *Environmental Engineering Research*, 19(2):139–143, 2014.

[252] Qiuyu Yang, Juntao Zhao, Yangyang Zheng, Tao Chen, and Zhiwen Wang. Microbial synthesis of heme b: biosynthetic pathways, current strategies, detection, and future prospects. *Molecules*, 28(8):3633, 2023.

[253] Gunhild Layer, Martina Jahn, Jurgen Moser, and Dieter Jahn. Radical sam enzymes involved in tetrapyrrole biosynthesis and insertion. *ACS bio & med Chem Au*, 2(3):196–204, 2022.

[254] Laurie K Jackson, Tammy A Dailey, Brenden Anderle, Martin J Warren, Hector A Bergonia, Harry A Dailey, and John D Phillips. Exploiting differences in heme biosynthesis between bacterial species to screen for novel antimicrobials. *Biomolecules*, 13(10):1485, 2023.

[255] Jianzhong Ge, Xiaolu Wang, Yingguo Bai, Yaru Wang, Yuan Wang, Tao Tu, Xing Qin, Xiaoyun Su, Huiying Luo, and Bin Yao. Engineering escherichia coli for efficient assembly of heme proteins. *Microbial Cell Factories*, 22(1):59, 2023.

[256] Tugba Sagir, Salih Gencer, Nurufe Kemikli, M Fatih Abasiyanik, Sevim Isik, and Ramazan Ozturk. Photodynamic activities of protoporphyrin ix and its dopamine conjugate against cancer and bacterial cell viability. *Medicinal Chemistry Research*, 21:4499–4505, 2012.

[257] Jacob E Choby and Eric P Skaar. Heme synthesis and acquisition in bacterial pathogens. *Journal of molecular biology*, 428(17):3408–3428, 2016.

[258] Aidan Tinafar, Katariina Jaenes, and Keith Pardee. Synthetic biology goes cell-free. *BMC biology*, 17:1–14, 2019.

[259] Jan-Karl Guterl and Volker Sieber. Biosynthesis "debugged": novel bioproduction strategies. *Engineering in Life Sciences*, 13(1):4–18, 2013.

[260] C Eric Hodgman and Michael C Jewett. Cell-free synthetic biology: thinking outside the cell. *Metabolic engineering*, 14(3):261–269, 2012.

[261] Quentin M Dudley, Ashty S Karim, and Michael C Jewett. Cell-free metabolic engineering: biomanufacturing beyond the cell. *Biotechnology journal*, 10(1):69–82, 2015.

[262] Jonas Fyrestam and Conny Östman. Determination of heme in microorganisms using hplc-ms/ms and cobalt (iii) protoporphyrin ix inhibition of heme acquisition in escherichia coli. *Analytical and Bioanalytical Chemistry*, 409:6999–7010, 2017.

[263] Mats Hansson and Lars Hederstedt. Purification and characterisation of a water-soluble ferrochelatase from bacillus subtilis. *European journal of biochemistry*, 220(1):201–208, 1994.

[264] Tamara A Dailey, Tye O Boynton, Angela-Nadia Albetel, Svetlana Gerdes, Michael K Johnson, and Harry A Dailey. Discovery and characterization of hemq: an essential heme biosynthetic pathway component. *Journal of Biological Chemistry*, 285(34):25978–25986, 2010.

[265] Harry A Dailey and Svetlana Gerdes. Hemq: an iron-coproporphyrin oxidative decarboxylase for protoheme synthesis in firmicutes and actinobacteria. *Archives of biochemistry and biophysics*, 574:27–35, 2015.

[266] Christine Nicole S Santos and Gregory Stephanopoulos. Combinatorial engineering of microbes for optimizing cellular phenotype. *Current opinion in chemical biology*, 12(2):168–176, 2008.

[267] Márcia R Couto, Joana L Rodrigues, Adelaide Braga, Oscar Dias, and Lígia R Rodrigues. Optimization of chondroitin production in e. coli using genome scale models. *Molecular Systems Design & Engineering*, 9(6):597–611, 2024.

[268] Harry A Dailey. Enzymes of heme biosynthesis. *JBIC Journal of Biological Inorganic Chemistry*, 2(4):411–417, 1997.

[269] Adam D Silverman, Ashty S Karim, and Michael C Jewett. Cell-free gene expression: an expanded repertoire of applications. *Nature Reviews Genetics*, 21(3):151–170, 2020.

[270] David C Mauzerall. Evolution of porphyrins. *Clinics in dermatology*, 16(2):195–201, 1998.

[271] Hagen Schlicke, Andreas Richter, Maxi Rothbart, Pawel Brzezowski, Boris Hedtke, and Bernhard Grimm. Function of tetrapyrroles, regulation of tetrapyrrole metabolism and methods for analyses of tetrapyrroles. *Procedia Chemistry*, 14:171–175, 2015.

[272] David Shemin. Porphyrin synthesis: some particular approaches. *Annals of the New York Academy of Sciences*, 244(1):348–355, 1975.

[273] Samuel I Beale. Biosynthesis of hemes. *EcoSal Plus*, 2(2):10.1128/ecosalplus. 3.6. 3.11, 2007.

[274] Peter M Jordan and David Shemin. *11 δ-Aminolevulinic Acid Synthetase*, volume 7, pages 339–356. Elsevier, 1972.

[275] DAE-KYUNG KANG, SANG-SUK KIM, WON-JAE CHI, SOON-KWANG HONG, HA-KUN KIM, and HYUN-UK KIM. Cloning and expression of the rhodobacter capsulatus hema gene in e. coli for the production of s-aminolevulinic acid. *Journal of microbiology and biotechnology*, 14(6):1327–1332, 2004.

[276] David Shemin. On the synthesis of heme. *Naturwissenschaften*, 57(4):185–190, 1970.

[277] Michael AA Mathews, Heidi L Schubert, Frank G Whitby, Kelly J Alexander, Kevin Schadick, Hector A Bergonia, John D Phillips, and Christopher P Hill. Crystal structure of human uroporphyrinogen iii synthase. *The EMBO journal*, 2001.

[278] Amy E Medlock and Harry A Dailey. Heme biosynthesis. *Metalloprotein Active Site Assembly*, page 299, 2017.

[279] Elena Di Pierro, Valentina Brancaleoni, and Francesca Granata. Advances in understanding the pathogenesis of congenital erythropoietic porphyria. *British Journal of Haematology*, 173(3):365–379, 2016.

[280] GH Elder and AG Roberts. Uroporphyrinogen decarboxylase. *Journal of bioenergetics and biomembranes*, 27:207–214, 1995.

[281] Marcel Zámocký, Stefan Hofbauer, Thomas Gabler, and Paul G Furtmüller. The molecular evolution, structure, and function of coproporphyrinogen oxidase and protoporphyrinogen oxidase in prokaryotes. *Biology*, 12(12):1527, 2023.

[282] Volodymyr I Lushchak. Free radicals, reactive oxygen species, oxidative stress and its classification. *Chemico-biological interactions*, 224:164–175, 2014.

[283] Jevas C Ozougwu. The role of reactive oxygen species and antioxidants in oxidative stress. *International Journal of Research*, 1(8):1–8, 2016.

[284] Canghai Lu, William E Bentley, and Govind Rao. Comparisons of oxidative stress response genes in aerobic escherichia coli fermentations. *Biotechnology and bioengineering*, 83(7):864–870, 2003.

[285] YS Voronkova, OS Voronkova, VA Gorban, and KK Holoborodko. Oxidative stress, reactive oxygen species, antioxidants: a review. *Ecology and noospherology*, 29(1):52–55, 2018.

[286] Agnieszka Gegotek and Elzbieta Skrzydlewska. Ascorbic acid as antioxidant. *Vitamins and Hormones*, 121:247–270, 2023.

[287] Vitaly A Roginsky and Hartmut B Stegmann. Ascorbyl radical as natural indicator of oxidative stress: quantitative regularities. *Free radical biology and medicine*, 17(2):93–103, 1994.

[288] Yoshinobu Matsumura, Masahiro Takagi, and Tadayuki Imanaka. Regulation of escherichia coli superoxide dismutase genes (soda and sodb) by oxygen. *Biotechnology letters*, 15:229–234, 1993.

[289] JA Fee. Regulation of sod genes in escherichia coli: relevance to superoxide dismutase function. *Molecular microbiology*, 5(11):2599–2610, 1991.

[290] Noriyuki Doukyu and Katsuya Taguchi. Involvement of catalase and superoxide dismutase in hydrophobic organic solvent tolerance of escherichia coli. *Amb Express*, 11(1):97, 2021.

[291] H Sakamoto and D Touati. Cloning of the iron superoxide dismutase gene (sodb) in escherichia coli k-12. *Journal of bacteriology*, 159(1):418–420, 1984.

[292] DANItLE Touati. Cloning and mapping of the manganese superoxide dismutase gene (soda) of escherichia coli k-12. *Journal of bacteriology*, 155(3):1078–1087, 1983.

[293] KR Imlay and James A Imlay. Cloning and analysis of sodc, encoding the copper-zinc superoxide dismutase of escherichia coli. *Journal of bacteriology*, 178(9):2564–2571, 1996.

[294] Karen A Hopkin, Madeline A Papazian, and Howard M Steinman. Functional differences between manganese and iron superoxide dismutases in escherichia coli k-12. *Journal of Biological Chemistry*, 267(34):24253–24258, 1992.

[295] HE Schellhorn. Regulation of hydroperoxidase (catalase) expression in escherichia coli. *FEMS microbiology letters*, 131(2):113–119, 1995.

[296] Scott M Brown, Michael L Howell, Michael L Vasil, Anne J Anderson, and Daniel J Hassett. Cloning and characterization of the katb gene of pseudomonas aeruginosa encoding a hydrogen peroxide-inducible catalase: purification of katb, cellular localization, and demonstration that it is essential for optimal resistance to hydrogen peroxide. *Journal of Bacteriology*, 177(22):6536–6544, 1995.

[297] A Heimberger and A Eisenstark. Compartmentalization of catalases in escherichia coli. *Biochemical and biophysical research communications*, 154(1):392–397, 1988.

[298] Alex Hillar, Larry Van Caeseele, and Peter C Loewen. Intracellular location of catalase-peroxidase hydroperoxidase i of escherichia coli. *FEMS microbiology letters*, 170(2):307–312, 1999.

[299] Ingemar von Ossowski, Michael R Mulvey, Pamela A Leco, Andrew Borys, and Peter C Loewen. Nucleotide sequence of escherichia coli kate, which encodes catalase hpii. *Journal of bacteriology*, 173(2):514–520, 1991.

[300] JA Imlay and Irwin Fridovich. Assay of metabolic superoxide production in escherichia coli. *Journal of Biological Chemistry*, 266(11):6957–6965, 1991.

[301] James A Imlay. Cellular defenses against superoxide and hydrogen peroxide. *Annu. Rev. Biochem.*, 77(1):755–776, 2008.

[302] DP Shivaprasad, Neetu Kumra Taneja, Anupama Lakra, and Divya Sachdev. In vitro and in situ abrogation of biofilm formation in e. coli by vitamin c through ros generation, disruption of quorum sensing and exopolysaccharide production. *Food Chemistry*, 341:128171, 2021.

[303] M Tabak, R Armon, G Rosenblat, E Stermer, and I Neeman. Diverse effects of ascorbic acid and palmitoyl ascorbate on helicobacter pylori survival and growth. *FEMS microbiology letters*, 224(2):247–253, 2003.

[304] Mehrdad Tajkarimi and Salam A Ibrahim. Antimicrobial activity of ascorbic acid alone or in combination with lactic acid on escherichia coli o157: H7 in laboratory medium and carrot juice. *Food Control*, 22(6):801–804, 2011.

[305] Sai Zhang, Yaodong He, Biswarup Sen, Xiaohong Chen, Yunxuan Xie, Jay D Keasling, and Guangyi Wang. Alleviation of reactive oxygen species enhances

pufa accumulation in schizochytrium sp. through regulating genes involved in lipid metabolism. *Metabolic engineering communications*, 6:39–48, 2018.

[306] Zhexian Geng, Jinxia Ge, Wei Cui, Hui Zhou, Jieying Deng, and Baocai Xu. Efficient de novo biosynthesis of heme by membrane engineering in escherichia coli. *International Journal of Molecular Sciences*, 23(24):15524, 2022.

[307] Maria D Esteve-Gassent, Trever C Smith, Christina M Small, Derek P Thomas, and J Seshu. Absence of soda increases the levels of oxidation of key metabolic determinants of borrelia burgdorferi. *PloS one*, 10(8):e0136707, 2015.

[308] Amy Strohmeier Gort, Daniel M Ferber, and James A Imlay. The regulation and role of the periplasmic copper, zinc superoxide dismutase of escherichia coli. *Molecular microbiology*, 32(1):179–191, 1999.

[309] ShylajaNaciyar Mohandass, Mangalalakshmi Ragavan, Dineshbabu Gnanasekaran, Uma Lakshmanan, Prabaharan Dharmar, and Sushanta Kumar Saha. Overexpression of cu/zn superoxide dismutase (cu/zn sod) in synechococcus elongatus pcc 7942 for enhanced azo dye removal through hydrogen peroxide accumulation. *Biology*, 10(12):1313, 2021.

[310] Heather R Panek and Mark R O'Brian. Katg is the primary detoxifier of hydrogen peroxide produced by aerobic metabolism in bradyrhizobium japonicum. *Journal of bacteriology*, 186(23):7874–7880, 2004.

[311] Ludmil Benov and Fatima Sequeira. Escherichia coli Δ fur mutant displays low hpii catalase activity in stationary phase. *Redox report*, 8(6):379–383, 2003.

[312] Stefano Mancini and James A Imlay. The induction of two biosynthetic enzymes helps e scherichia coli sustain heme synthesis and activate catalase during hydrogen peroxide stress. *Molecular microbiology*, 96(4):744–763, 2015.

[313] Surabhi Mishra and James Imlay. Why do bacteria use so many enzymes to scavenge hydrogen peroxide? *Archives of biochemistry and biophysics*, 525(2):145–160, 2012.

[314] Oreste Arrigoni and Mario C De Tullio. Ascorbic acid: much more than just an antioxidant. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1569(1-3):1–9, 2002.

APPENDICES

Supplementary Tables

Table S3.1: Oligomers used in Chapter 3

Primers/ oligo name	**Primer/oligo sequence (5' → 3')**
P001	GACATTTTCCTCGAGGTCATTTCCGTTAAGACCCACTTTCACATTTAAG
P002	GGCTTTCTTCTGTTCACCTG
P003	TTCTTTCAGGTGAACAGAAGAAAGC
P004	AGCAAAACTACTGCAGTAAGGACTCGGATTCTTCGTCTGTTTCTACTGG
P005	GAGTCCTTACTGCAGTAGTTTTGC
P006	GCTGCAGGAATTCAAGGCGATTAAGTTGGGTAACG
P007	GAACCCCAGAGTCCCGCTGTCAGACCAAGTTTAC
P008	CGGAAATGACCTCGAGGAAAATG
P009	ACTTAATCGCCTTGAATTCCTGCAGCCCTGGCG
P010	GGTCTGACAGCGGGAGCTCTCGAACCCCAGAGTC
P011	TAATTTTGTTTAACTTTAAGAAGGAGATATACATATGACGCGTGAAGTGGTAG
P012	TTAATGCTAGCTCAGATACGCTCGAAGATGG
P013	AGCGTATCTGAGCTAGCATTAACTTTAAAAAGGAGGAAGAATTCATGACTCAG
P014	CTCATTAGGCGGGCTGCAC
P015	ATTAAAAAAATTATAAGAGCTCTCGCTGTCAGACCAAGTTTACTCATATATAC
P016	TTATCATTTTATGGTTGTATCAACACAGATCAATTCGCGCTAACTC
P017	GAACCGGTGCAGCCCGCCTAATGAGC
P018	CGAGAGCTCTTATAATTTTTTTAATCTG
P019	TGTTGATACAACCATAAAATGATAATTAC
P020	CTTCTTAAAGTTAAACAAAATTATTTCTAGAAATTGTTATCCGCTCACAATTC
P021	TAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCGACCGGCAAAGGC
P022	TTAATGCTAGCTCATGCCTTGGCTTTGACGTATCG
P023	CCAAGGCATGAGCTAGCATTAACTTTAAAAAGGAGGATAAGATAATGACTGACG
P024	GCGGGCTGCACCGGTTTATTTGCGCTCGACTGCC
P025	CGCAAATAAACCGGTGCAGCCCGCCTAATGAGC
P026	GAACCGCCAGAGCTACTTGGG
P027	ACTTGGCTTGGCCGCACCATAGC

P028	ATGCTATGGTGCGGCCAAGCCAAG
P029	CCTTGCTGCATCTACTGGCTTCGC
P030	CGAAGCCAGTAGATGCAGCAAGG
P031	AATCACTGGATATACCACCGTTG
P032	ATATCAACGGTGGTATATCCAG
P033	ATTGCCAACCCAAGTAGCTC
P034	ATAACGAAATGCACTTCTATATTCCGGCCCTGAAAGCG
P035	TTCACGCTTTCAGGGCCGGAATATAGAAGTGCATTTCG
P036	ATAACATAGAACAGCGTGCCTTCGTTCGGCGTTTTACG
P037	TTCACGTAAAACGCCGAACGAAGGCACGCTGTTCTATG
P038	ATAACCTTCGCATGACTATTGGCAACATTAACGACTGG
P039	TTCACCAGTCGTTAATGTTGCCAATAGTCATGCGAAGG
P040	GCTTTGTTGATACATCCGCATATCAACGCCTCGC
P041	TCTCGTTCAGGCAGCCACCAGATATAACGGCAAG
P042	ATGCGGGTAACCCATGCAGACGCCGCACCACTTC
P043	CGAGCTCGGTACCCCCGACAAGAAGTGGAACATGGC
P044	TATATCTGGTGGCTGCCTGAACGAGAAGCTATC
P045	GCGGCGTCTGCATGGGTTACCCGCATAGGTG
P046	TTCTTGTCGGGGGTACCGAGCTCGTCGATCAGACCAGTTTTTAATTTGTG
P047	GATATGCGGATGTATCAACAAAGCCAGATCAATTCGCGCTAACTC
P048	ACGCGTACCATGGGCCAGCGGCATGATGGTAAATAC
P049	GTCTTAAGACGTCGCTGGAAGTAAAGGTCAAGTGC
P050	AGACGAAGAATCCGCGAATAACTTTGAACAGCTG
P051	GCTGCAGGAATTCCGAAATTATCCAGACAGGCG
P052	CTTTACTTCCAGCGACGTCTTAAGACCCACTTTCAC
P053	CAAAGTTATTCGCGGATTCTTCGTCTGTTTCTACTGG
P054	GGATAATTTCGGAATTCCTGCAGCGCAAACATTCTTGACACTCC
P055	ATGCCGCTGGCCCATGGTACGCGTGCACAATATAAATGTGAAATACTTCAC
P056	AACGAAACGCTGGATATGTGGTTTTAGAGCTAGAAATAGCAAG
P057	AGAACGATCGCGTTGGTACGGTTTTAGAGCTAGAAATAGCAAG
P058	CAGGTTCAGTCTTTAACGCGGTTTTAGAGCTAGAAATAGCAAG
P059	AAACTTCGGCGGCGAGCAGGGTTTTAGAGCTAGAAATAGCAAG
P060	GCATGCATTATAACAAACCTC

P061	TTCAGCAAGAAATGGTACCGGTTTTAGAGCTAGAAATAGCAAG
P062	ACCATTCAAAACAGCAAAAAAAGCACCGACTCGG
P063	GTCGGTGCTTTTTTTGCTGTTTTGAATGGTTCCAACAAG
P064	TGCGTCATCGCCATTTCAGTCACCTCCTAGCTGACTC
P065	CTAGGAGGTGACTGAAATGGCGATGACGCATCC
P066	TTTGATATCGAGCTCTGTCAGACCAAGTTTACTC
P067	ACTTGGTCTGACAGAGCTCGATATCAAATTACGCC
P068	CTAAAACCGGTACCATTTCTTGCTGAATCTCTATCACTGATAGGGAGTG
P069	GTTCCGCTCACACTTCCAC
P070	TATATGCACTCAGCGCTGTG
P071	CGCTCGGCTGTTTCTGGATG
P072	CGGTCTGGATCACGTTGCC
P073	GAGTTTGACTTCCTGATGACC
P074	ATATCGTTGCCATCCTGCAC
P075	CAGATGTAACAACCCTGGCATGC
P076	CGATCTTGACGCCTGCCAG
P077	TTGCTGCTCACACTTGCTCC
P078	AGACACCAGCGTTGCGTG
P079	TGCTCTCTAAAGTCTTCCCG
P080	CAAGCGATTTCACCAACGG

Table S3.2: Statistical analysis data for integration efficiency based on transposon-associated CRISPR-Cas system vs. DNA cargo size

ANOVA: Single Factor (α = 0.05)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	200.617284	2	100.308642	1.625	**0.27291572**	5.14325285
Within Groups	370.37037	6	61.7283951			
Total	570.987654	8				

Table S3.3: Statistical analysis data for volumetric CFU based on transposon-associated CRISPR-Cas system vs. DNA cargo size

ANOVA: Single Factor (α = 0.05)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1069500247	2	534750123	1.90985892	**0.22811651**	5.14325285
Within Groups	1679967407	6	279994568			
Total	2749467654	8				

Table S3.4: Statistical analysis data for pDonor eviction efficiency based on transposon-associated CRISPR-Cas system vs. DNA cargo size

ANOVA: Single Factor (α = 0.05)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	8101.85185	2	4050.92593	20.1923077	**0.00216437**	5.14325285
Within Groups	1203.7037	6	200.617284			
Total	9305.55556	8				

Table S3.5: Statistical analysis data for comparing the efficiency of the integrating methods with increase in the DNA cargo size

ANOVA: Two-Factor With Replication (α = 0.05)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Method of Integration	4218.75	1	4218.75	36.45	**0.00031022**	5.31765507
DNA size	4866.89815	1	4866.89815	42.05	0.00019119	5.31765507
Interaction of method vs. DNA size	2552.08333	1	2552.08333	22.05	**0.00154997**	5.31765507
Within	925.925926	8	115.740741			
Total	12563.6574	11				

Table S4.1: Oligomers used in Chapter 4

Primers/ oligo name	**Primer/oligo sequence (5' → 3')**
P001	AAGGAGGAATAGAAAATGGACTACAATCTGGCACTCG
P002	TTTCGTACCTCCTTTGCTAGCTCAGGCAACGACCTCGGC
P003	CTAGCAAAGGAGGTACGAAATATGACAGACTTAATCCAACGC
P004	ATGTAAATTTCCTCCTTTGGTACCTTAACGCAGAATCTTCTTCTCAG
P005	CCAAAGGAGGAAATTTACATGTTAGACAATGTTTTAAGAATTGCCAC
P006	TAATAATCCTCCTTTGCATGCTCATGCCGGGGCGTCTCC
P007	CATGCAAAGGAGGATTATTATATGAGTATCCTTGTCACCC
P008	CTGCGGATCCTTATTGTAATGCCCGTAAAAGC
P009	ATTACAATAAGGATCCGCAGCCCGCCTAATG
P010	TCCATTTTCTATTCCTCCTTTAATTGTTATCCGCTC
P011	GGATCCGCAGCCCGCCTAATG
P012	TCAGGCAACGACCTCGGCG
P013	TTAACGCAGAATCTTCTTCTCAGCC
P014	CCCGGCATGAGGATCCGCAGCCCGCCTAATGAGC
P015	CTGCGGATCCTCATGCCGGGGCGTCTCCGTTATAG
P016	CCAAAGGAGGAAATTTACATGATGAGAACGATTAAAGTAGG
P017	CTGCGGATCCTCATTTTCCATCCTCGTCAAGC
P018	TGGAAAATGAGGATCCGCAGCCCGCCTAATGAGC
P019	GTCCATTTTCTATTCCTCCTTTAATTGTTATCCGC
P020	TACCAAAGGAGGAAATTTACATGAGTATCCTTGTCACCCG
P021	GTCCATTTTCTATTCCTCCTTTAATTGTTATCCG

Table S4.2: Strains and plasmids used in Supplementary Figures

Name	Description or Relevant Genotype	Source
Host strains		
ST002	CPC-Sbm *endA*::(P_{BAD}::*γ*:*β*:*α*) *yjcS*::(P_{tetA}::spc.-gRNA.P279T-*cas9*)	Chapter 3
ST-AB	CPC-Sbm *endA*::(P_{BAD}::*γ*:*β*:*α*) *yjcS*::(P_{tetA}::spc.-gRNA.P279T-*cas9*) Δ*iclR* Δ*sdhA* *intF*::(P_{trc}::*hemAB*)	This study
ST-ABC	CPC-Sbm *endA*::(P_{BAD}::*γ*:*β*:*α*) *yjcS*::(P_{tetA}::spc.-gRNA.P279T-*cas9*) Δ*iclR* Δ*sdhA* *intF*::(P_{trc}::*hemABC*)	This study
BL21(DE3)	F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dcm* (DE3)	Lab stock
BLS	BL21(DE3) Δ*iclR*	This study
BLD	BL21(DE3) Δ*iclR* Δ*sdhA*	This study
BL-AB	BL21(DE3)/ pK-hemAB	This study
BLS-AB	BLS/ pK-hemAB	This study
BLD-AB	BLD/ pK-hemAB	This study
BLD-A	BLD/ pK-hemA	This study
BLD-ABC	BLD/ pK-hemABC	This study
BLD-ABE	BLD/ pK-hemAB-E	This study
Plasmids		
pK-hemA	p15A ori, Km^R, P_{trc}::*hemA*	Chapter 4
pK-hemAB	p15A ori, Km^R, P_{trc}::*hemAB*	Chapter 4
pK-hemABC	p15A ori, Km^R, P_{trc}::*hemABC*	Chapter 4
pK-hemAB-E	p15A ori, Km^R, P_{trc}:: *hemAB* -$P_{gracmax}$::*hemE*	Chapter 5

Table S5.1: Oligomers used in Chapter 5

Primers/ oligo name	**Primer/oligo sequence (5' → 3')**
P001	AAGGAGGAATAGAAAATGGACTACAATCTGGCACTCG
P002	TTTCGTACCTCCTTTGCTAGCTCAGGCAACGACCTCGGC
P003	CTAGCAAAGGAGGTACGAAATATGACAGACTTAATCCAACGC
P004	ATGTAAATTTCCTCCTTTGGTACCTTAACGCAGAATCTTCTTCTCAG
P005	CCAAAGGAGGAAATTTACATGTTAGACAATGTTTTAAGAATTGCCAC
P006	TAATAATCCTCCTTTGCATGCTCATGCCGGGGCGTCTCC
P007	CATGCAAAGGAGGATTATTATATGAGTATCCTTGTCACCC
P008	CTGCGGATCCTTATTGTAATGCCCGTAAAAGC
P009	ATTACAATAAGGATCCGCAGCCCGCCTAATG
P010	TCCATTTTCTATTCCTCCTTTAATTGTTATCCGCTC
P011	GTATCACCGCTAAGGATCCGCAGCCCGCCTAATGAGC
P012	TTTCGGATGTTATTGTAATGCCCGTAAAAGCGCATCG
P013	GGAATCAATAATGACCGAACTTAAAAACGATCG
P014	CATCGACTGTCTGAACAGTATCACCGCTAAGGATCCG
P015	CGGGCATTACAATAACATCCGAAAAGAATGATGGATC
P016	TTAAGTTCGGTCATTATTGATTCCTCCTTTAATTGG
P017	TACCAAAGGAGGAAATTTACATGAGTATCCTTGTCACCCG
P018	GTCCATTTTCTATTCCTCCTTTAATTGTTATCCG
P019	TCTTTTCGGATGTTAACGCAGAATCTTCTTCTCAG
P020	TCTGCGTTAACATCCGAAAAGAATGATGG
P021	GGATACTCATTATTGATTCCTCCTTTAATTGGG
P022	GGAGGAATCAATAATGAGTATCCTTGTCACCCG
P023	AAATTTCCTCCTTTGGTACCTTATTGTAATGCCCGTAAAAGCGCATCG
P024	GGTACCAAAGGAGGAAATTTACATGACCGAACTTAAAAACGATCG
P025	TTGAATTTTCCTCCTTTTCATGATTAGCGGTGATACTGTTCAG
P026	TGAAAAGGAGGAAAATTCAATGAGTGACGGCAAAAAACATG
P027	CTGCGGATCCTTAGCTGAATAAATAGGTAAGCG
P028	TTATTCAGCTAAGGATCCGCAGCCCGCCTAATGAG
P029	GTCCATTTTCTATTCCTCCTTTAATTGTTATCCGC
P030	TGAAAAGGAGGAAAATTCAATGCTAAAAGTGTTATTGCTCTTTGTG

P031	CTGCGGATCCCTACTGTGGCGGGTTATTCTGC
P032	GCCACAGTAGGGATCCGCAGCCCGCCTAATGAGC
P033	CAATTAAAGGAGGAATAGAAAATGGAC
P034	TTTCTATTCCTCCTTTAATTGTTATCCG

Table S6.1: Oligomers used in Chapter 6

Primers/ oligo name	**Primer/oligo sequence (5' → 3')**
P1	TTATATGTATAAGGATCCGCAGCCCGCCTAATGAGC
P2	TATTTCGTACCTCCTTTGCTAGCTCAGG
P3	GCTAGCAAAGGAGGTACGAAATATGACAGAC
P4	TGAATTTTCCTCCTTTTCATGATTAGCGGTGATACTGTTCAGACAG
P5	TCATGAAAAGGAGGAAAATTCAGTGAGTAGAAAGAAAATGGGGCTTC
P6	CTATTTAACCTCCTTTACTAGTTTAACGTCCTAATTTTTTTAATACGACTGTTGC
P7	ACTAGTAAAGGAGGTTAAATAGATGAGTGAACAGCAAATGACG
P8	GCTGCGGATCCTTATACATATAAGAATTGAGGAAGTGTATCAAGCG
P9	TTAGGACGTTAAGGATCCGCAGCCCGCCTAATGAGC
P10	TATTTCGTACCTCCTTTGCTAGCTCAGGC
P11	GCTGCGGATCCTTAACGTCCTAATTTTTTTAATACGACTGTTGC
P12	CAATTAAAGGAGGAATAGAAAATGGAC
P13	TGAATTTTCCTCCTTTTCATGATTAGCGGTGATACTGTTCAGACAG
P14	TCCATTTTCTATTCCTCCTTTAATTGTTATCCGCTC
P15	GCTGCGGATCCTTAACGTCCTAATTTTTTTAATACGACTGTTGC
P16	AAGGAGGAATAGAAAATGGACTACAATCTGGCACTCG
P17	ATGTAAATTTCCTCCTTTGGTACCTTAACGCAGAATCTTCTTCTCAG
P18	GGTACCAAAGGAGGAAATTTACATGACCGAACTTAAAAACGATCG
P19	TCTTTTCGGATGTTAGCGGTGATACTGTTCAG
P20	TATCACCGCTAACATCCGAAAAGAATGATGGATC
P21	CTTTCTACTCACTATTGATTCCTCCTTTAATTGGG
P22	GAGGAATCAATAGTGAGTAGAAAGAAAATGGGGCTTC
P23	GTAGTCCATTTTCTATTCCTCCTTTAATTGTTATCCGCTC
P24	ACTAGTAAAGGAGGTTAAATAGATGAGTCAAGCAGCCGAAACATTAGATGG
P25	GCTGCGGATCCTTAAGAAATCGCAAAGAATTGATCG
P26	CGATTTCTTAAGGATCCGCAGCCCGCCTAATGAG
P27	AGGAAGAATAAGTGAGTAGAAAGAAAATGGGGCTTC
P28	TTAGTGATGATGATGATGGTGACGTCCTAATTTTTTTAATACGACTGTTGC
P29	GTGCACCATCATCATCATCACAGTAGAAAGAAAATGGGGCTTCTCG
P30	CCAAAACAGCCTTAACGTCCTAATTTTTTTAATACGACTGTTGC

P31	TCACCATCATCATCATCACTAAGGCTGTTTTGGCGGATGAGAG
P32	TTTCTACTCACTTATTCTTCCTCCTTTTTAAAGTTAATGCTAGC
P33	TAGGACGTTAAGGCTGTTTTGGCGGATGAGAG
P34	TGTGATGATGATGATGGTGCACTTATTCTTCCTCCTTTTTAAAGTTAATGCTAGC
P35	AGGAAGAATAAATGAGTGAACAGCAAATGACG
P36	CTTAGTGATGATGATGATGGTGTACATATAAGAATTGAGGAAGTGTATCAAGCG
P37	AATGCACCATCATCATCATCACAGTGAACAGCAAATGACGAATGAAGCG
P38	CCAAAACAGCCTTATACATATAAGAATTGAGGAAGTGTATCAAGCG
P39	AGGAAGAATAAATGAGTCAAGCAGCCGAAAC
P40	CTTAGTGATGATGATGATGGTGAGAAATCGCAAAGAATTGATCG
P41	AATGCACCATCATCATCATCACAGTCAAGCAGCCGAAACATTAGATGG
P42	CCAAAACAGCCTTAAGAAATCGCAAAGAATTGATCGAATTCG
P43	CACCATCATCATCATCACTAAGGCTGTTTTGGCGGATGAGAG
P44	TGTTCACTCATTTATTCTTCCTCCTTTTTAAAGTTAATGCTAGC
P45	TATATGTATAAGGCTGTTTTGGCGGATGAGAG
P46	GTGATGATGATGATGGTGCATTTATTCTTCCTCCTTTTTAAAGTTAATGCTAGC
P47	GCTTGACTCATTTATTCTTCCTCCTTTTTAAAGTTAATGCTAGC
P48	CGATTTCTTAAGGCTGTTTTGGCGGATGAGAG

Table S7.1: Oligomers used in Chapter 7

Primers/ oligo name	**Primer/oligo sequence (5' → 3')**
P1	AAGGAGGAATAGAAAATGGACTACAATCTGGCACTCG
P2	ATGTAAATTTCCTCCTTTGGTACCTTAACGCAGAATCTTCTTCTCAG
P3	TAATTAAGTAAGGATCCGCAGCCCGCCTAATGAG
P4	GTAGTCCATTTTCTATTCCTCCTTTAATTGTTATCCGCTC
P5	TACCAAAGGAGGAAATTTACATGAGCTATACCCTGCCATCC
P6	ATAATAATCCTCCTTTGCATGCTTATTTTTTCGCCGCAAAACGTG
P7	GCATGCAAAGGAGGATTATTATATGTCATTCGAATTACCTGC
P8	CTATTTAACCTCCTTTACTAGTTTATGCAGCGAGATTTTTCG
P9	ACTAGTAAAGGAGGTTAAATAGATGAAACGTTTTAGTCTGGC
P10	GCTGCGGATCCTTACTTAATTACACCACAGG
P11	CGAAAAAATAAGGATCCGCAGCCCGCCTAATGAG
P12	GCTGCGGATCCTTATTTTTTCGCCGCAAAACGTG
P13	TCGCTGCATAAGGATCCGCAGCCCGCCTAATGAG
P14	TACCAAAGGAGGAAATTTACATGTCATTCGAATTACCTGC
P15	GCTGCGGATCCTTATGCAGCGAGATTTTTCG
P16	TACCAAAGGAGGAAATTTACATGAAACGTTTTAGTCTGGC
P17	ACCTGCTGTAAGGATCCGCAGCCCGCCTAATGAG
P18	TACCAAAGGAGGAAATTTACATGAGC
P19	CTTTTCGGATGTTACTTAATTACACCACAGGCATAGCG
P20	TAATTAAGTAACATCCGAAAAGAATGATGGATC
P21	TGTTGCGACATTATTGATTCCTCCTTTAATTGGG
P22	AGGAATCAATAATGTCGCAACATAACGAAAAGAACC
P23	TGAATTTTCCTCCTTTTCATGATCAGGCAGGAATTTTGTCAATCTTAGG
P24	TCATGAAAAGGAGGAAAATTCAATGAGCACGTCAGACGATATCC
P25	GCTGCGGATCCTTACAGCAGGTCGAAACGGTC
P26	GGAGGAATAGAAAATGGACTACAATCTGG
P27	TTCCTGCCTGAGGATCCGCAGCCCGCCTAATGAG
P28	GCTGCGGATCCTCAGGCAGGAATTTTGTCAATCTTAGG
P29	GACGTGCTCATTATTGATTCCTCCTTTAATTGGG
P30	AGGAATCAATAATGAGCACGTCAGACGATATCC

P31	CTTTTCGGATGTTATTTTTTCGCCGCAAAACGTG
P32	CGAAAAAATAACATCCGAAAAGAATGATGGATC

Table S7.2: Statistical analysis data for the UP-III/UP-I ratio for 0.5 g/L ascorbic acid supplementation vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*0.5 g/L Ascorbic acid*
Mean	0.62764495	1.44557979
Variance	0.00868759	0.00282797
Observations	2	2
Pearson Correlation	-1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-7.9019515	
P(T≤t) one-tail	**0.04006944**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.08013888	
t Critical two-tail	12.7062047	

Table S7.3: Statistical analysis data for the UP-III/UP-I ratio for 1 g/L ascorbic acid supplementation vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*1 g/L Ascorbic acid*
Mean	0.62764495	2.57398428
Variance	0.00868759	0.00896544
Observations	2	2
Pearson Correlation	-1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-14.649488	
P(T≤t) one-tail	**0.02169474**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.04338948	
t Critical two-tail	12.7062047	

Table S7.4: Statistical analysis data for the UP-III/UP-I ratio for 2.0 g/L ascorbic acid supplementation vs the 0.5 g/L ascorbic acid supplementation

t-Test: Paired Two Sample for Means (α = 0.05)

	0.5 g/L Ascorbic acid	*2.0 g/L Ascorbic acid*
Mean	1.44557979	1.5469349
Variance	0.00282797	0.04561101
Observations	2	2
Pearson Correlation	-1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-0.5373567	
P(T≤t) one-tail	**0.34304634**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.68609268	
t Critical two-tail	12.7062047	

Table S7.5: Statistical analysis data for the UP-III/UP-I ratio for SOD2 vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*SOD2*
Mean	0.62764495	1.94063513
Variance	0.00868759	0.01188752
Observations	2	2
Pearson Correlation	1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-117.35372	
P(T≤t) one-tail	**0.00271233**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.00542466	
t Critical two-tail	12.7062047	

Table S7.6: Statistical analysis data for the UP-III/UP-I ratio for SOD3 vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*SOD3*
Mean	0.62764495	1.75227929
Variance	0.00868759	9.6012E-05
Observations	2	2
Pearson Correlation	-1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-15.440617	
P(T≤t) one-tail	**0.02058635**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.0411727	
t Critical two-tail	12.7062047	

Table S7.7: Statistical analysis data for the UP-III/UP-I ratio for SOD2 + 1 g/L ascorbic acid supplementation vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*SOD2 + 1 g/L Ascorbic acid*
Mean	0.62764495	1.4131059
Variance	0.00868759	0.00014286
Observations	2	2
Pearson Correlation	-1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-10.563096	
P(T≤t) one-tail	**0.0300446**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.06008921	
t Critical two-tail	12.7062047	

Table S7.8: Statistical analysis data for the UP-III/UP-I ratio for KAT1 vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*KAT1*
Mean	0.62764495	6.20477723
Variance	0.00868759	0.03422162
Observations	2	2
Pearson Correlation	1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-85.933143	
P(T≤t) one-tail	**0.00370399**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.00740798	
t Critical two-tail	12.7062047	

Table S7.9: Statistical analysis data for the UP-III/UP-I ratio for KAT2 vs the control strain

t-Test: Paired Two Sample for Means (α = 0.05)

	Control	*KAT2*
Mean	0.62764495	1.28336217
Variance	0.00868759	0.00098766
Observations	2	2
Pearson Correlation	-1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-7.4403655	
P(T≤t) one-tail	**0.04252664**	
t Critical one-tail	6.31375151	
P(T≤t) two-tail	0.08505328	
t Critical two-tail	12.7062047	

Supplementary Figures

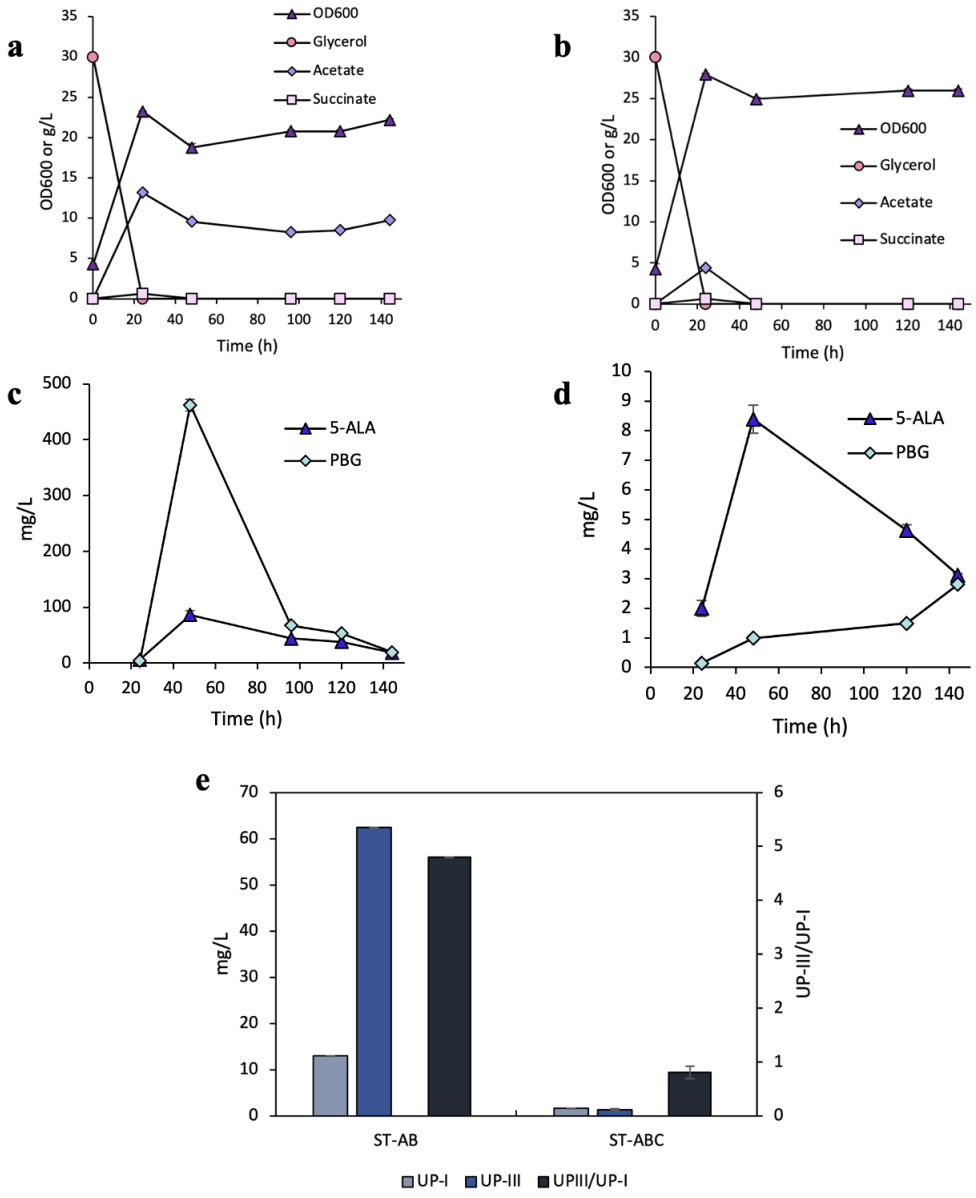


Figure S4.1: Aerobic bioreactor cultivation of ST-AB and ST-ABC strains for UP biosynthesis. Time profiles of (**a** and **b**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**c** and **d**) 5-ALA and PBG biosynthesis, and (**e**) UP biosynthesis at 144 h. All values are reported as means±SD (n=2).

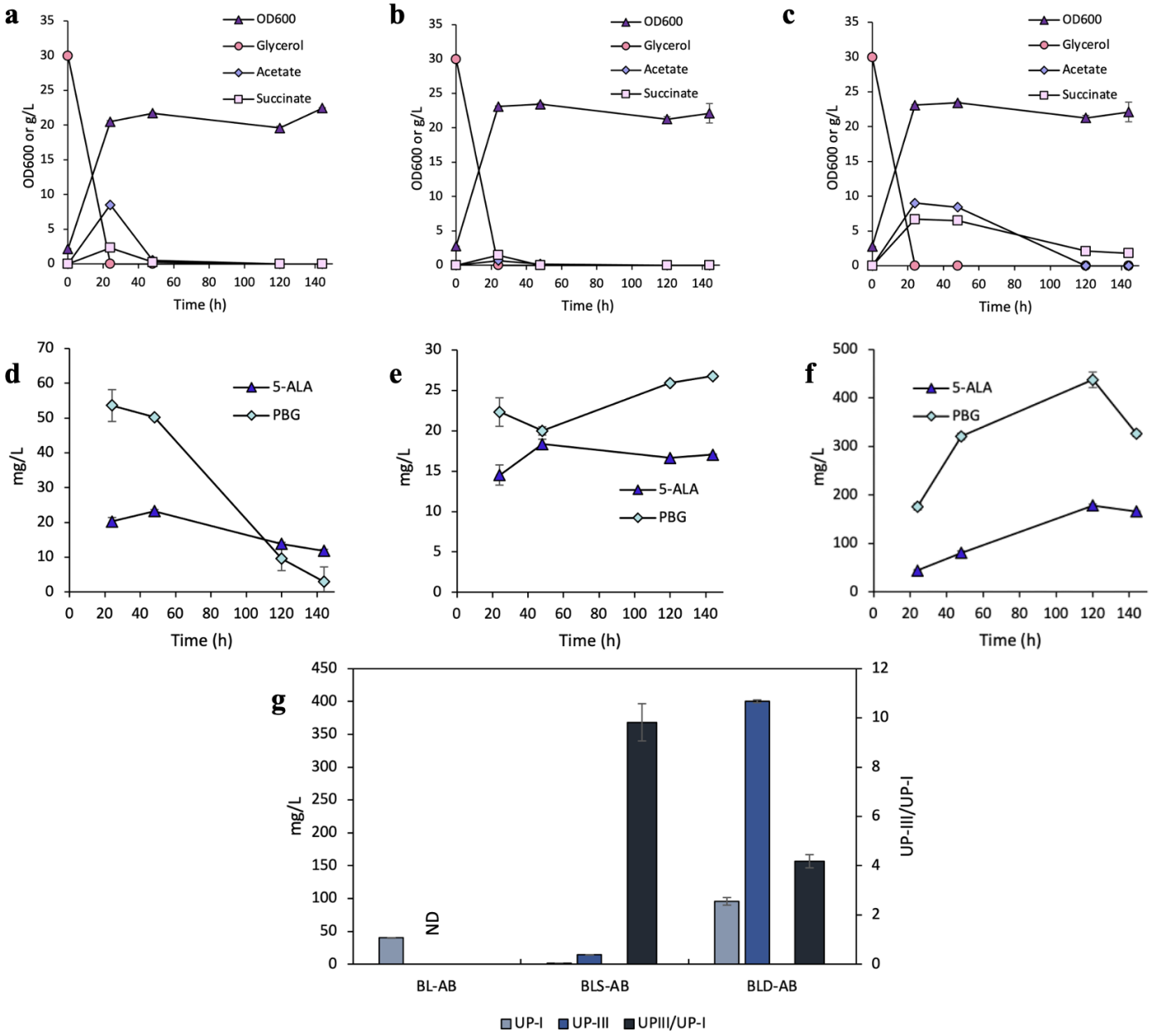


Figure S4.2: Aerobic bioreactor cultivation of BL21(DE3)-derived strains for UP biosynthesis. Time profiles of (**a-c**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**d-f**) 5-ALA and PBG biosynthesis, and (**g**) UP biosynthesis at 144 h. All values are reported as means±SD (n=2).

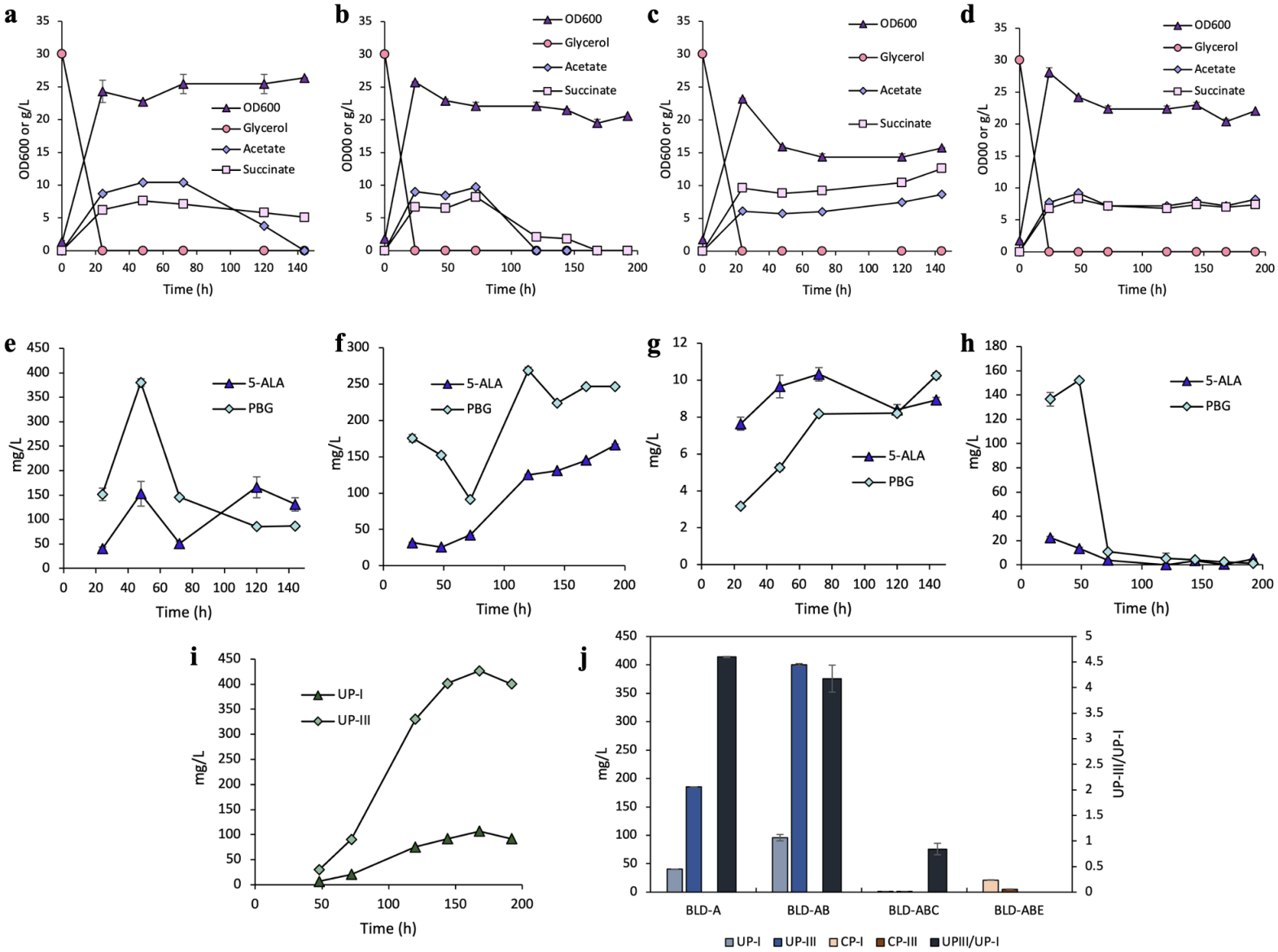


Figure S4.3: Aerobic bioreactor cultivation of BLD-derived strains for porphyrin biosynthesis. Time profiles of (**a-d**) cell density (OD_{600}), glycerol consumption, and acetate/succinate formation (**e-h**) 5-ALA and PBG biosynthesis, (**i**) UP biosynthesis time profile for BLD-AB, and (**j**) porphyrin biosynthesis at 144 h. All values are reported as means±SD (n=2).