

SYNTHESIS OF 3-HYDROXYPROPIONIC ACID FROM ACETYLENECARBOXYLIC ACID

By

Amaya Nethmini Sirinimal Mathes Hewage

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry—Doctor of Philosophy

2022

ABSTRACT

SYNTHESIS OF 3-HYDROXYPROPIONIC ACID FROM ACETYLENECARBOXYLIC ACID

By

Amaya Nethmini Sirinimal Mathes Hewage

Throughout the past several decades, the petroleum industry has remained a dominant player in the production of fuels, chemicals, and consumer products. However, rising global temperatures and declining fossil fuel reserves demand the need for sustainable processes that can compete with the existing crude oil-based economy. Employing microbes such as *E. coli* for the chemical production using renewable starch-derived feedstocks such as glucose has been identified as a preferred alternative. Although lignocellulosic feedstocks have been explored as an alternative to renewable sugars, these processes are yet to be successfully implemented in an industrial setting. Despite more recent research efforts to incorporate C1 feedstocks such as methane, methanol and carbon dioxide in microbial catalysis, development of efficient bacterial metabolic pathways using these feedstocks have proven to be challenging. Herein, acetylenecarboxylic acid (ACA) is proposed as a unique feedstock for microbial catalysis. ACA can be derived from dehydrodimerization of methane to acetylene and subsequent carboxylation of acetylene. Cg10062 was previously identified as an enzyme capable of hydrating ACA to form a mixture of malonate semialdehyde (MSA) and acetaldehyde. In this study, novel variant Cg10062(E114N)) that forms exclusively MSA from ACA was discovered using rational mutagenesis. Cg10062(E114N) was coupled with NADPH-dependent dehydrogenase YdfG to develop a unique biocatalytic route to building block chemical 3-hydroxypropionic acid (3-HP). *In vitro* synthesis of 3-HP from ACA was demonstrated using catalytic amounts of NADP(H) where cofactor regeneration was achieved using NADP⁺-dependent phosphite dehydrogenase PTDH.

This dissertation is dedicated in memory and honor of my beloved parents.
Thank you for your unconditional love and support. I hope I made you proud.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my research advisor Dr. Karen Draths for what has been a most enriching experience. I will always remain grateful to you for creating a safe learning environment as I started research in your lab. As a first year graduate student, I was incredibly thankful for the opportunity to learn directly from a seasoned scientist and mentor. Thank you for believing in me and giving me countless opportunities to develop my skills as a research mentor. I will never forget your kind understanding and patience during difficult times over the past five years and for going above and beyond your responsibilities as a PhD supervisor. I can only hope to pass this kindness onto others. A special thank you to my committee, Dr. Babak Borhan, Dr. James Jackson and Dr. Aaron Odom for their continued guidance throughout the program.

My sincere thanks to Dr. Kelly Miller, Dr. Hadi Nayebi and Kiyoto Tanemura for their guidance during my first year of graduate school. Thank you for sharing valuable research techniques and knowledge, and yet always maintaining a safe distance to allow me to grow as an independent researcher. A very special thank you to my colleagues and friends, Megan Gruenberg, Katie Kwiatkowski, Katelyn Silva, Bismarck Amaniampong, Yasheen Jadidi, Austin Jones, Dave Sreedhar, Daniel Chi and Esther Lee for all your support. You are all friends turned family and I wouldn't have made it through this difficult journey without you all.

My heartfelt gratitude goes to my father, whose undying efforts made it all possible. You always gave us the best of everything in life. I'm thankful to my late mother, who taught me to be strong in the face of adversity. I never did get the chance to thank you for that. My love goes to my family, Hansamali, Yasas, Suchira, Malaka, Yuween and Yuara for their unparalleled support. Thank you to all my friends for helping me persevere through difficult times, the list is simply too long but know that I truly appreciate you all.

TABLE OF CONTENTS

LIST OF TABLES.....	vii
LIST OF FIGURES	ix
KEY TO ABBREVIATIONS	xv
CHAPTER ONE: Cg10062 CATALYSIS FORGES A LINK BETWEEN ACETYLENECARBOXYLIC ACID AND BACTERIAL METABOLISM	1
1.1. Introduction	1
1.2. Synthesis of ACA from CH ₄ and CO ₂	3
1.3. ACA and bacterial metabolism	4
1.4. Hydration of ACA by Cg10062	8
1.5. Rational mutagenesis of the Cg10062 active site	14
1.6. Kinetic characterization of Cg10062 and variants	17
1.7. ¹ H NMR characterization of Cg10062 and variants.....	29
1.8. Structural analysis of Cg10062 and variants	39
1.9. Proposed mechanism for hydration of ACA by Cg10062.....	46
1.10. Assessing the growth of <i>E. coli</i> expressing Cg10062 or variant on ACA	50
1.11. Growth of <i>E. coli</i> on aldehyde indicator plates.....	54
1.12. Conclusion	60
REFERENCES	62
CHAPTER TWO: SYNTHESIS OF 3-HYDROXYPROPIONIC ACID.....	69
2.1. Introduction	69
2.2. Synthetic routes to 3-HP	72
2.2.1. 3-HP production from glycerol.....	73
2.2.2. 3-HP production from glucose	75
2.2.3. Other biocatalytic routes to 3-HP.....	77
2.3. Synthesis of 3-HP from ACA	78
2.3.1. Characterization of RutE and YdfG for MSA reduction to 3-HP	80
2.3.2. Steady-state kinetics and product confirmation of YdfG	88
2.4. Cofactor regeneration for 3-HP synthesis.....	93
2.4.1. Steady-state kinetics of PTDH.....	94
2.5. Buffer dependence of Cg10062(E114N), YdfG and PTDH.....	97
2.6. <i>In vitro</i> synthesis of 3-HP from ACA with cofactor recycling	101
2.7. Cg10062(E114N) inhibition by NADP(H).....	122
2.8. <i>In vivo</i> synthesis of 3-HP	129
2.9. Conclusion	134
REFERENCES	137
CHAPTER THREE: 3-HP BIOSENSOR FOR EVOLUTION OF Cg10062.....	146
3.1. Introduction	146
3.1.1. 3-HP degradation in <i>Pseudomonas denitrificans</i>	148
3.1.2. Genetically-encoded 3-HP biosensor in <i>Pseudomonas denitrificans</i>	149
3.2. Development of a 3-HP biosensor for evolution of Cg10062 in <i>E. coli</i> MG1655.....	152
3.3. Characterization of 3-HP biosensor in <i>E. coli</i> MG1655.....	157
3.4. ¹ H NMR characterization of <i>in vivo</i> synthesis of 3-HP from ACA.....	164

3.5. Conclusion	170
REFERENCES	173
CHAPTER FOUR: EXPERIMENTAL	177
4.1. Materials and equipment	177
4.2. Culture media and stock solutions	178
4.3. Preparation and transformation of electrocompetent <i>E. coli</i>	179
4.4. Isolation of plasmid DNA	180
4.5. Isolation of genomic DNA	181
4.6. Restriction digestion of DNA.....	181
4.7. DNA sequencing	181
4.8. PCR amplification.....	182
4.9. Protein expression	183
4.10. Protein purification	183
4.11. Protein quantification.....	184
4.12. Polyacrylamide gel electrophoresis (SDS-PAGE).....	186
4.13. Distillation of acetylenecarboxylic acid (ACA).....	187
4.14. Bacterial strains and plasmids.....	188
4.15. Chapter One: Characterization of Cg10062 and variants	190
4.15.1. Genes and plasmids.....	190
4.15.2. Expression and purification of TEV protease	195
4.15.3. His ₆ -tag removal by TEV protease.....	195
4.15.4. Q5 site-directed mutagenesis	195
4.15.5. Kinetic characterization of Cg10062 and variants	197
4.15.6. Specific activity of MSAD.....	199
4.15.7. ¹ H NMR characterization of Cg10062 and variants	200
4.15.8. Protein crystallization and structure determination.....	201
4.15.9. Growth of <i>E. coli</i> on ACA.....	201
4.15.10. Growth of <i>E. coli</i> on ACA.....	202
4.15.11. Growth of <i>E. coli</i> expressing Cg10062 on aldehyde indicator plates.....	203
4.16. Chapter Two: Synthesis of 3-Hydroxypropionic Acid	203
4.16.1. Genes and plasmids.....	203
4.16.2. Purification and characterization of YdfG and RutE	210
4.16.3. YdfG coupled enzyme assay.....	211
4.16.4. ¹ H NMR characterization of YdfG	212
4.16.5. Kinetic characterization of PTDH.....	212
4.16.6. Buffer dependence of Cg10062(E114N), YdfG and PTDH	213
4.16.7. <i>In vitro</i> synthesis of 3-HP with cofactor regeneration	215
4.16.8. <i>In vivo</i> synthesis of 3-HP in <i>E. coli</i>	216
4.16.9. HPLC quantification of ACA and 3-HP.....	216
4.16.10. ¹ H NMR analysis of <i>in vitro</i> and <i>in vivo</i> 3-HP synthesis.....	217
4.16.11. Cg10062(E114N) inhibition by NADP(H)	218
4.17. Chapter Three: 3-HP Biosensor for evolution of Cg10062(E114N).....	222
4.17.1. Genes and plasmids.....	222
4.17.2. Preparation of <i>E. coli</i> 3HPC, 3HPC-WT and 3HPC-E114N	228
4.17.3. Determination of RFP fluorescence as a function of IPTG concentration.....	228
4.17.4. Determination of RFP fluorescence as a function of 3-HP concentration.....	229
4.17.5. Determination of RFP fluorescence as a function of ACA concentration.....	230
4.17.6. Determination of RFP fluorescence with increased 3-HP and ACA concentrations	230
4.17.7. ¹ H NMR characterization of 3-HP biosensor	230
REFERENCES	232

LIST OF TABLES

Table 1.1. Steady state kinetic parameters of Cg10062 and variants	9
Table 1.2. Product distribution of Cg10062 variants using the coupled enzyme assay	22
Table 1.3. Kinetic parameters of Cg10062 and variants.	28
Table 1.4. Genes expressed by <i>E. coli</i> RB791 <i>serA</i> transformants.	50
Table 2.1. Biochemical characterization of YdfG and RutE for reduction of MSA	87
Table 2.2. Specific activity of PTDH.	96
Table 2.3. Michaelis Menten kinetic parameters of enzymes used for 3-HP synthesis.	96
Table 2.4. Synthesis of 3-HP from 100 mM ACA with cofactor regeneration.	110
Table 4.1. Components of a routine restriction digest.....	181
Table 4.2. Components of a routine PCR reaction	182
Table 4.3. Standard temperature cycling parameters for PCR	183
Table 4.4. Molecular weights and molar extinction coefficients of enzymes	185
Table 4.5. Strains and plasmids used in this study	188
Table 4.6. Components of a routine Q5 site-directed mutagenesis PCR	196
Table 4.7. Primers used to generate Cg10062 variants.....	196
Table 4.8. KLD treatment of PCR product in Q5 site-directed mutagenesis	197
Table 4.9. Components used in the Cg10062 coupled enzyme assay	198
Table 4.10. Components used in the MSAD coupled enzyme assay.....	200
Table 4.11. Components used in the RutE/YdfG coupled enzyme assay	211
Table 4.12. Components used in the YdfG coupled enzyme assay	212
Table 4.13. Components used in the PTDH enzyme assay.....	213
Table 4.14. Components used in pH dependence assays.....	214
Table 4.15. Conversion of 100 mM ACA to 3-HP with cofactor regeneration.....	215

Table 4.16. Conversion of 500 mM ACA with cofactor regeneration.....	215
Table 4.17. PicoProbe™ components used for NADH standard curve preparation.	220
Table 4.18. Components used in the Cg10062 coupled enzyme assay	221
Table 4.19. Primers used in chapter three of this study.....	227

LIST OF FIGURES

Figure 1.1. Anodic oxidation of propargyl alcohol.....	4
Figure 1.2. General conditions for the synthesis of ACA and ADCA from CH ₄ and CO ₂	4
Figure 1.3. Base-dependent carboxylation of acetylene followed by hydrogenation.....	4
Figure 1.4. The 1,3-dichloropropene catabolic pathway	6
Figure 1.5. An overlay of the active sites of cis-CaaD (PDB 3MF8) and Cg10062 (PDB 3N4G). ..	7
Figure 1.6. Varying catalytic activity of Cg10062 and cis-CaaD with ACA.	8
Figure 1.7. Mechanism 1 proposed by Huddleston et al.....	11
Figure 1.8. Mechanism 2 proposed by Huddleston et al.....	12
Figure 1.9. Mechanism 3 proposed by Huddleston et al.....	13
Figure 1.10. Codon-optimized nucleotide sequence of <i>cg10062</i>	15
Figure 1.11. Purification of Cg10062 from a 1L culture of BL21(DE3)/pAS1.046 via affinity chromatography..	16
Figure 1.12. SDS-PAGE of Cg10062 (19 kDa) and variants.....	17
Figure 1.13. Codon-optimized nucleotide sequence of <i>msad</i>	18
Figure 1.14. Purification of MSAD from a 1 L culture of BL21(DE3)/pAS2.031 via affinity chromatography.	19
Figure 1.15. The coupled enzyme assay used to determine the ratio of products formed by Cg10062 and variants.	20
Figure 1.16. Cg10062 activity was determined using the Cg10062 coupled enzyme assay	20
Figure 1.17. Comparison of Cg10062 activity before and after removal of His ₆ -tag.....	21
Figure 1.18. The relative ratios of MSA and acetaldehyde produced by Cg10062 and variants.....	23
Figure 1.19. Michaelis Menten kinetics of Cg10062	25
Figure 1.20. Michaelis Menten kinetics of Cg10062(E114Q).....	25
Figure 1.21. Michaelis Menten kinetics of Cg10062(E114D).....	26
Figure 1.22. Michaelis Menten kinetics of Cg10062(E114N).....	26

Figure 1.23. Michaelis Menten kinetics of Cg10062(E114D-Y103F).....	27
Figure 1.24. Michaelis Menten kinetics of Cg10062(Y103F).....	27
Figure 1.25. ¹ H NMR of reaction of Cg10062 (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 0 h.....	31
Figure 1.26. ¹ H NMR of reaction of Cg10062 (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 1 h.....	32
Figure 1.27. ¹ H NMR of reaction of Cg10062(E114Q) (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 0 h.....	33
Figure 1.28. ¹ H NMR of reaction of Cg10062(E114Q) (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 1 h.....	34
Figure 1.29. ¹ H NMR of reaction of Cg10062(E114D) (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 0 h.....	35
Figure 1.30. ¹ H NMR of reaction of Cg10062(E114D) (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 1 h.....	36
Figure 1.31. ¹ H NMR of reaction of Cg10062(E114N) (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 0 h.....	37
Figure 1.32. ¹ H NMR of reaction of Cg10062(E114N) (0.27 mg mL ⁻¹) incubated with ACA (111 mM) at t = 1 h.....	38
Figure 1.33. Structure of apo-Cg10062 (PBD 7MS0)	40
Figure 1.34. Overlay of wild type Cg10062 structures	40
Figure 1.35. The active site of wild type Cg10062 and covalently-bound intermediates captured in Cg10062 variants.	41
Figure 1.36. Asymmetric unit of Cg10062(E114D) with twelve chains in four trimers.	43
Figure 1.37. 3-(<i>N</i> -prolyl)-acrylate (enamine) intermediate in chains A, B, H, I and K of Cg10062(E114D).	44
Figure 1.38. Comparison of active sites of Cg10062 variants.....	45
Figure 1.39. Active site intermediates in Cg10062(E114N)	46
Figure 1.40. Proposed catalytic cycle of Cg10062 hydration and hydration/decarboxylation of ACA.....	49
Figure 1.41. Growth of <i>E. coli</i> on M9/ACA/glu plates.....	52
Figure 1.42. A comparison of <i>E. coli</i> growth on M9 plates after 24 h	53

Figure 1.43. “Film”-like growth of <i>E. coli</i> expressing Cg10062 and variants on M9/ACA.....	53
Figure 1.44. <i>p</i> -Rosaniline reacts with bisulfite to produce "decolorized" Schiff reagent	55
Figure 1.45. Growth of <i>E. coli</i> RB791 <i>serA</i> on aldehyde indicator plates	56
Figure 1.46. Color changes observed on M9/ACA/glu/ros after 48 h of growth.....	57
Figure 1.47. Dark red rings around RB791 <i>serA</i> /pAS1.163.	57
Figure 1.48. Growth of <i>E. coli</i> RB791 <i>serA</i> on aldehyde indicator plates	59
Figure 1.49. Growth of <i>E. coli</i> RB791 <i>serA</i> on M9/ACA/ros	60
Figure 2.1. Secondary chemicals that can be produced from the 3-HP platform chemical	72
Figure 2.2. 3-HP synthesis from glycerol via CoA-dependent and CoA-independent pathways.....	74
Figure 2.3. 3-HP synthesis from glucose via the malonyl-CoA pathway	76
Figure 2.4. Synthesis of 3-HP from glucose via the β -alanine pathway	77
Figure 2.5. Substrate specificity of YdfG and YMR226C	79
Figure 2.6. Nucleotide sequences of (A) YdfG and (B) RutE	81
Figure 2.7. Plasmids used for the expression of His ₆ -tagged (A) YdfG and (B) RutE.	82
Figure 2.8. Purification of His ₆ -tagged YdfG from BL21(DE3)/pAS2.141 (1 L) by affinity chromatography.	83
Figure 2.9. Purification of His ₆ -tagged RutE from BL21(DE3)/pAS2.144 (1 L) by affinity chromatography.	84
Figure 2.10. Isolation of YdfG (flowthrough) from His ₇ -tagged TEV protease via affinity chromatography.	85
Figure 2.11. Isolation of RutE (flowthrough) from His ₇ -tagged TEV protease via affinity chromatography.	86
Figure 2.12. Coupled enzyme assay used to determine activity of RutE/YdfG	88
Figure 2.13. Purification of YdfG from BL21(DE3)/pAS2.084 (1 L) via affinity chromatography	89
Figure 2.14. Michaelis Menten kinetic parameters of YdfG	90
Figure 2.15. ¹ H NMR characterization of YdfG	92
Figure 2.16. Enzyme assay used to determine measure PTDH activity	94

Figure 2.17. Purification of PTDH from BL21(DE3)/pET15b-12x (1 L) via affinity chromatography.	95
Figure 2.18. Michaelis Menten kinetics of PTDH.....	97
Figure 2.19. Relative activity of Cg10062(E114N) in various buffers.....	98
Figure 2.20. Relative activity of YdfG in various buffers.	100
Figure 2.21. Relative activity of PTDH in various buffers.....	101
Figure 2.22. Biocatalytic route to 3-HP via ACA	102
Figure 2.23. SDS-PAGE of enzymes used for 3-HP synthesis.....	102
Figure 2.24. HPLC chromatogram of (A) ACA and (B) 3-HP	104
Figure 2.25. HPLC calibration curve of ACA	105
Figure 2.26. HPLC calibration curve of ACA	105
Figure 2.27. HPLC chromatogram of (A) glycerol (RT = 13.5 min) and (B) ethylene glycol (RT = 16 min).....	106
Figure 2.28. HPLC chromatogram of 3-HP synthesis from 100 mM ACA with 1 mM NADP(H) at (A) 1 h and (B) 5 h after initiation.....	108
Figure 2.29. <i>In vitro</i> synthesis of 3-HP from ACA	109
Figure 2.30. <i>In vitro</i> synthesis of 3-HP from ACA without cofactor regeneration.....	109
Figure 2.31. HPLC chromatogram of 3-HP synthesis from 100 mM ACA with 0.1 mM NADP(H) (A) immediately after initiation and (B) 30 h after initiation	111
Figure 2.32. <i>In vitro</i> synthesis of 3-HP from 100 mM ACA with cofactor regeneration.....	112
Figure 2.33. ¹ H NMR of 3-HP synthesis with 100 mM ACA and 10 mM NADP(H).....	114
Figure 2.34. ¹ H NMR of 3-HP synthesis with 100 mM ACA and 1 mM NADP(H).....	115
Figure 2.35. ¹ H NMR of 3-HP synthesis with 100 mM ACA and 0.1 mM NADP(H).....	116
Figure 2.36. <i>In vitro</i> synthesis of 3-HP from 500 mM ACA	118
Figure 2.37. ¹ H NMR of 3-HP synthesis with 500 mM ACA and 50 mM NADP(H)	119
Figure 2.38. ¹ H NMR of 3-HP synthesis with 500 mM ACA and 5 mM NADP(H).....	120
Figure 2.39. ¹ H NMR of 3-HP synthesis with 500 mM ACA and 0.5 mM NADP(H).....	121
Figure 2.40. Modes of enzyme inhibition	123

Figure 2.41. The coupled enzyme assay used to measure Cg10062(E114N) activity.....	124
Figure 2.42. Quantification of NADH using PicoProbe™	125
Figure 2.43. PicoProbe™- NADH standard curve	126
Figure 2.44. NADPH Inhibition Screen for Cg10062(E114N) using PicoProbe™	127
Figure 2.45. IC ₅₀ of NADPH for Cg10062(E114N) inhibition	128
Figure 2.46. Plasmid map of pAS3.112 was used for <i>in vivo</i> synthesis of 3-HP.	130
Figure 2.47. <i>In vivo</i> 3-HP synthesis by BL21/pAS3.112 in LB media supplemented with ACA	131
Figure 2.48. <i>In vivo</i> 3-HP synthesis by BL21/pAS3.112 in M9 media supplemented with ACA.....	132
Figure 2.49. Expression of Cg10062(E114N) and YdfG by BL21/pAS3.112 in M9 media	133
Figure 3.1. 3-HP degradation pathway in <i>P. denitrificans</i>	148
Figure 3.2. A genetically-encoded 3-HP biosensor in <i>P. denitrificans</i>	151
Figure 3.3. Plasmid map of pAS5.003 used for expression of MmsR and RFP.	153
Figure 3.4. Plasmid map of pAS5.008 used for expression of Cg10062(E114N) and YdfG.	154
Figure 3.5. Plasmid map of pAS5.005 used for expression of Cg10062 and YdfG.	155
Figure 3.6. 3-HP biosensor for evolution of Cg10062(E114N) in <i>E. coli</i> MG1655.....	156
Figure 3.7. RFP fluorescent signal in <i>E. coli</i> MG1655/pBbA1a-RFP.....	159
Figure 3.8. Growth inhibition of 3HPC with increasing 3-HP concentrations.....	161
Figure 3.9. Growth inhibition of 3HPC-WT with increasing 3-HP concentrations.	162
Figure 3.10. Growth of 3HPC with increasing ACA concentrations.....	163
Figure 3.11. Growth of 3HPC-WT with increasing ACA concentrations.....	164
Figure 3.12. ¹ H NMR of 3HPC-E114N in M9/glu/ACA media after 48 h.	167
Figure 3.13. ¹ H NMR of 3HPC-WT in M9/glu/ACA media after 48 h.....	168
Figure 3.14. ¹ H NMR of 3HPC in M9/glu/ACA media after 48 h.....	169
Figure 4.1. Standard calibration curve for the Bradford protein assay	185
Figure 4.2. Plasmid map of pAS1.046.....	191

Figure 4.3. Nucleotide sequence of <i>cg10062</i> optimized for expression in <i>E. coli</i>	191
Figure 4.4. Codon-optimized nucleotide sequence of MSAD.....	192
Figure 4.5. Plasmid map of pAS2.031	192
Figure 4.6. Construction of plasmid pAS1.163 expressing Cg10062.....	194
Figure 4.7. The coupled enzyme assay used for Cg10062 and variants	198
Figure 4.8. Plasmid map of pAS2.141.....	204
Figure 4.9. Plasmid map of pAS2.144.....	205
Figure 4.10. Plasmid map of pAS2.084.....	206
Figure 4.11. Plasmid map of pET15b-12x	207
Figure 4.12. Construction of pAS3.112 using NEB HiFi DNA Assembly.....	209
Figure 4.13. RutE/YdfG coupled enzyme assay for initial screening.....	210
Figure 4.14. YdfG coupled enzyme assay.....	211
Figure 4.15. Enzyme assay used to determine measure PTDH activity	213
Figure 4.18. Modified coupled enzyme assay for monitoring Cg10062(E114N) inhibition by NADPH using PicoProbe™	219
Figure 4.19. The DNA fragment containing promoter J23119.....	222
Figure 4.20. Construction of pAS5.003 using NEB HiFi DNA Assembly.....	223
Figure 4.21. The DNA fragments synthesized for pAS5.005	224
Figure 4.22. Construction of pAS5.005 using NEB HiFi DNA Assembly.....	225
Figure 4.23. Final construction of pAS5.005 using NEB HiFi DNA Assembly	227

KEY TO ABBREVIATIONS

3-HP	3-hydroxypropionic acid/3-hydroxypropionate
3-HPA	3-hydroxypropionaldehyde
3-HP-CoA	3-hydroxypropionyl-CoA
ACA	acetylenecarboxylic acid/ acetylenecarboxylate
ACC	acetyl-CoA carboxylase
AcOH	acetic acid
ADCA	acetylenedicarboxylic acid/ acetylenedicarboxylate
ADH	alcohol dehydrogenase
ALDH	aldehyde dehydrogenase
Ap	ampicillin
cCA	<i>cis</i> -3-chloroacrylic acid/ <i>cis</i> -3-chloroacrylate
<i>cis</i> -CaaD	<i>cis</i> -3-chloroacrylate dehydrogenase
CoA	coenzyme A
DNA	deoxyribonucleic acid
DTT	dithiothreitol
GDHt	glycerol dehydratase
glu	glucose
h	hours(s)
IPTG	isopropyl- β -D-1-thiogalactopyranoside
Kan	kanamycin
kb	kilobase
kDa	kilodalton
LB	Luria-Bertani

LTTR	LysR-type transcriptional regulator
M	molar
MCR	malonyl-CoA reductase
MeOH	methanol
mg	milligram
min	minute(s)
mL	milliliter
mM	millimolar
MOPS	3-morpholinopropanesulfonic acid
MSA	malonic/ malonate semialdehyde
MSAD	malonate semialdehyde decarboxylase
NAD ⁺	nicotinamide adenine dinucleotide, oxidized form
NADH	nicotinamide adenine dinucleotide, reduced form
NADP ⁺	nicotinamide adenine dinucleotide phosphate, oxidized form
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NB	nutrient broth
NMR	nuclear magnetic resonance
OD	optical density
PCR	polymerase chain reaction
<i>pdu</i>	propanediol utilization
PduL	phosphate propanoyltransferase
PduP	propionaldehyde dehydrogenase
PduW	propionate kinase
ros	<i>p</i> -rosaniline
s	second(s)
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis

ser	L-serine
TAE	tris-acetate EDTA
TCA	tricarboxylic acid
Th	thiamine
U	unit; $\mu\text{mol min}^{-1}$

CHAPTER ONE: Cg10062 CATALYSIS FORGES A LINK BETWEEN ACETYLENECARBOXYLIC ACID AND BACTERIAL METABOLISM

Reprinted with permission from Mathes Hewage, A.; Gavgani, H. N.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochemistry* 2021, 60 (51), 3879–3886.¹ Copyright 2022 American Chemical Society.

1.1. Introduction

Identification of optimal sources of carbon from which performance-advantaged, life-enhancing goods can be sustainably and inexpensively made is a critical challenge for the chemical enterprise.² Throughout the 20th century and continuing to the present, petroleum-derived carbon has dominated the fuels, chemicals, and consumer products industries, meeting over 80% of the world's energy needs.^{2,3} However, rapid industrialization driven by the crude oil-based economy has led to alarming levels of greenhouse gas (GHG) emissions.^{1,2} The increase in atmospheric levels of GHGs cause rising global temperatures and climate change, with irreparable impacts on biodiversity.^{4,5} In the United States, refinement, storage and transportation of petroleum is one of largest sources of GHG emissions.⁶ Replacing these traditional practices which rely on finite petroleum reserves⁷ is of paramount importance. A consensus to establish a sustainable chemical industry^{8,9} has prompted a shift in emphasis to biomass as a preferred carbon source.¹⁰

Efforts to move away from petroleum-derived carbon and create a bioeconomy by developing sustainable syntheses using renewable carbon and microbial catalysis, began a few decades ago.^{11,12} The shift to a bio-based economy is reliant on biotechnological advances that are not only environmental-friendly but also economically feasible, to make them more attractive relative to current practices.¹² Biocatalytic methods ranging from utilization of a single isolated enzyme to a consortium of microorganisms undoubtedly offer enormous potential for fuel and chemical production.¹³ Extensive use of natural and engineered enzymes in pharmaceutical

syntheses highlight exquisite selectivity and stereospecificity in their catalytic mechanisms.^{14,15} Biocatalytic methodologies afford access to specialty chemicals, building block chemicals, and molecules with enhanced physical properties^{16–19} and biomass, such as renewable sugars or lignocellulosic feedstocks have been explored for the replacement of petroleum-derived feedstocks in the chemical industry.²⁰ While a vast majority of the early biocatalytic processes rely on polysaccharide and polyol feedstocks obtained from edible crops (first generation feedstock),²¹ the food versus chemicals debate has prompted the scientific community to consider alternate carbon sources.²² Nevertheless, despite enormous research and investment, the lignocellulosic biorefinery in which non-edible biomass (second generation feedstocks) is converted into low-cost fuels and high-value chemicals is yet to be successfully implemented.^{2,20} Additionally, the environmental and ecological impacts of industrial-scale farming required to generate feedstocks sufficient for chemical biosynthesis deter the use of biomass as the primary source of carbon.²³

As an alternative to biomass, bioconversion of carbon dioxide (CO₂)^{24–26} and methane (CH₄),^{27–29} (third generation feedstocks) are being explored as promising candidates due to their natural abundance. For example, the production of muconic acid and 3-hydroxypropionic acid have been demonstrated using CH₄³⁰ and CO₂,³¹ respectively. CH₄, in particular is an abundant source of carbon which can be sourced from non-renewable sources such as fossil fuels (e.g., natural gas, shale gas) or from renewable biogas (anaerobic digestion of waste).³² The use of CO₂ and CH₄ directly for fuel and chemical synthesis would have the added benefit of attenuating increasing atmospheric concentrations of these potent GHGs. Unfortunately, CO₂ and CH₄ bioconversion are hindered by multiple challenges including enzyme mechanism elucidation and uncertainty related to kinetic improvement, elaboration of metabolic pathways, and bioreactor design.³³ Given the challenges faced by the implementation of C1 gaseous feedstocks for biocatalysis, the possibility of transforming CO₂ and CH₄ into one or more molecules that may be used as alternate substrates for biocatalytic transformation was investigated. In this study,

acetylenecarboxylic acid (ACA) is proposed as a unique source of carbon that can be derived from CH₄ and CO₂ and be utilized more efficiently in microbial syntheses.

1.2. Synthesis of ACA from CH₄ and CO₂

Examination of the literature led to acetylenecarboxylic acid (ACA) as a potential carbon feedstock for microbial catalysis. At present, ACA **2** synthesis is carried out via oxidation of propargyl alcohol (Figure 1.1).³⁴⁻³⁷ Although ACA **2** synthesis is not currently practiced on an industrial scale, sufficient literature indicates that ACA could be derived from CH₄ and CO₂ in high yield. Recent technology reported by Dinh et al. describes the dehydrodimerization of CH₄ to yield acetylene **1** with high selectivity and conversion rates.³⁸ In this plasma-mediated process, CH₄ is transported through a reactor at supersonic speeds using H₂ as a carrier gas, where the thermal plasma activates CH₄ for acetylene **1** formation (Figure 1.2). With rapid thermal quenching, acetylene was produced with >95% selectivity and 70% conversion, with unreacted CH₄ being recycled through the reactor. However, although acetylene **1** can be derived from the C1 waste gas in high yield, the combination of its instability and the inability to store or transport acetylene in condensed phase prohibits its utilization in biocatalysis. However, reports describing the derivatization of terminal alkynes with CO₂ indicate that acetylene **1** may be carboxylated to form ACA **2** or acetylenedicarboxylic acid (ADCA) **3**.³⁹⁻⁴⁷ Early studies demonstrated carboxylation of monosubstituted alkynes using Cu(I) or Ag(I) salts with stoichiometric addition of an inorganic base.^{39,45,48} Velazquez et al. reported the carboxylation of acetylene using Cu(I) or Ag(I) *bis*-(*N*-heterocyclic carbene)-metal complexes with 40% yield.⁴⁰ More recently Wang et al. showed evidence of ACA **2** and ADCA **3** formation from acetylene in a crude mixture using organic bases DBU and TBD, respectively (Figure 1.3).⁴¹ In this study, the alkynoic acids were not isolated but subjected to Pd-catalyzed hydrogenation. Relative to acetylene, ACA **2** and ADCA **3** are a stable liquid and solid that can be safely stored and transported. The missing link between bacterial metabolism and ACA **2** or ADCA **3**, is an enzyme that can transform these compounds into

metabolites that are found in common platform organisms such as *E. coli*, *Corynebacterium glutamicum* or *Pseudomonas* species.

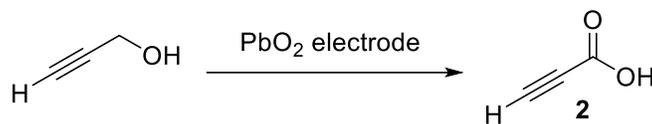


Figure 1.1. Anodic oxidation of propargyl alcohol.

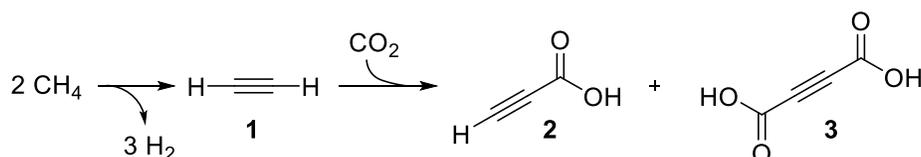


Figure 1.2. General conditions for the synthesis of ACA and ADCA from CH₄ and CO₂.

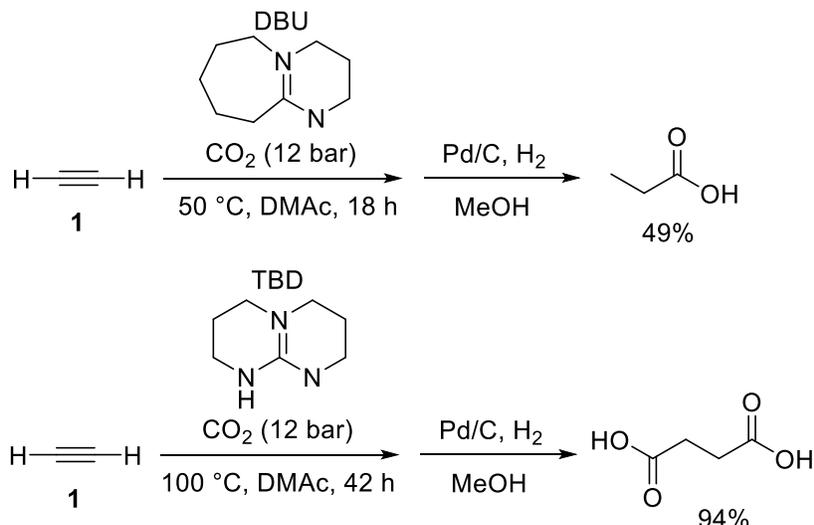


Figure 1.3. Base-dependent carboxylation of acetylene followed by hydrogenation.

1.3. ACA and bacterial metabolism

The first report of an ACA-utilizing enzyme was described by Yamada and Jakoby in 1959.⁴⁹ The “acetylenemonocarboxylic acid hydrolase” was isolated from the cell-free lysate of a microorganism obtained by enrichment culturing in minimal media containing 0.1% ACA as the sole carbon source.⁴⁹ Incubation of ACA with the isolated enzyme resulted in formation of malonate semialdehyde (MSA). It was reported that the enzyme did not require addition of

exogenous coenzymes or metal cofactors for catalytic activity.⁴⁹ The reported organism was able to grow on additional carbon sources including D-glucose and malonic acid; however, it was noted that ACA hydratase was expressed only when the strain was cultured in media containing ACA as the exclusive carbon source.⁴⁹ Similarly, the enzyme capable of transforming ADCA to pyruvate via oxaloacetate was also purified from a different organism.⁵⁰ Unfortunately, the organisms were only identified as belonging to the Gram-negative *Pseudomonas* sp., and were not archived in a culture collection. In this research study, organisms capable of utilizing ACA or ADCA as an exclusive carbon source for growth were isolated but the enzymes responsible for this growth were not successfully identified (data not shown).

As a result, research focused on the more recently described *cis*-CaaD and Cg10062 catalyzed hydration of ACA described extensively by Whitman and coworkers.⁵¹ Both enzymes belong to the tautomerase superfamily (TSF), a group of over 11,000 proteins from all kingdoms of life that have been identified thus far through biochemical and computational methods.⁵² A key characteristic of this superfamily is the presence of an unusual catalytic N-terminal Pro-1 residue in a vast majority of enzymes (346 TSF sequences were recently identified as lacking Pro-1).⁵¹ Enzymes are classified as members of the TSF based on structural homology, where they assume a β - α - β protein fold, and the presence of the catalytic Pro-1.⁵¹ Cg10062 and *cis*-CaaD form homotrimers mediated by interaction of the C-terminus of one monomer with the N-terminus of a second monomer to form the active site, resulting in three active sites per trimer.⁵³

Tautomerases *cis*-CaaD, *trans*-3-chloroacrylic acid dehalogenase (CaaD) and malonate semialdehyde dehydrogenase (MSAD) were discovered as part of the well-established 1,3-dichloropropene degradation pathway soil bacteria.⁵⁴ 1,3-Dichloropropene 4 is an active ingredient in soil fumigants used in the agricultural industry, and the 1,3-dichloropropene catabolic pathway (Figure 1.4) allows certain bacterial species to use the compound as a sole source of carbon for biomass.⁵⁵ CaaD catalyzes the hydrolytic dehalogenation of *trans*-3-chloroacrylate 5 to produce malonate semialdehyde (MSA) 7⁵⁶ while *cis*-CaaD catalyzes the same reaction with

cis-3-chloroacrylate 6.⁵⁵ The MSA 7 intermediate produced by the two isomer-specific dehalogenases are then decarboxylated to acetaldehyde 8 by MSAD.⁵⁷

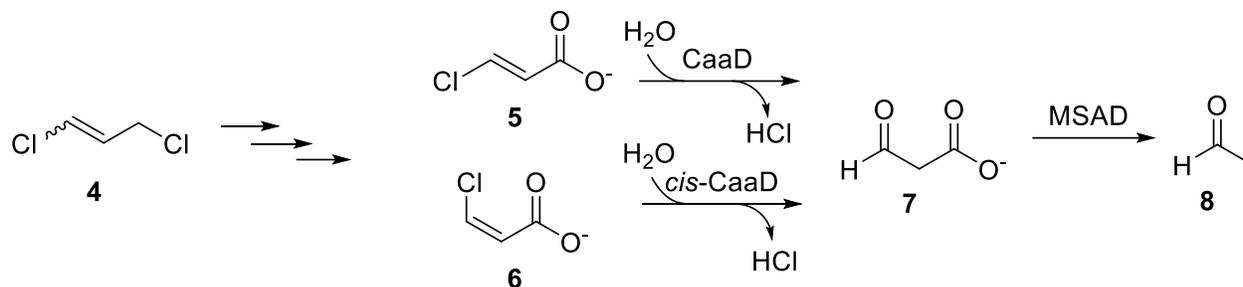


Figure 1.4. The 1,3-dichloropropene catabolic pathway. CaaD, *trans*-3-chloroacrylate dehalogenase; *cis*-CaaD, *cis*-3-chloroacrylate dehalogenase; MSAD, malonate semialdehyde decarboxylase; 4, 1,3-dichloropropene; 5, *trans*-3-chloroacrylate; 6, *cis*-3-chloroacrylate; 7, malonate semialdehyde; 8, acetaldehyde.

Whitman and coworkers previously carried out detailed characterization of CaaD, MSAD and *cis*-CaaD tautomerase.^{55,56} In particular, *cis*-CaaD and MSAD enzymes native to *Pseudomonas pavonaceae* 170 and *Coryneform* strain FG41 have been studied extensively.^{51,55,57,58} In 2008, Cg10062 native to *Corynebacterium glutamicum* was identified as a *cis*-CaaD homologue using a sequence homology search.⁵¹ While *cis*-CaaD has been implicated in the 1,3-dichloropropene 4 catabolic pathway, the location of the *cg10062* gene on the chromosome does not provide any clues regarding its physiological function.^{51,54} Like other members of the TSF, Cg10062 includes the characteristic β - α - β protein fold and catalytic Pro-1. Interestingly, *cis*-CaaD and Cg10062 are both homotrimers (149 amino acids per monomer) that share 53% sequence similarity and six residues implicated in *cis*-CaaD catalytic activity are also

found in Cg10062.⁵¹ The active site of both enzymes is composed of Pro-1, His-28, Arg-70, Arg-73, Glu-114 from one monomer and Tyr-103 from a neighboring monomer (Figure 1.5).⁵¹

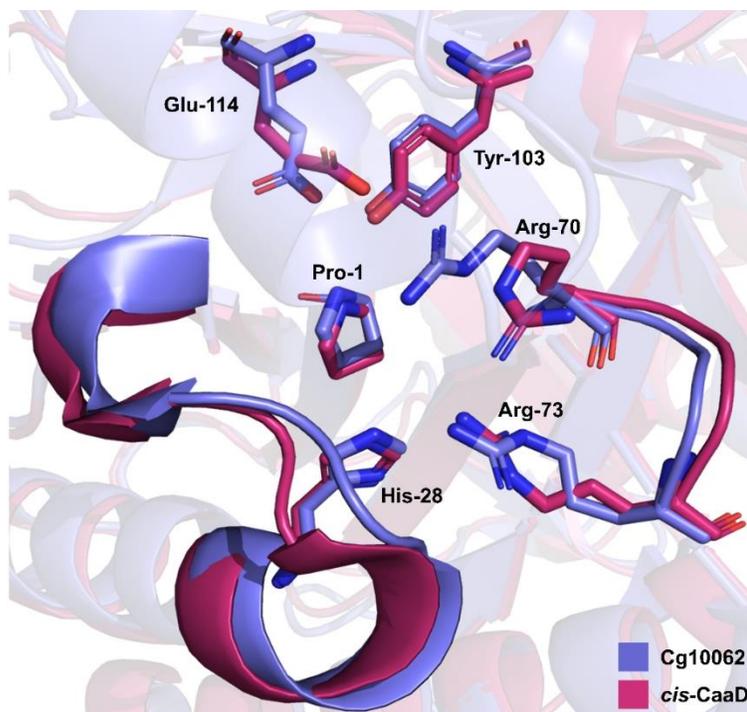


Figure 1.5. An overlay of the active sites of *cis*-CaaD (PDB 3MF8) and Cg10062 (PDB 3N4G).

Despite these striking structural and sequence similarities, the catalytic activity of the two tautomerases vary significantly.⁵⁹ For example, k_{cat}/K_m of *cis*-3-chloroacrylate **6** is strikingly different for the two enzymes: $3.0 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ for *cis*-CaaD and $14 \text{ M}^{-1} \text{ s}^{-1}$ for Cg10062.⁵¹ Furthermore, while *cis*-CaaD does not catalyze the hydration of *trans*-3-chloroacrylate **5** to a measurable extent, Cg10062 catalyzes the hydrolytic dehalogenation of both *cis* and *trans* isomers. However, kinetic data supports a stronger preference for the *cis* isomer (18-fold) relative to the *trans* isomer.⁵¹ More importantly, both *cis*-CaaD and Cg10062 were reported to utilize ACA **2** as a substrate (Figure 1.6).⁵⁹

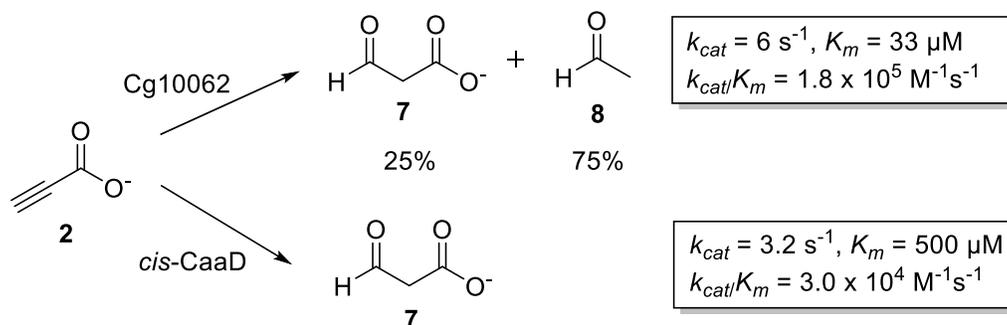


Figure 1.6. Varying catalytic activity of Cg10062 and *cis*-CaaD with ACA.

1.4. Hydration of ACA by Cg10062

In order to access value-added C3 chemicals from ACA, it is highly desirable to identify an enzyme that acts exclusively as a hydratase when ACA **2** is the substrate. Although both *cis*-CaaD and Cg10062 catalyze the hydration of ACA **2**, their catalytic activities vary significantly in terms of overall efficiency and product distribution (Figure 1.6). *cis*-CaaD produces exclusively MSA **7** while Cg10062 produces a mixture of MSA **7** and acetaldehyde **8**. It was reported that the formation of acetaldehyde is hydration-dependent since the addition of exogenous MSA **7** did not result in the enzyme-catalyzed formation of acetaldehyde **8**. Cg10062 was selected for this study due to its catalytic superiority (6-fold higher k_{cat}/K_m) with ACA **2** relative to *cis*-CaaD. However, the hydration-dependent decarboxylation of MSA **7** to acetaldehyde **8** by Cg10062 represents the immediate loss of a one carbon. Huddleston et al. report two active site variants, Cg10062(E114Q) and Cg10062(E114D), that catalyze hydration without also catalyzing decarboxylation to produce exclusively MSA **7**, albeit at a slower rate than the wild type enzyme.⁵⁹ The steady-state kinetic parameters previously reported for Cg10062 (and variants) by Whitman and coworkers⁵⁹ are summarized below (Table 1.1).

Table 1.1. Steady state kinetic parameters of Cg10062 and variants⁵⁹

variant	substrate	activity ^a	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{M}^{-1} \text{s}^{-1}$)
wild type	ACA 2	H/D	33	6	1.8×10^5
	<i>cis</i> -3-chloroacrylate 6	H/D	72,000	1	14
	<i>trans</i> -3-chloroacrylate 5	H	78,000	0.06	0.8
E114Q	ACA 2	H	3	0.8	2.7×10^5
	<i>cis</i> -3-chloroacrylate 6	H	4,000	0.4	100
E114D	ACA 2	H	90	1	1.1×10^4
	<i>cis</i> -3-chloroacrylate 6	H	40,100	0.1	2.5
Y103F	ACA 2	H/D	5	0.5	1.0×10^5
	<i>cis</i> -3-chloroacrylate 6	H/D	4,700	0.3	60

^aH: hydratase; H/D: hydratase/decarboxylase

The immediate goal of this study was to optimize Cg10062 to efficiently convert ACA into MSA without decarboxylation to maximize carbon flux from CH₄- and CO₂-derived ACA into useful biocatalytic end products. In order to accomplish this goal, a good understanding of the mechanisms of Cg10062 and elucidation of which amino acid residues influence hydratase versus hydratase/decarboxylase activity is crucial (Figure 1.5). In *cis*-CaaD, the carboxylate group of ACA 2 is presumed to be stabilized by the positively charged Arg-70, Arg-73 and His-28 residues.⁵¹ This interaction can induce a partial positive charge on C-3 of the substrate to facilitate hydration. Glu-114 and Tyr-103 have been implicated in the activation of the water molecule, allowing the Michael addition onto the C-3 of the substrate. These residues are presumed to play a similar role in Cg10062 catalysis of ACA 2 and Huddleston et al. suggest that a mechanism involving the formation of covalent enzyme–substrate intermediates may be at play, at least in the case of Cg10062.⁵⁹ Collectively, these studies have highlighted three potential mechanisms for the Cg10062-catalyzed hydration of ACA.

In mechanism 1, the conversion of ACA **2** to MSA **7** is catalyzed by Cg10062, followed by decarboxylation to acetaldehyde **8** (Figure 1.7) under the assumption that like in *cis*-CaaD, Pro-1 acts as a proton donor.^{59,60} The water molecule is activated by Glu-114 and Tyr-103 residues for Michael addition on C-3 of ACA **2** to form intermediate **9**, which was hypothesized to accept a proton from the Pro-1 residue. Tautomerization of the resulting intermediate produces MSA **7**, some of which may be released from the active site, accounting for MSA **7** formation (25% in native Cg10062). The remaining MSA **7** is decarboxylated in the active site by an unknown mechanism, yielding acetaldehyde **8**. However, this mechanism argues against the fact that Cg10062 does not decarboxylate exogenously added MSA **7** although the authors suggest that the hydration of ACA **2** may position MSA **7** in a favorable position for decarboxylation, that may not be achieved via exogenous addition of MSA **7**.

In mechanism 2, direct hydration of ACA **2** is followed by decarboxylation via Schiff base intermediate (Figure 1.8).⁵⁹ The covalent intermediate **11** is a result of the nucleophilic attack by Pro-1 (deprotonated after the hydration step) on C-3 of MSA **7**. Decarboxylation of Schiff base **11** results in the formation of enamine **12**, which can undergo rearrangement to form **13**. Hydrolysis of intermediate **14** results in the formation of acetaldehyde **8** and the free enzyme. The MSA **7** (25%) observed with Cg10062 may be due to release of a fraction of the hydration product from the active site. Since the initial hydration of ACA **2** results in a nucleophilic Pro-1, the lack of activity with exogenously added MSA **7** was presumably due to the inability of the protonated Pro-1 to form a Schiff base with MSA **7**.

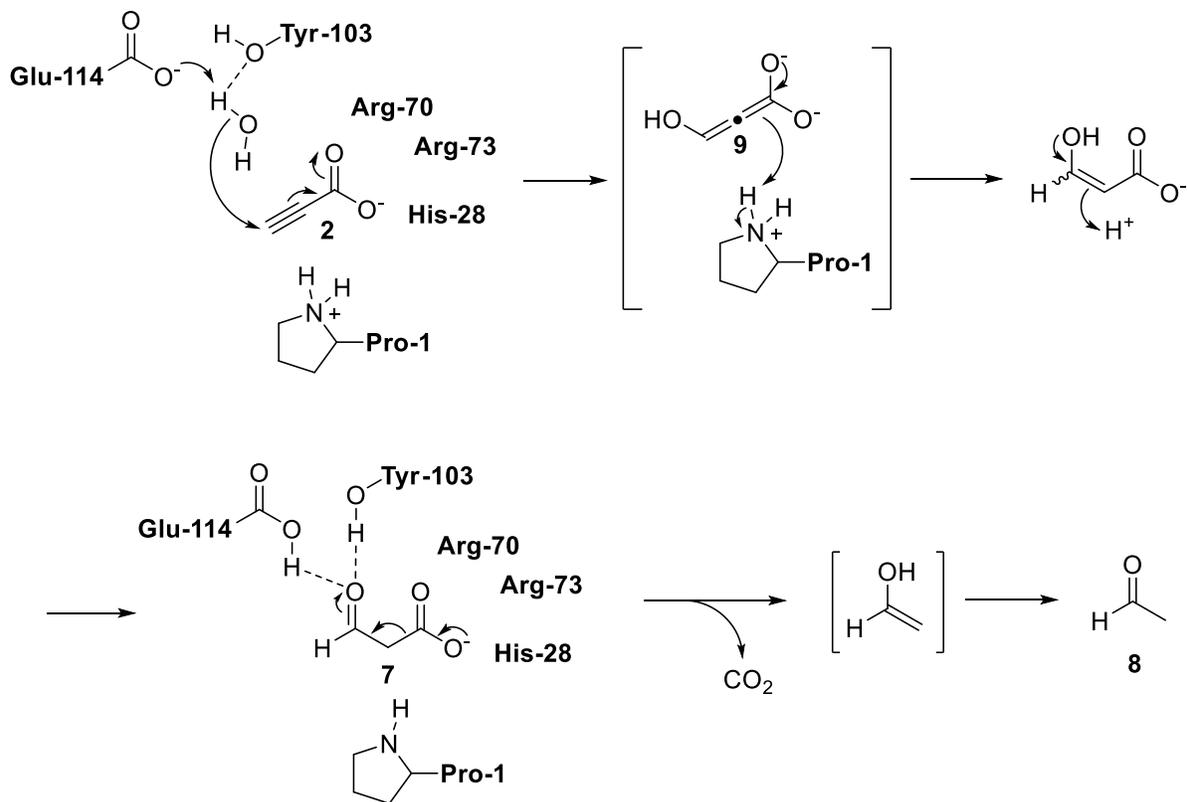


Figure 1.7. Mechanism 1 proposed by Huddleston et al.⁵⁹ Cg10062-catalyzed ACA hydration followed by decarboxylation with Pro-1 acting as a general acid catalyst.

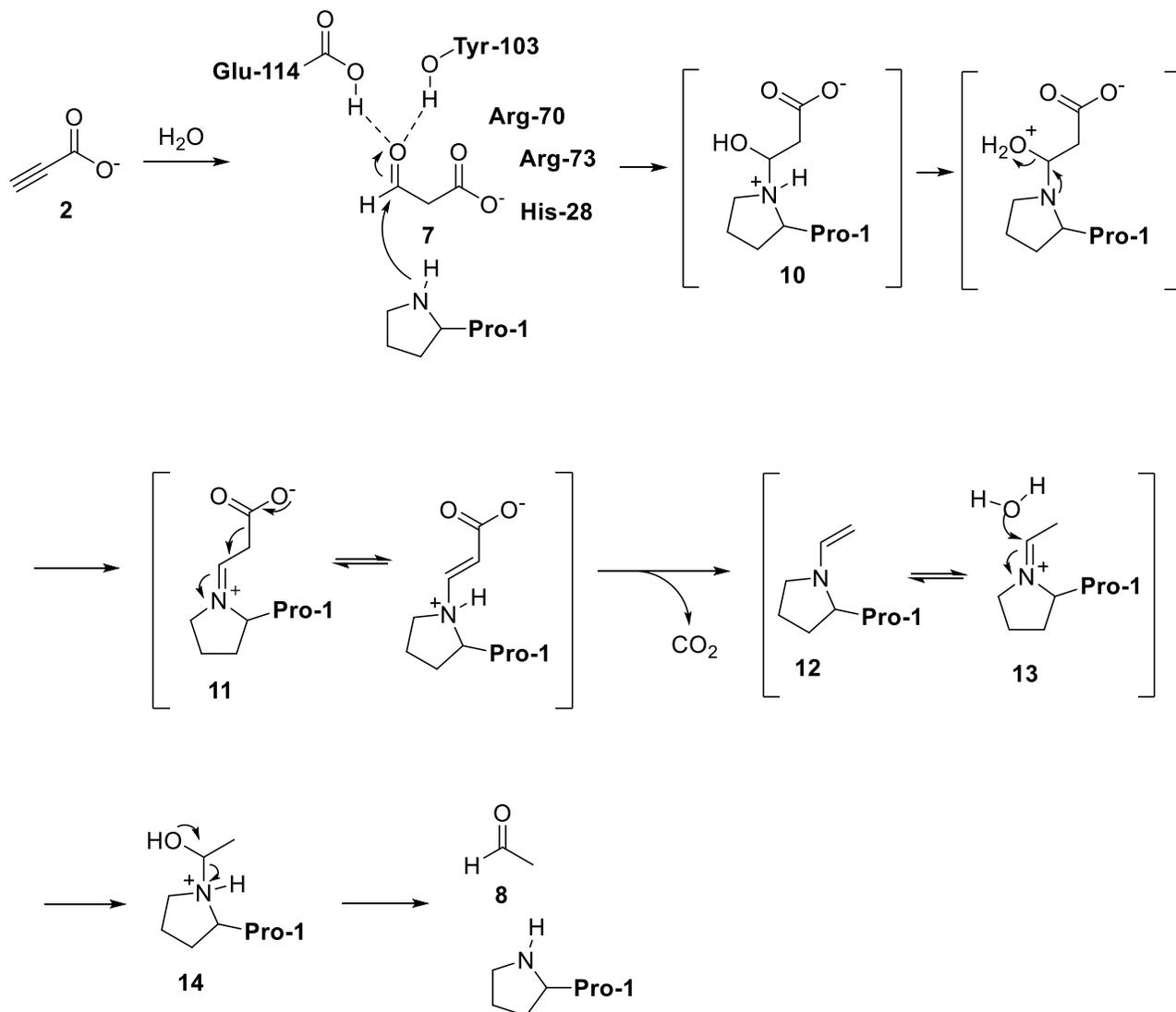


Figure 1.8. Mechanism 2 proposed by Huddleston et al.⁵⁹ Cg10062-catalyzed ACA hydration proceeds by direct attack of water and decarboxylation via a Schiff base intermediate.

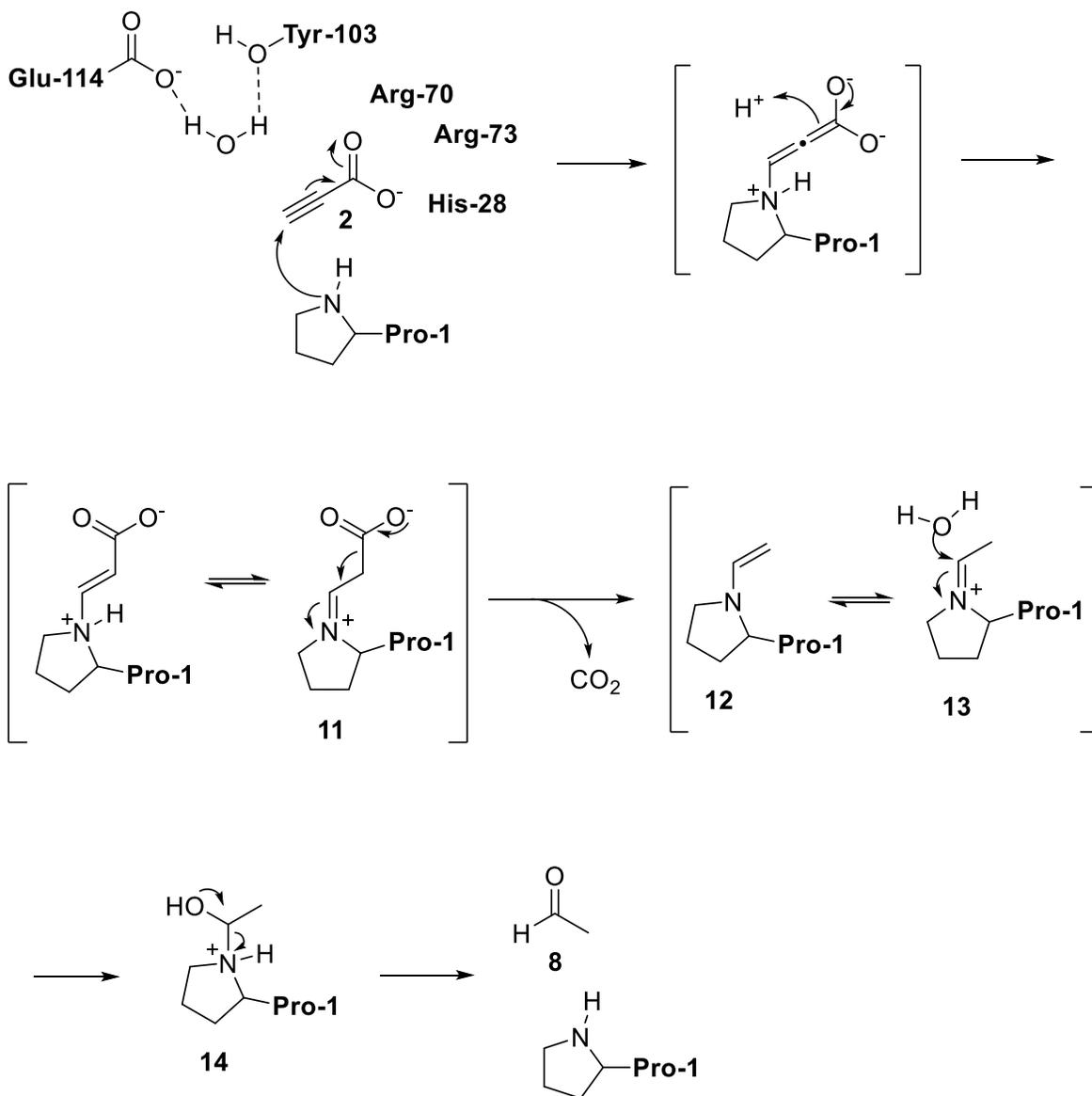


Figure 1.9. Mechanism 3 proposed by Huddleston et al.⁵⁹ Cg10062-catalyzed hydration of ACA proceeds via covalent catalysis.

In mechanism 3 (Figure 1.9), Pro-1 is proposed to initiate nucleophilic attack on C-3 of the substrate to form intermediate **11**, which can decarboxylate to form intermediate **12**.⁵⁹ Tautomerization of **12** can result in intermediate **13** which can undergo hydrolysis to yield acetaldehyde **8** and free enzyme. Any MSA **7** formed was attributed to the hydrolysis of intermediate **11**. The authors note that a limitation of this mechanism is the assumption that the pK_a of Pro-1 in Cg10062 is similar to the measured pK_a (~9.2) in *cis*-CaaD, suggesting that Pro-1

is in its protonated state and unable to form a covalent intermediate with the substrate.⁵⁹ However, a pH rate profile of Cg10062 was not carried out and the pK_a is an assumption based on prior studies with *cis*-CaaD.⁶⁰ If catalysis indeed proceeds through a covalent intermediate, it has been suggested that the pH rate profile may not reflect the true pK_a of the prolyl nitrogen or that a fraction of the Pro-1 may be present in the unprotonated form.^{59,60} Any addition of exogenous MSA 7 may not be decarboxylated due to improper orientation within the active site, thereby preventing Cg10062-catalyzed decarboxylation.

In order to utilize hydratase-only Cg10062 variants for production of C3 chemicals via ACA, a good understanding of the mechanism is crucial. Herein, a combination of mutagenesis, enzymology and structural biology was used to clarify these mechanistic questions and to ultimately obtain a novel hydratase-only variant of Cg10062 with improved kinetic parameters relative to those previously reported in the literature. Structural elucidation of substrate-soaked Cg10062 variants, capturing key intermediates bound in the active site at atomic resolution, provide a clear three-dimensional picture of this multistep mechanism.

1.5. Rational mutagenesis of the Cg10062 active site

In previous reports, incubation of Cg10062(E114Q) and Cg10062(E114D) with ACA 2, followed by analyses of the resulting products by ¹H NMR confirmed that the two variants only displayed hydratase activity.⁵⁹ Given our interest in identifying enzymes lacking hydratase/decarboxylase activity, Q5 site-directed mutagenesis was used to generate several active site variants of Cg10062, including the previously identified E114Q and E114D variants. Additional active site variants, E114N, Y103F, Y103A, E114D-Y103F, H28A, R70A, R73A, R70K, R73K, E114A and E114S were generated. The P1A variant was not examined in this study as previous accounts indicated complete loss of catalytic activity.^{51,59} Cg10062 from *Corynebacterium glutamicum*^{51,59} was modified with a TEV protease recognition site (ENLYFQG)⁶¹ at the C-terminus (Figure 1.10). The gene was cloned into the commercial vector pET-21a(+) containing a C-terminal His₆-tag and a T7 promoter to enable high-level expression of recombinant Cg10062.

Purification of His₆-tagged Cg10062 and variants were carried using nickel affinity chromatography (Figure 1.11). The N-terminus of Cg10062 was not modified due to the importance of Pro-1 for catalytic activity. Although affinity tags are common tools that are beneficial in protein purification, there is potential for these tags to interfere with protein structure and folding and/or its function or catalytic activity.⁶² To overcome this limitation, researchers often incorporate methods for removal of affinity tags from the target protein. The canonical Tobacco Etch Virus (TEV) protease recognition sequence ENLYFQG is recognized and cleaved by TEV protease between the glutamate (Q) and glycine (G) residues.⁶³ The TEV protease recognition site was included for removal of the His₆-tag (if necessary) should it interfere with catalysis or crystallization efforts. Optimization of expression and purification conditions was not required due to high levels of protein expression. An additional band consistent with the dimer of wild type Cg10062 was analyzed by SDS-PAGE (Figure 1.12). The band was excised from the gel for protein identification and characterization at the MSU Proteomics Facility. Proteomic analysis of the band at approximately 37 kDa was matched against all non-redundant protein sequences on the NCBI database and confirmed to match wild type Cg10062 (data not shown). Interestingly, similar bands were not observed for any of the other variants.

```

5' – CATATGCCGACCTACACCTGCTGGAGCCAACGCATTCGTATTAGCCGTGAAG
CGAAGCAACGCATCGCGGAAGCGATTACCGACGCGCACCATGAACTGGCGCAC
GCGCCGAAGTACCTGGTGCAGGTTATTTTCAACGAAGTGGAGCCGGACAGCTATT
TTATCGCGGCGCAGAGCGCGAGCGAGAACCACATTTGGGTTCAAGCGACCATCC
GTAGCGGCCGTACCGAAAAGCAGAAAGAGGAACTGCTGCTGCGTCTGACCCAAG
AGATCGCGCTGATTCTGGGTATCCCGAACGAGGAAGTGTGGGTTTACATTACCGA
AATCCCGGGTAGCAACATGACCGAATATGGCCGTCTGCTGATGGAGCCGGGCGA
GGAAGAGAAATGGTTCAACAGCCTGCCGGAGGGCCTGCGTGAGCGTCTGACCG
AACTGGAGGGTAGCAGCGAAGAGAACCTGTATTTTCAAGGCCTCGAG – 3'

```

Figure 1.10. Codon-optimized nucleotide sequence of *cg10062*. The NdeI and XhoI restriction sites are highlighted in yellow, and the TEV protease recognition site (ENLYFQG) is highlighted in blue.

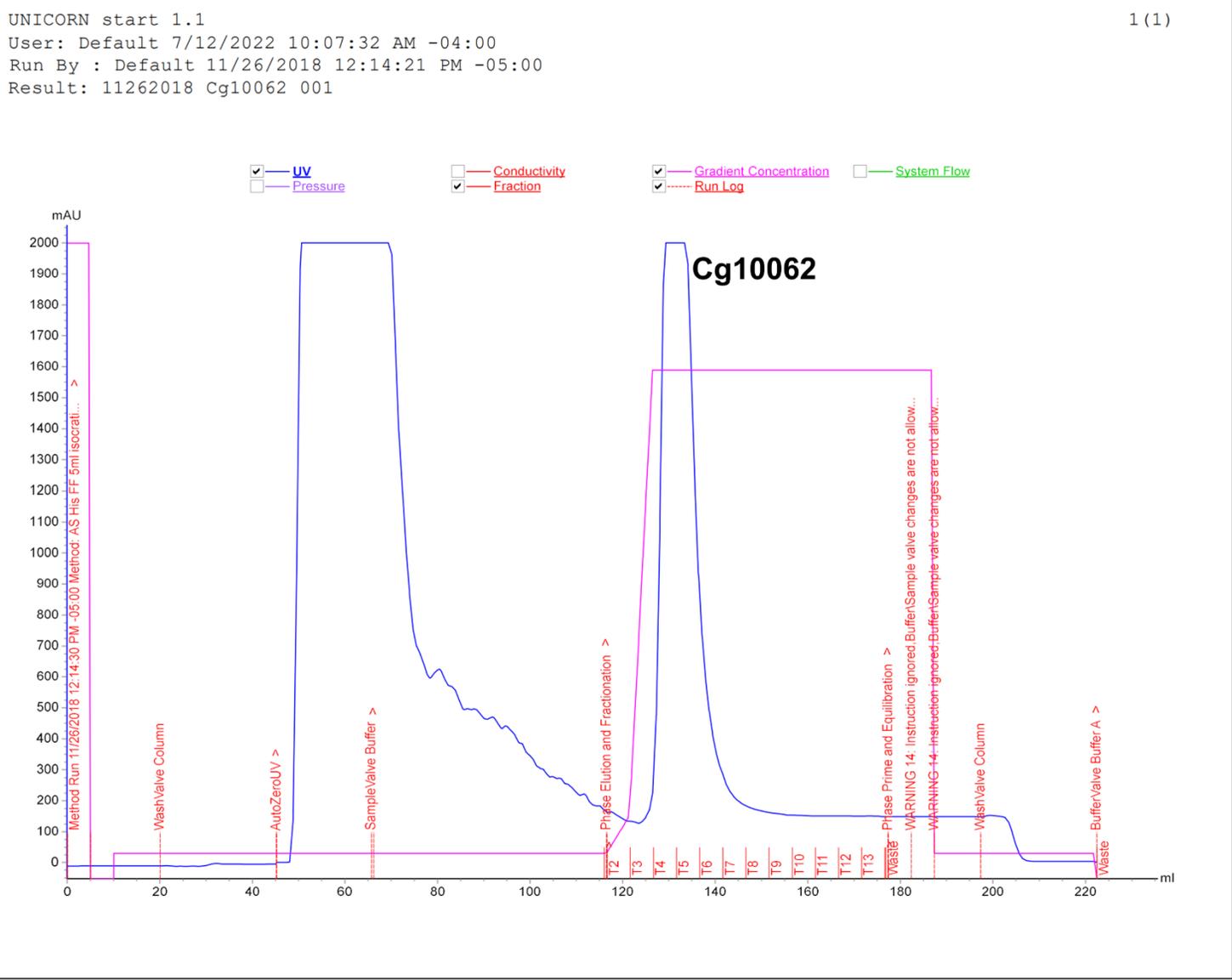


Figure 1.11. Purification of Cg10062 from a 1L culture of BL21(DE3)/pAS1.046 via affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

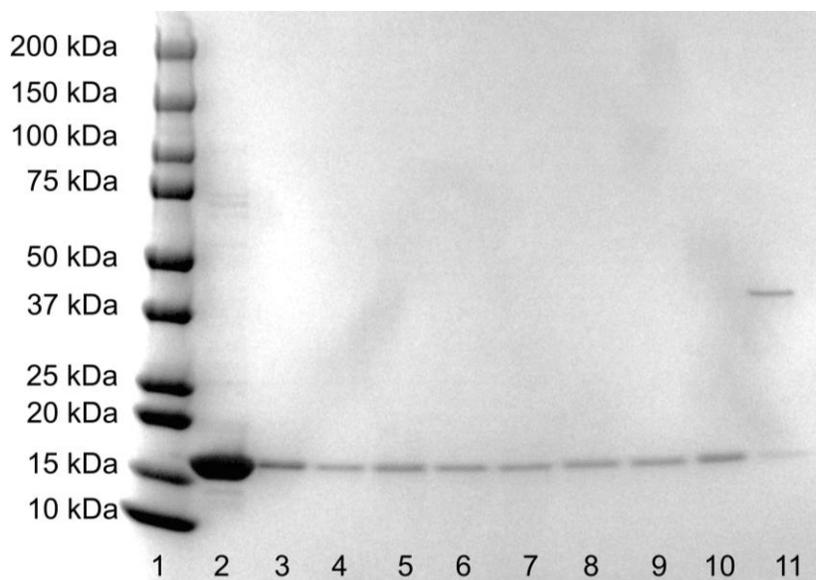


Figure 1.12. SDS-PAGE of Cg10062 (19 kDa) and variants. Lanes 1) 10 kDa ladder; 2) lysate of IPTG-induced BL21(DE3)/pAS1.046 expressing wild type Cg10062, followed by purified variants 3) E114Q; 4) E114D; 5) Y103F; 6) E114D-Y103F; 7) H28A; 8) R70A; 9) R73A; 10) E114A and wild-type Cg10062. Each lane contained a total of 2.5 μ g of protein. Wild type Cg10062 formed a dimer that are stable to SDS-PAGE. A dimer was not observed for other variants analyzed by SDS-PAGE.

1.6. Kinetic characterization of Cg10062 and variants

It was essential to have a uniform comparison of all variants examined while also using the same enzyme samples for characterization of product distribution and crystallization trials. It was also necessary to measure the ratio of products under the same conditions that kinetics parameters were determined. Measurement of product distribution using ^1H NMR requires significantly higher amounts of enzyme (0.1–0.25 mg) and higher substrate concentrations (50–100 mM) relative to the UV assay (0.005 – 0.05 mg enzyme, 2 mM ACA) for accuracy. Instead, purified enzymes were analyzed within an hour of protein purification and concentration, using a coupled enzyme assay described previously by Poelarends et al.⁵¹ The enzyme reactions were monitored by UV spectroscopy (340 nm) to determine the relative ratios of MSA **7** and acetaldehyde **8** produced by each Cg10062 variant (Figure 1.15). MSA **7** is subsequently decarboxylated by MSAD to form acetaldehyde **8**. The reduction of acetaldehyde **8** by alcohol dehydrogenase (ADH) from *Saccharomyces cerevisiae* (available commercially) was coupled to

the oxidation of NADH at 340 nm. The reaction rate in the presence of MSAD and ADH accounts for the combination of hydratase and hydratase/decarboxylase activity catalyzed by Cg10062 or its variant. The reaction rate in the absence of MSAD measure only the hydratase/decarboxylase activity. Excess quantities of coupling enzymes MSAD and ADH were included to ensure that Cg10062 or its variant was rate-limiting. Each assay contained the same concentration of ACA 2. The difference in the two rates was used to determine the ratio of MSA 7 to acetaldehyde 8 formed (Figure 1.16, Table 1.2). For the assays in this study, the gene encoding MSAD from *Coryneform* strain FG41⁵⁸ was modified to contain a TEV recognition sequence and cloned into pET-28a(+) using NdeI and XhoI restriction sites (Figure 1.13). MSAD was purified using BL21(DE3)/pAS2.031 using the same methods described above for Cg10062 and variants (Figure 1.14).

5' – CATATGCCCTTAATCCGTATAGATCTTACCAGTGATCGTTTCGAGAGAGCAACGGC
 GGGCGATTGCTGATGCAGTCCATGACGCTTTAGTAGAAGTTTTAGCGATTCCGGCTC
 GTGATCGCTTCCAGATACTGACTGCGCACGATCCCTCTGATATTATAGCCGAAGATG
 CTGGACTTGGCTTTCAGCGGTCCCCCAGTGTAGTCATCATAACGTCTTTACACAGG
 CAGGTAGAACTATTGAAACGAAACAGAGAGTATTTGCAGCGATAACAGAAAGTCTGG
 CTCCAATCGGTGTTGCAGGATCTGATGTTTTTATCGCAATCACCGAAAATGCACCCC
 ATGACTGGAGCTTTGGGTTTGGCAGTGCACAATATGTCACGGGTGAACTTGCGATTC
 CAGCCACTGGTGC GGCTGAGAACCTGTATTTCAAGGCCTCGAG – 3'

Figure 1.13. Codon-optimized nucleotide sequence of *msad*. The NdeI and XhoI restriction sites are highlighted in yellow, and the TEV protease recognition site (ENLYFQG) is highlighted in blue.

UNICORN start 1.1
User: Default 7/12/2022 10:10:31 AM -04:00
Run By : Default 8/19/2021 4:01:15 PM -04:00
Result: EL053 MSAD 001

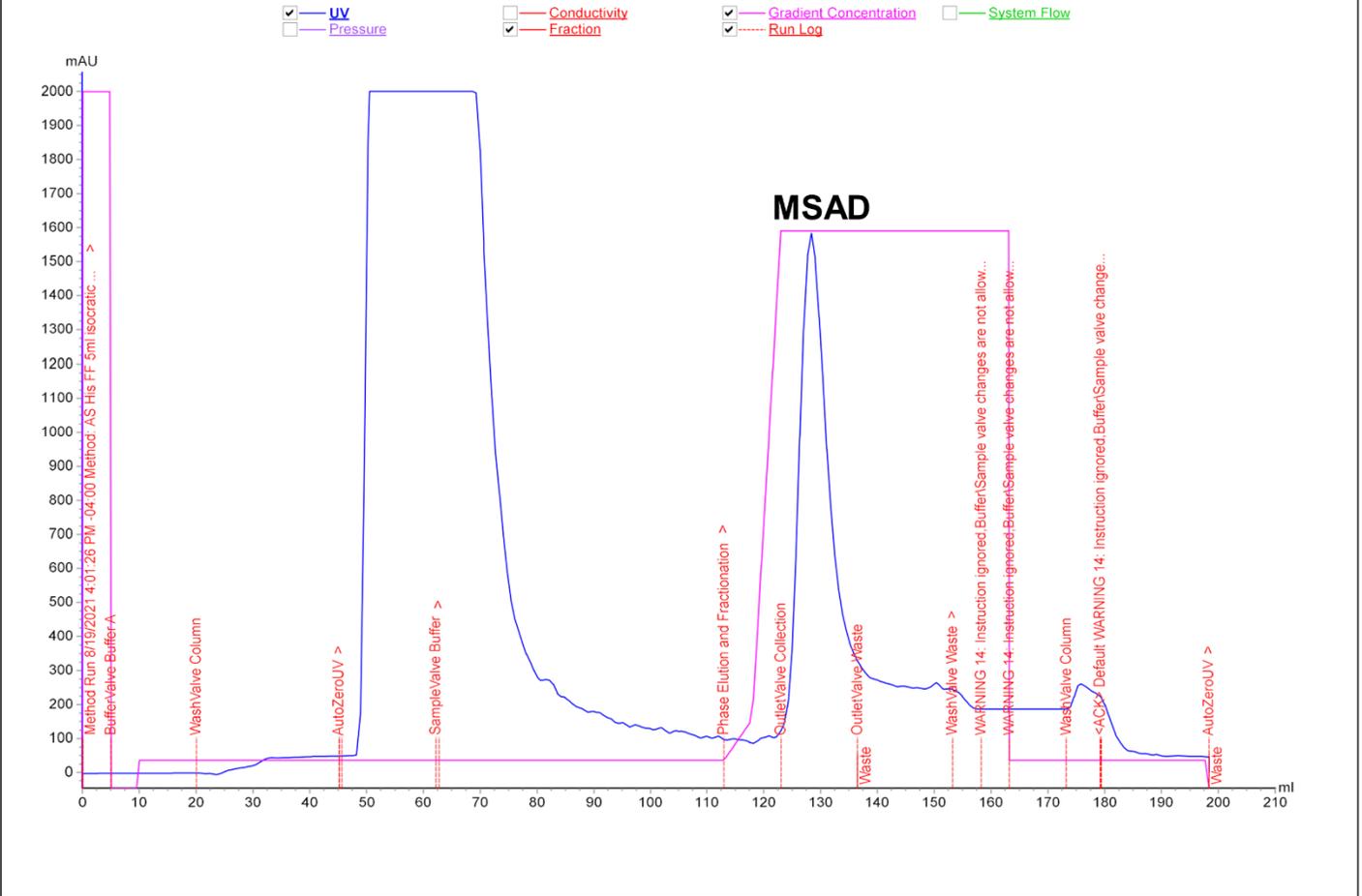


Figure 1.14. Purification of MSAD from a 1 L culture of BL21(DE3)/pAS2.031 via affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

To ensure an excess of coupling enzymes were present in each sample, the specific activity of MSAD was measured using the Cg10062 coupled enzyme assay⁵⁹ (Figure 1.15) with some modifications (Section 4.15.5). MSA 7 was generated from ACA *in situ* using an excess of Cg10062(E114N) (1.5 U) and was followed by the addition of MSAD (after a 30 s delay) and excess ADH (30 U). Since ADH was commercially available, the specific activity (300 U mg⁻¹) was already made available. The specific activity of MSAD was determined to be 25-30 U mg⁻¹.

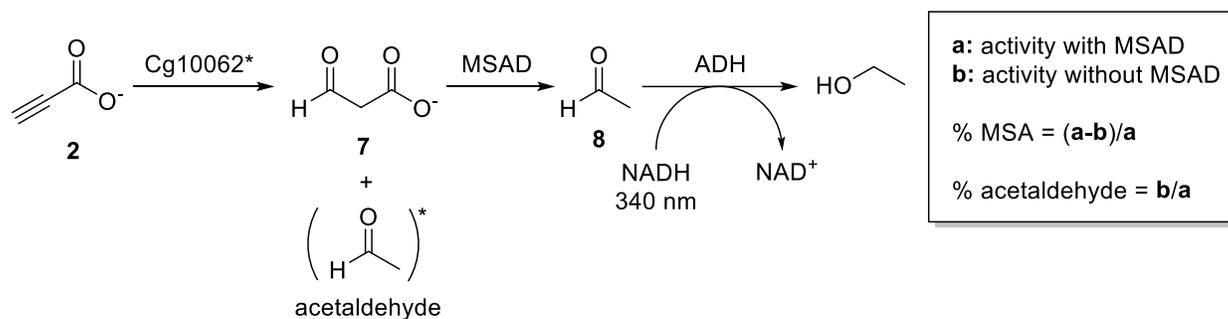


Figure 1.15. The coupled enzyme assay used to determine the ratio of products formed by Cg10062 and variants.

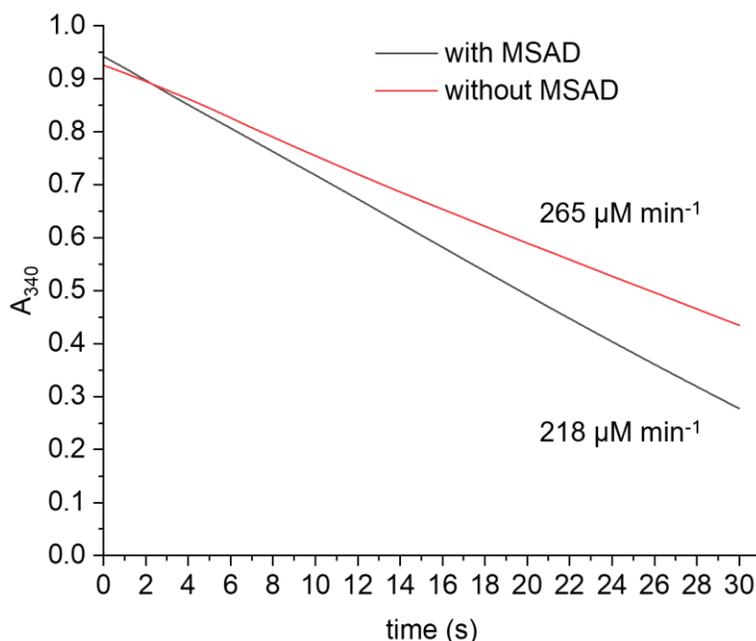


Figure 1.16. Cg10062 activity was determined using the Cg10062 coupled enzyme assay. Rates were measured with and without MSAD to determine the ratio of products formed.

The His₆-tag Cg10062 was removed by TEV protease to determine the difference in enzyme activity (if any) in the absence of the C-terminal His₆-tag. His₆-tagged Cg10062 was purified using affinity chromatography from cell lysates of BL21(DE3)/pAS1.046. The purified His₆-tagged Cg10062 was concentrated in 50 mM sodium phosphate, pH 8, with 0.5 mM EDTA and 1 mM DTT and incubated with His₇-tagged TEV protease (His₆-tagged Cg10062: TEV molar ratio = 1:50) at 4 °C for 12 h. The mixture of proteins was re-purified by nickel affinity chromatography to isolate the Cg10062. The activity of the His₆-tagged Cg10062 was compared to Cg10062 treated by TEV protease using the Cg10062 coupled enzyme assay (Figure 1.17). Upon removal of the His₆-tag, a significant change in enzyme activity was not observed with Cg10062 using the coupled enzyme assay. Therefore, the C-terminal His₆-tag of Cg10062 and its variants were not removed using TEV protease.

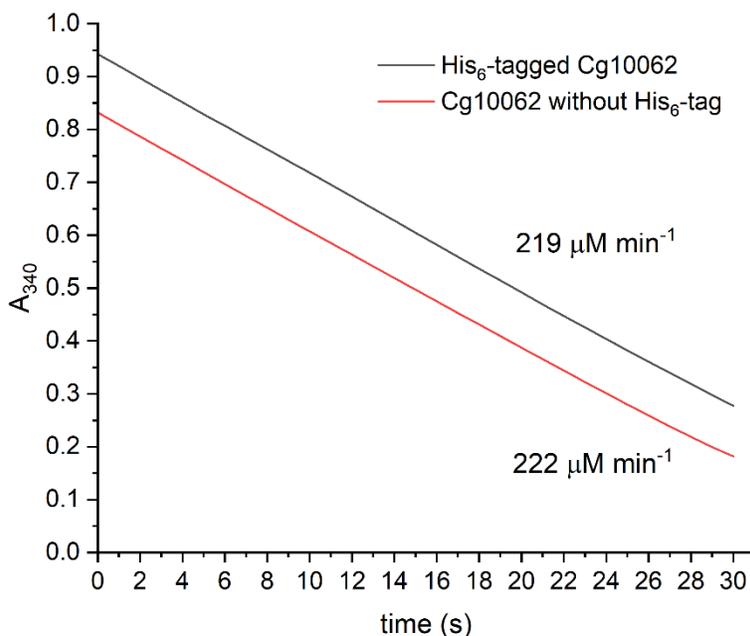


Figure 1.17. Comparison of Cg10062 activity before and after removal of His₆-tag

Table 1.2. Product distribution of Cg10062 variants using the coupled enzyme assay

Cg10062 variant	Cg10062 variant concentration (mg mL ⁻¹)	Rate of NADH loss ($\mu\text{M min}^{-1}$)		Product Ratio (%)	
		+ MSAD	-MSAD	13	14
Wild type	0.025	265	218	18	82
E114N	0.025	139	ND ^a	>99	<1
E114Q	0.10	185	ND ^a	>99	<1
E114D	0.10	164	ND ^a	>99	<1
Y103F	0.15	118	78	34	66
E114D-Y103F	0.10	71	ND ^a	>99	<1
H28A	0.25	22	8	64	36
E114A	1.50	27	20	26	74
Y103A	0.25	69	29	58	42
R70A	1.50	ND ^a	ND ^a	-	-
R73A	1.50	ND ^a	ND ^a	-	-
R70K	1.50	ND ^a	ND ^a	-	-
R73K	1.50	ND ^a	ND ^a	-	-
E114S	1.50	ND ^a	ND ^a	-	-

^aND: not detected

As expected, the mutagenesis of the active site residues resulted in variants with lower efficiency relative to wild type. To compensate for reduced activity while enabling adequate catalysis for accurate measurement, increased amounts of some Cg10062 variants were required (Table 1.2). Thus, the rates of NADH loss from this experiment were not used for calculation of specific activities. Product ratios calculated using this method are consistent with previously published data⁵⁹ for wild-type Cg10062 and showed a mixture of MSA **7** and acetaldehyde **8**, with the latter as the major product. In previous studies⁵⁹ and in this study, alanine scanning mutagenesis of active site residues His-28, Tyr-103, and Glu-114, resulting in variants H28A, Y103A, and E114A, showed diminished overall activity (Table 1.2) and reduction of

hydratase/decarboxylase activity relative to hydratase activity (Figure 1.18). Mutagenesis of Arg-70 and Arg-73 afforded variants R70A and R73A, which showed a complete loss of detectable activity. Taken together, these variants confirmed that these five residues played critical roles in Cg10062 activity.

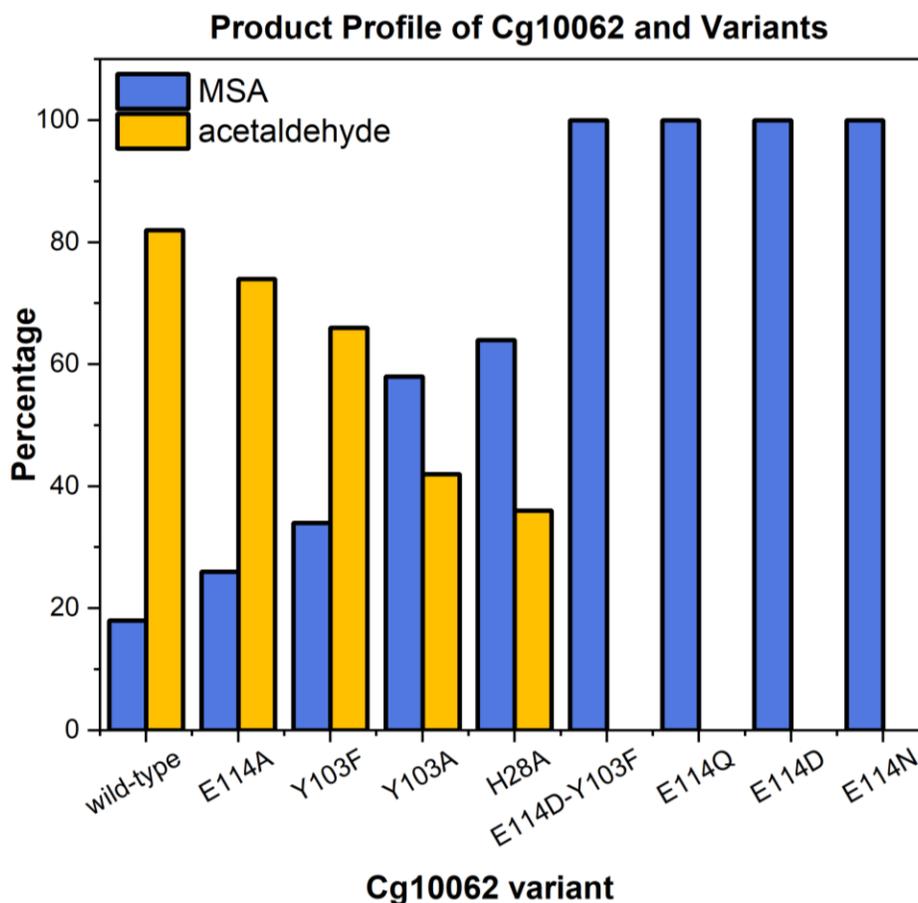


Figure 1.18. The relative ratios of MSA and acetaldehyde produced by Cg10062 and variants.

Previous accounts established the requirement of Pro-1 for catalysis; thus, variant P1A was not examined in this study.⁵¹ Interestingly, mutation of Glu-114 to a glutamine (E114Q) or an aspartate (E114D) completely abrogates decarboxylation, resulting in MSA 7 as essentially the sole product of these variants. Notably, both E114Q and E114D variants had lower hydratase activity relative to the combined hydratase and hydratase/decarboxylase activities of wild-type Cg10062. Further investigation of the Glu-114 residue prompted the generation of E114S and

E114N variants. Cg10062(E114N) was identified as a novel hydratase-only variant that has not been previously explored (Figure 1.18). Variants R70K, R73K and E114S did not display any activity detectable by the coupled enzyme assay and were not characterized further. In agreement with previous reports, substitution of Tyr-103 with phenylalanine resulted in a slight increase in MSA **7** (34 %) formation relative to acetaldehyde **8** (66%) in comparison with wild type Cg10062.⁵⁹ The combination of E114D-Y103F mutations in Cg10062 resulted in a new hydratase-only mutant that is less active than the previously reported E114D and E114Q and the newly-discovered E114N variant. Michaelis-Menten kinetic parameters were determined for wild type Cg10062, E114Q, E114D and Y103F mutants, as well as the newly constructed E114N and E114D-Y103F double mutant with ACA **2** using the coupled enzyme assay described above (Table 1.3, Figures 1.19–1.24).

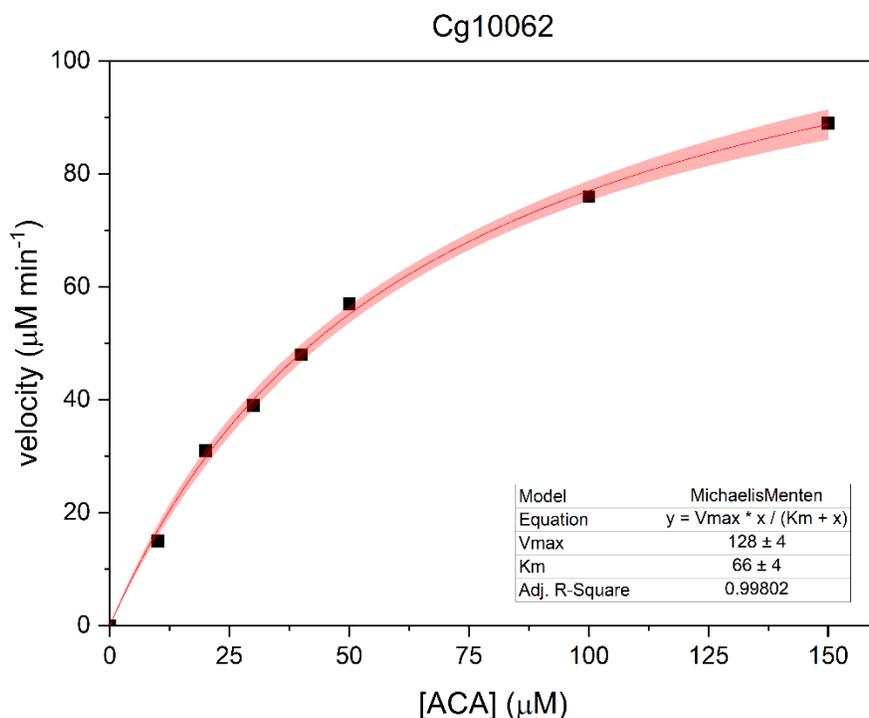


Figure 1.19. Michaelis Menten kinetics of Cg10062. The assay contained Cg10062 (0.005 mg mL⁻¹) in 100 mM sodium phosphate, pH 8. The red band represents the 95% confidence interval.

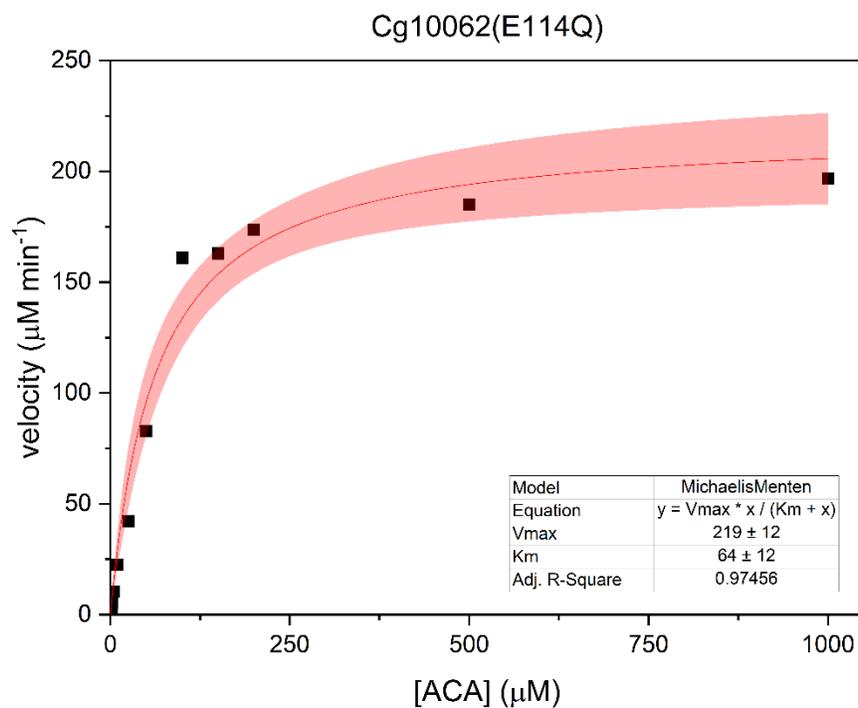


Figure 1.20. Michaelis Menten kinetics of Cg10062(E114Q). The assay contained Cg10062(E114Q) (0.1 mg mL⁻¹) in 100 mM sodium phosphate, pH 8. The red band represents the 95% confidence interval.

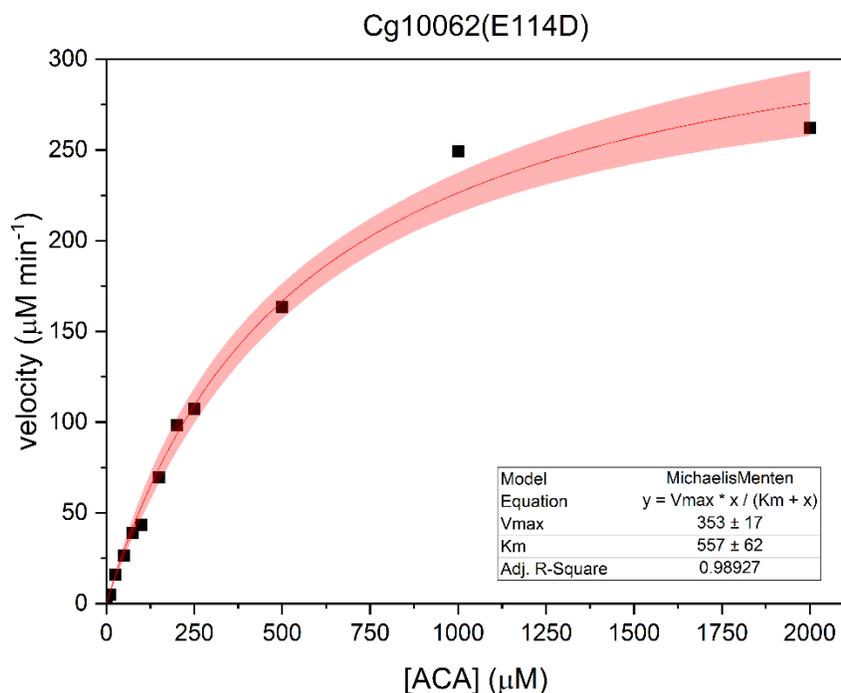


Figure 1.21. Michaelis Menten kinetics of Cg10062(E114D). The assay contained Cg10062(E114D) (0.1 mg mL^{-1}) in 100 mM sodium phosphate, pH 8. The red band represents the 95% confidence interval.

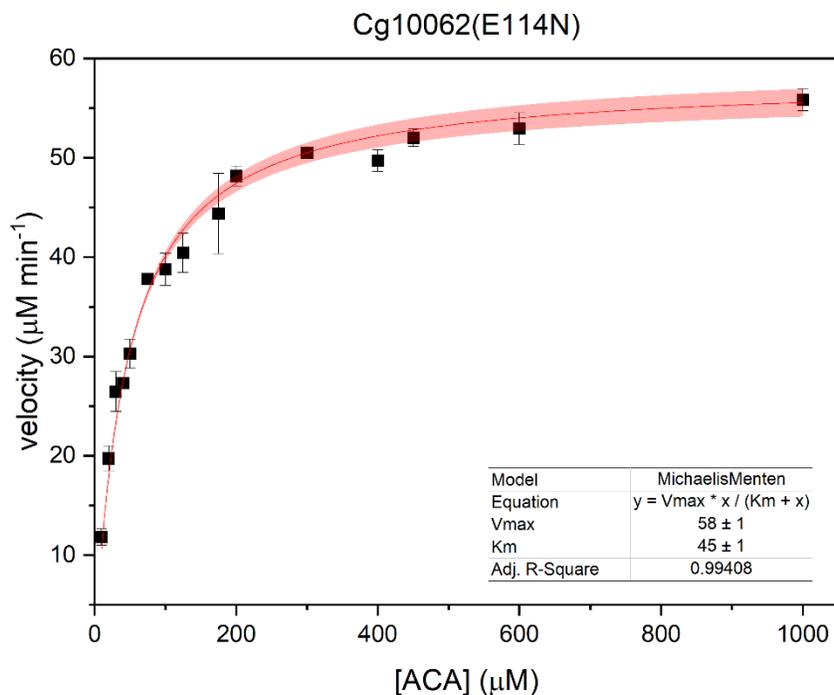


Figure 1.22. Michaelis Menten kinetics of Cg10062(E114N). The assay contained Cg10062(E114N) (0.01 mg mL^{-1}) in 100 mM sodium phosphate, pH 8. The red band represents the 95% confidence interval.

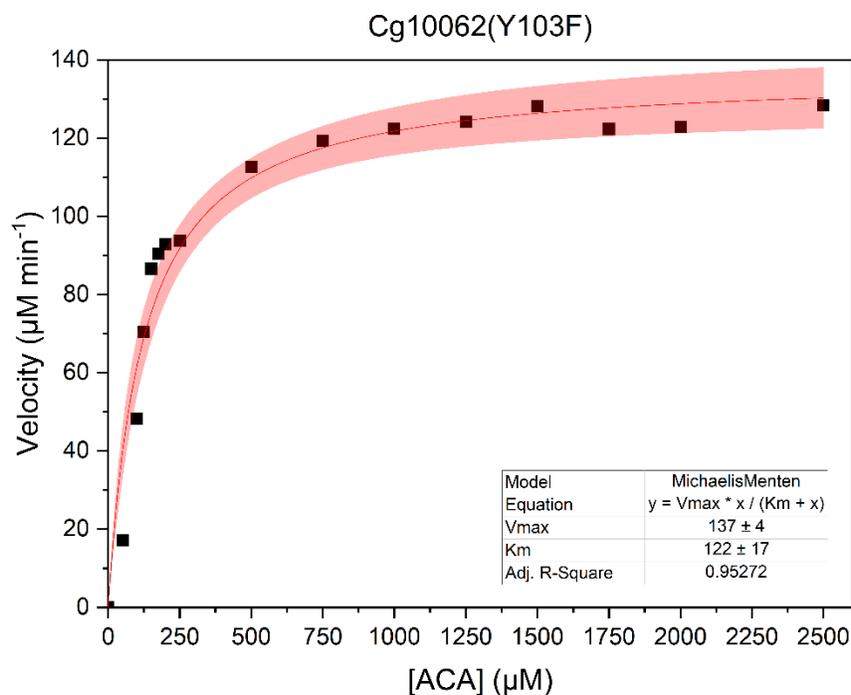


Figure 1.23. Michaelis Menten kinetics of Cg10062(Y103F). The assay contained Cg10062(Y103F) (0.1 mg mL^{-1}) in 100 mM sodium phosphate, pH 8. The red band represents the 95% confidence interval.

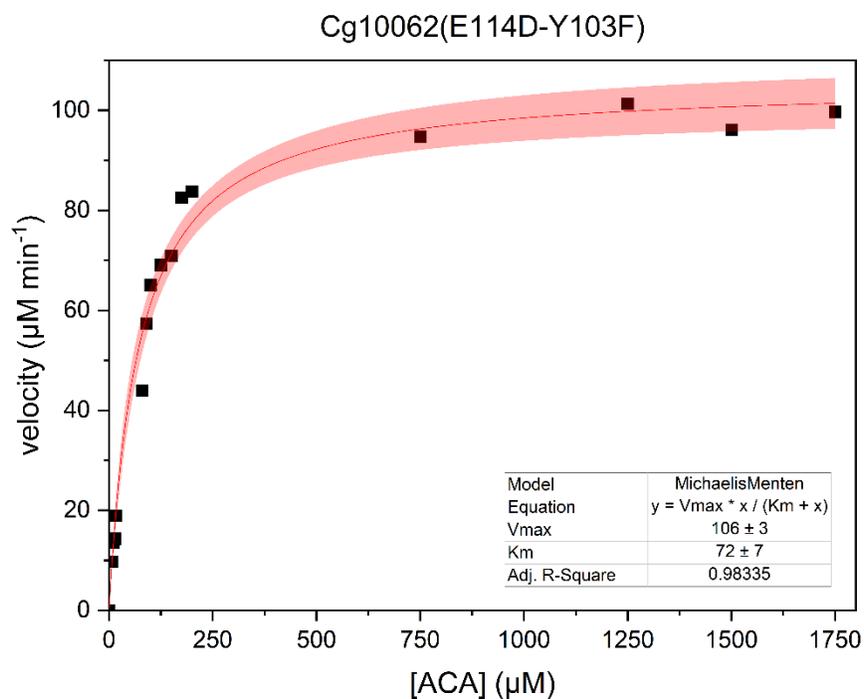


Figure 1.24. Michaelis Menten kinetics of Cg10062(E114D-Y103F). The assay contained Cg10062(E114D-Y103F) (0.1 mg mL^{-1}) in 100 mM sodium phosphate, pH 8. The red band represents the 95% confidence interval.

Table 1.3. Kinetic parameters of Cg10062 and variants.

variant	$K_m(\mu\text{M})$	$k_{cat} (\text{s}^{-1})$	$k_{cat}/K_m \times 10^4 (\text{M}^{-1} \text{s}^{-1})$
wild type Cg10062	66 ± 4	8.21 ± 0.26	12.4 ± 0.8
Cg10062(E114N)	45 ± 1	1.84 ± 0.03	4.08 ± 0.11
Cg10062(E114D)	557 ± 62	1.12 ± 0.05	0.20 ± 0.02
Cg10062(E114Q)	64 ± 12	0.69 ± 0.04	1.08 ± 0.21
Cg10062(Y103F)	122 ± 17	0.43 ± 0.01	0.36 ± 0.05
Cg10062(E114D-Y103F)	72 ± 7	0.34 ± 0.01	0.47 ± 0.05

Steady-state kinetics were measured by incubation of varying concentrations of ACA with Cg10062 or variant using a coupled enzyme assay containing 1.2 U MSAD, 12 U Adh, and 0.3 mM NADH in 100 mM sodium phosphate (pH 8.0) at 25 °C. The reduction of acetaldehyde was coupled to the oxidation of NADH (340 nm). All assays were carried out in triplicate.

Wild type Cg10062 had the highest k_{cat} of 8.21 s⁻¹. Although E114D-Y103F was confirmed to be a hydratase-only variant of Cg10062, the k_{cat} of this variant (0.34 s⁻¹) was lower than the previously discovered⁵⁹ hydratase-only E114Q (0.69 s⁻¹) and E114D (1.12 s⁻¹) mutants, respectively. In contrast, the novel variant Cg10062(E114N) discovered in this study displayed higher activity relative to the E114Q and E114D counterparts with a k_{cat} of 1.84 s⁻¹. The overall catalytic efficiencies of E114Q and E114D were 4-fold and 20-fold lower relative to E114N (k_{cat}/K_m 4.08 x 10⁴ M⁻¹ s⁻¹), respectively. The discovery of the more active E114N variant allows downstream conversion of MSA 7 produced Cg10062(E114N)-catalyzed hydration of ACA 2 into value-added C3 chemicals, as described earlier in this study. While this variant partially satisfies the goal of evolving Cg10062 to maximize carbon flux from ACA (by preventing decarboxylation) into value-added chemicals, efforts to engineer Cg10062 further for higher catalytic efficiency while maintaining hydratase-only activity are still underway (see chapter three). Steady-state kinetic parameters for previously described E114A and Y103A⁵¹ and newly constructed H28A were not determined due to low activity. Of all enzymes characterized in this study, wild-type Cg10062 remained superior in terms of overall catalytic efficiency. Note that k_{cat} values determined for the Cg10062 variants are in close agreement with previously reported values;⁵⁹

however, the catalytic efficiencies reported here are significantly lower. These differences are due to higher K_m values determined in this study. All variants of Cg10062 characterized in this study include a C-terminal His₆-tag to enable purification via affinity chromatography. Removal of the His₆-tag did not result in a change in catalytic activity in Cg10062 (data not shown). While inclusion of the His₆-tag is not expected to alter K_m values, comparison of proteins purified by affinity chromatography with those purified via multiple columns and precipitation^{51,55} may result in subtle differences.

1.7. ¹H NMR characterization of Cg10062 and variants

In previous accounts, the product ratios resulting from catalysis of Cg10062 and variants on ACA **2** were determined using ¹H NMR.⁵⁹ As described previously, given the need for high concentrations of enzyme and substrate, the coupled enzyme assay was deemed a better alternative for determining product distribution of the different Cg10062 variants. Furthermore, the MSA **7** intermediate is unstable and undergoes spontaneous decarboxylation.^{59,64} Thus, over the course of the NMR experiments, it can be challenging to distinguish between acetaldehyde **8** formed as a result of Cg10062 catalysis versus that formed from the spontaneous decarboxylation of MSA **7**. As a result, in this study, ¹H NMR analysis was used only as a second line of evidence for identification of products formed by Cg10062 and the key hydratase-only variants, E114N, E114Q and E114D. ACA (111 mM) was incubated with identical concentrations of each Cg10062 variant (0.27 mg mL⁻¹) in 100 mM sodium phosphate, pH 8. Samples were quenched with acid immediately after addition of enzyme and 1 h after the reaction progressed (room temperature). Since the assays were carried out in aqueous buffer, and due to the unstable nature of MSA **7**, samples were analyzed by NMR immediately after quenching by (H)wet1D for solvent suppression. All samples included DMSO-*d*₆ (δ 2.49) to provide a lock signal and TSP (3-(trimethylsilyl)propionate-2,2,3,3-*d*₄, sodium salt) (δ -0.21 (s, 9H)) to provide an internal standard. The resonances at δ 3.5 (s, 4H) corresponds to the cryoprotectant ethylene glycol. The resonance at δ 2.91 (s, 1H) corresponds to ACA. Resonances at δ 9.50 (t, 1H), δ 3.20 (d, 2H), and δ 5.13 (t,

1H), δ 2.30 (d, 2H) correspond to MSA **7** and its hydrate, respectively. Resonances at δ 9.47 (q, 1H), δ 2.03 (d, 3H), and δ 5.05 (q, 1H), δ 1.12 (d, 3H) correspond to acetaldehyde and its hydrate, respectively.

Consistent with kinetic characterization, Cg10062 catalysis with ACA **2** was significantly faster relative to the E114 variants studied. Spectra obtained immediately after the addition of wild type Cg10062 indicated the formation of more acetaldehyde **8** relative to MSA **7** (Figure 1.25) and complete consumption of ACA **2** was observed after 1 h of incubation (Figure 1.26). A majority of acetaldehyde observed (with most being in the hydrate form) is consistent with previous accounts⁵⁹ and data from the coupled enzyme assay in this study. On the other hand, variants E114Q (Figure 1.27–1.28) and E114D (Figure 1.29–1.30) were both slower relative to wild type as no product formation was observed initially ($t = 0$ h). In both cases, MSA **7** and its hydrate were observed as the major product after 1 h (Figures 1.28 and 1.30). Trace acetaldehyde **8** observed could correspond to non-enzyme catalyzed decarboxylation of MSA **7**. Similar results were observed for the novel hydratase-only variant, Cg10062(E114N) (Figure 1.31–1.32). The results from the coupled enzyme assay for Cg10062 and variants, described above, confirm that the hydratase-only variants E114Q, E114D, E114D-Y103F and E114N were all selective for MSA **7** formation. Unfortunately, it was not possible to measure MSA **7** decarboxylation in the absence of enzyme due to the challenge of preparing and isolating MSA **7** using non-enzymatic methods.

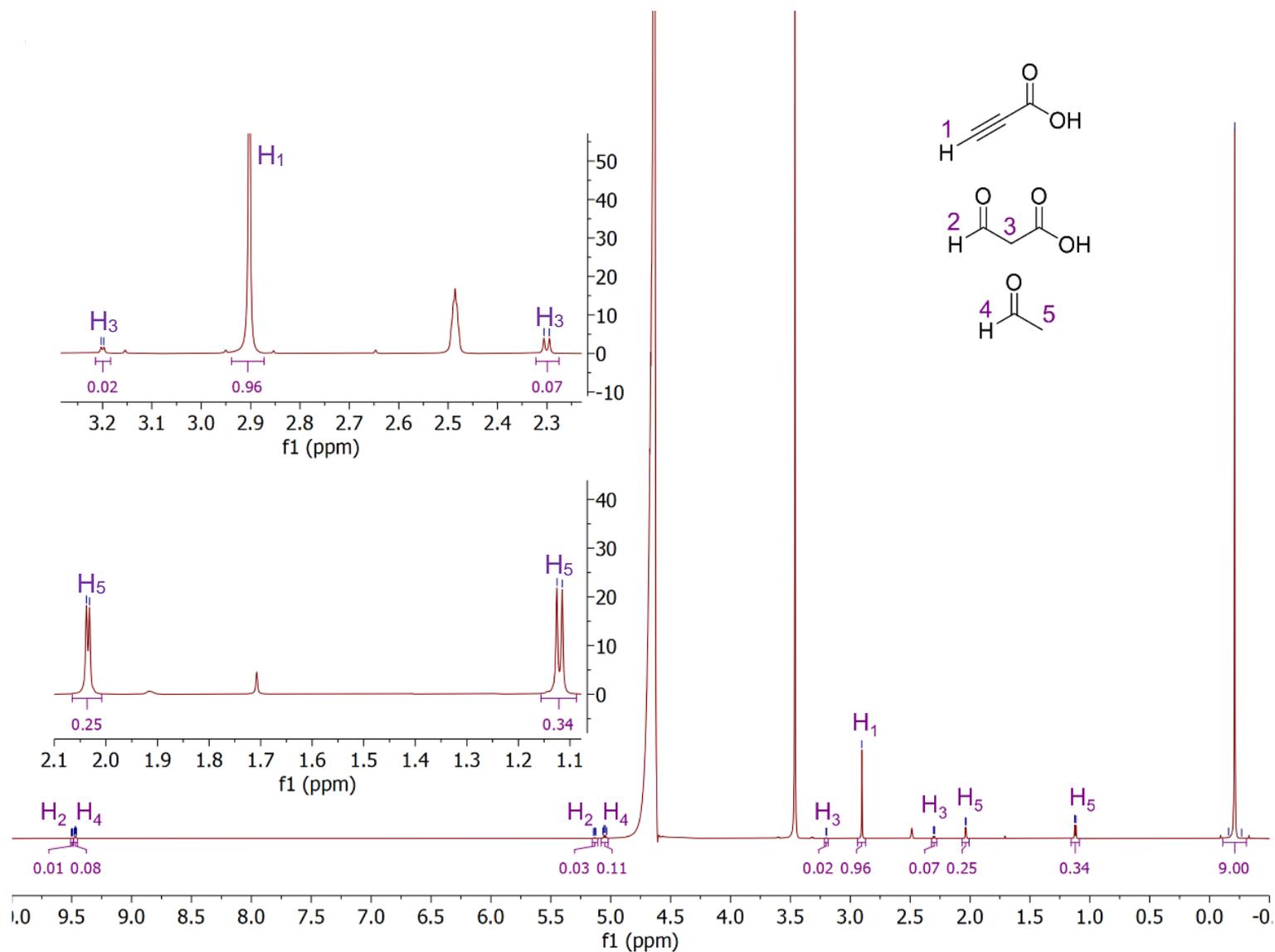


Figure 1.25. ^1H NMR of reaction of Cg10062 (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 0 \text{ h}$. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

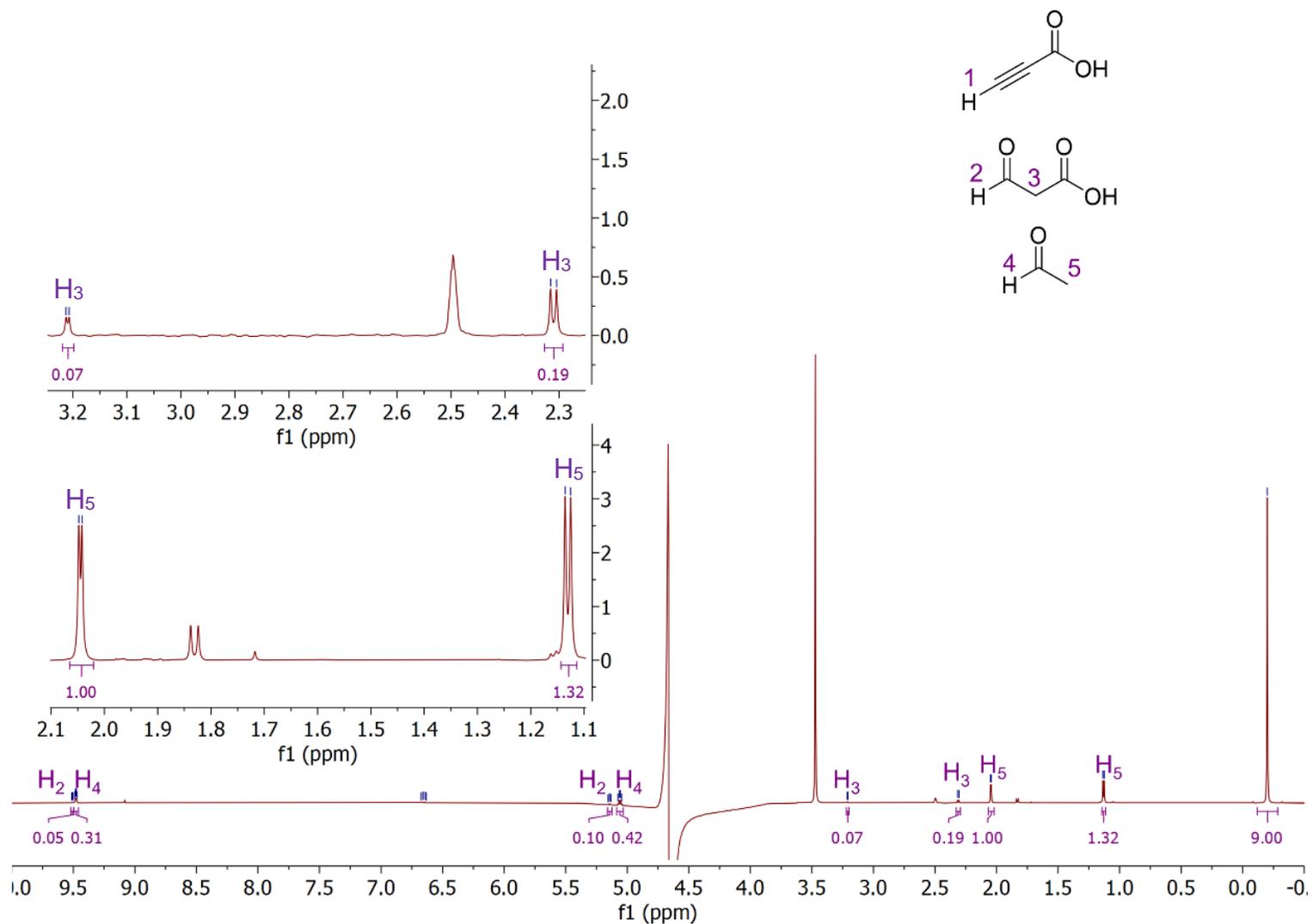


Figure 1.26. ^1H NMR of reaction of Cg10062 (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 1 \text{ h}$. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

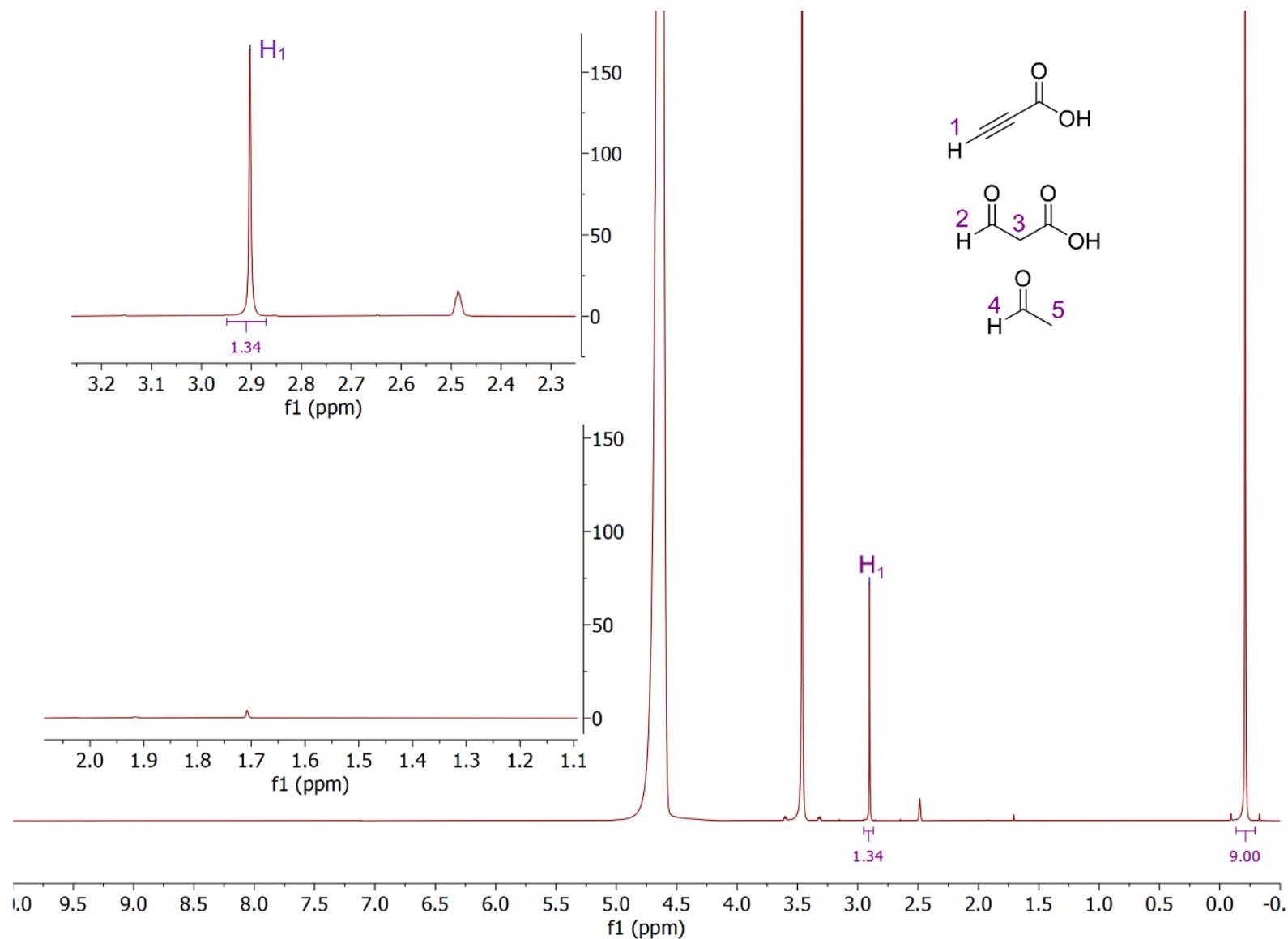


Figure 1.27. ^1H NMR of reaction of Cg10062(E114Q) (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 0 \text{ h}$. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

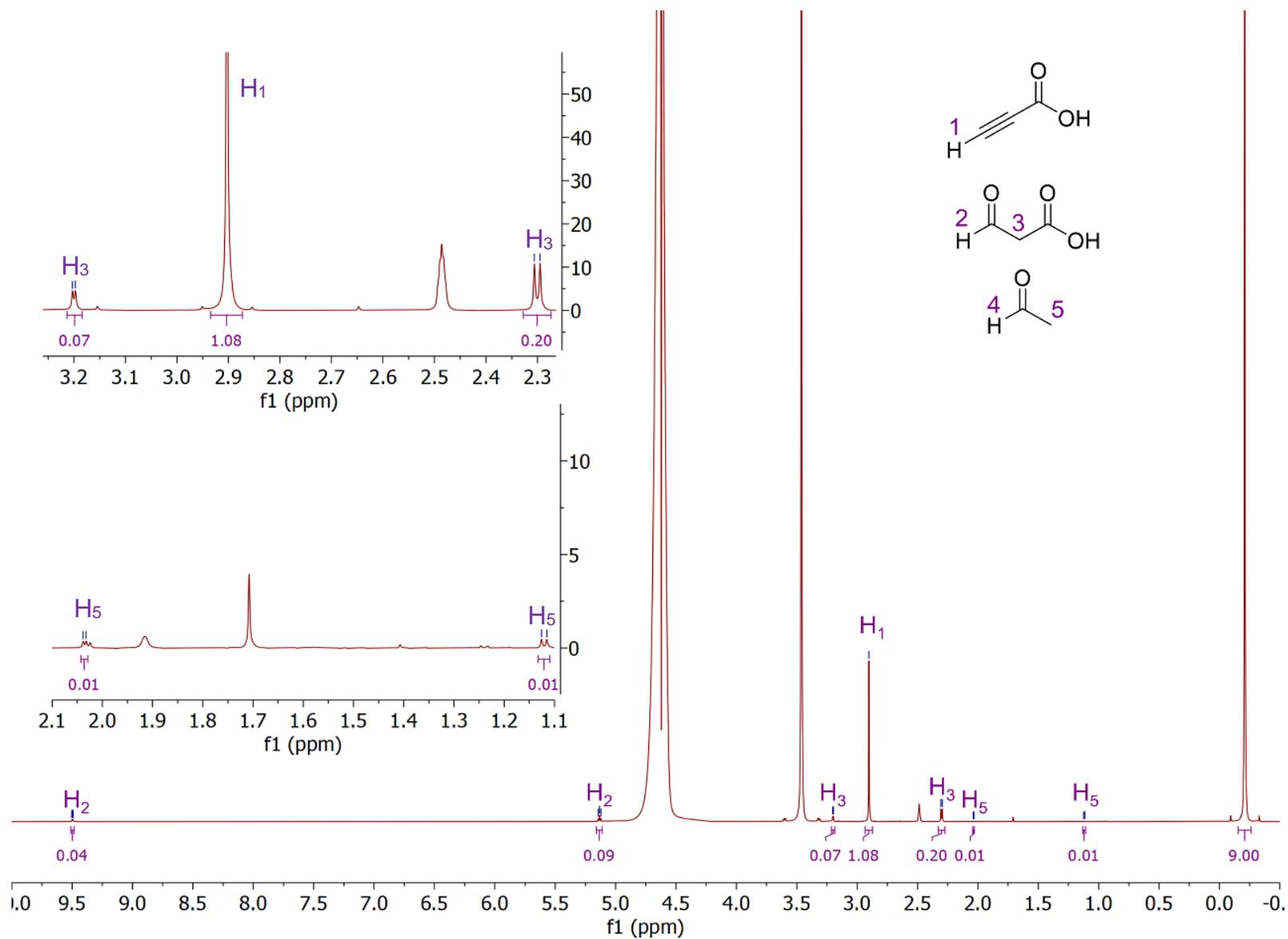


Figure 1.28. ^1H NMR of reaction of Cg10062(E114Q) (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 1 \text{ h}$. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

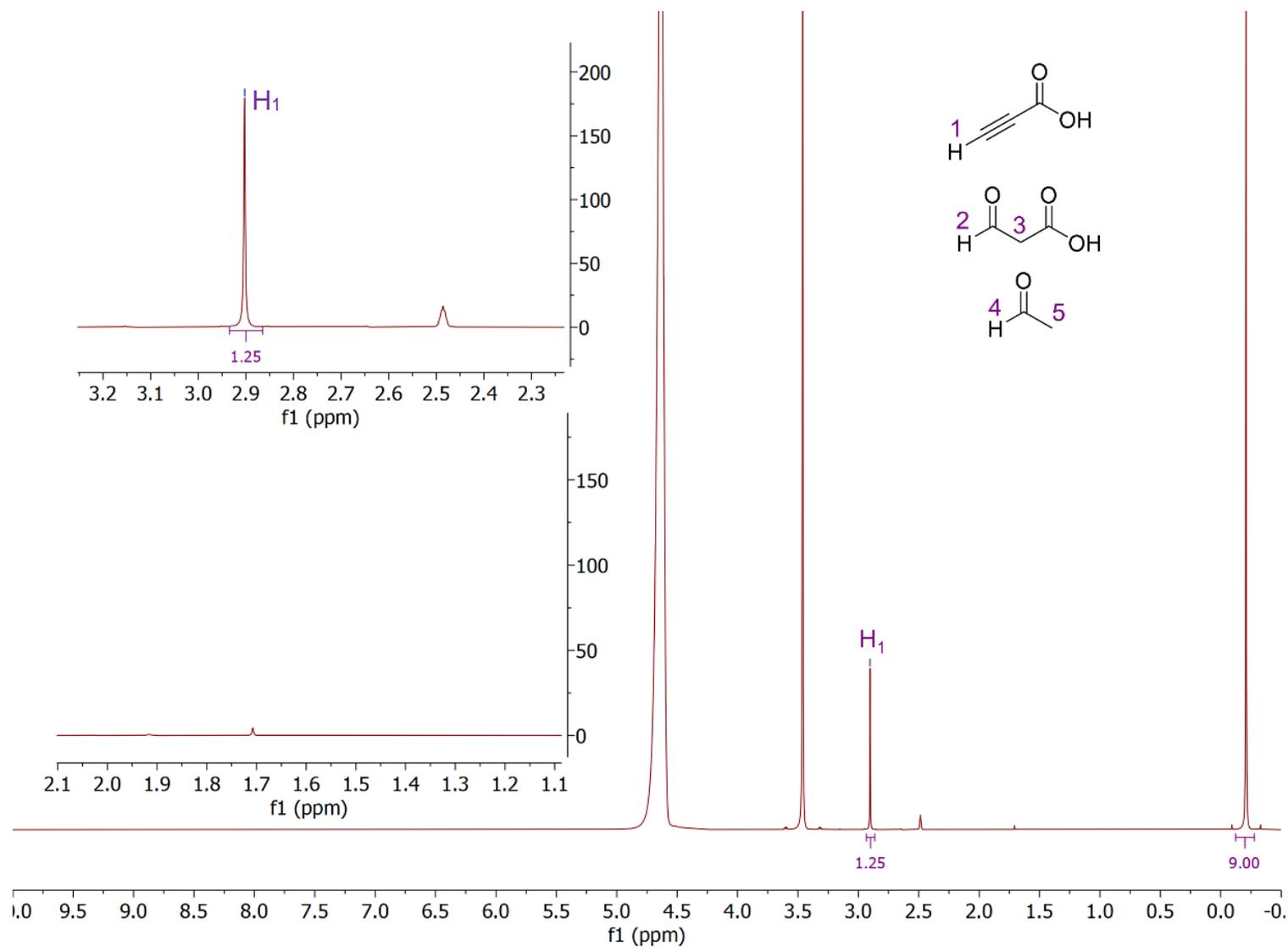


Figure 1.29. ^1H NMR of reaction of Cg10062(E114D) (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 0$ h. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

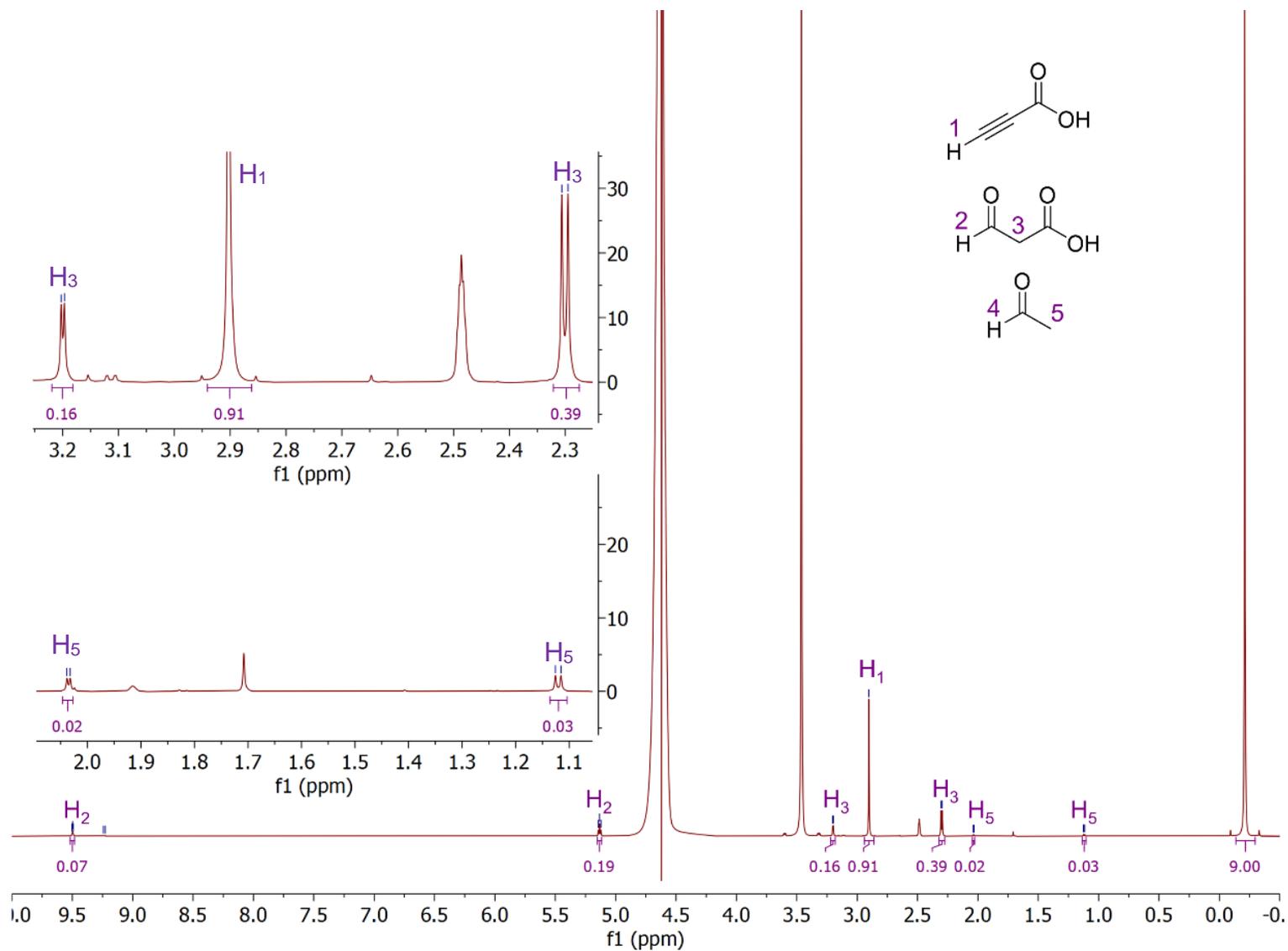


Figure 1.30. ^1H NMR of reaction of Cg10062(E114D) (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 1 \text{ h}$. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

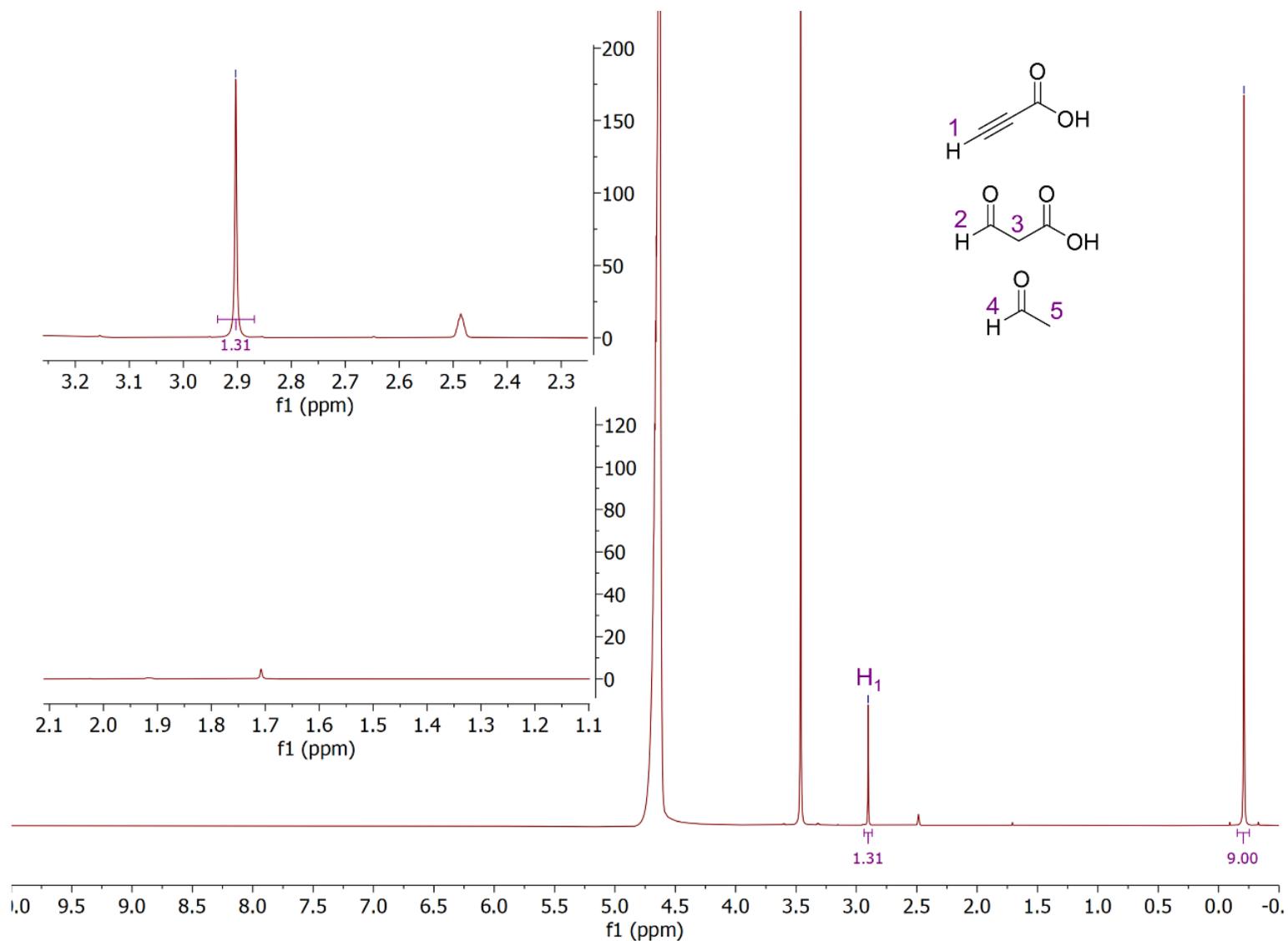


Figure 1.31. ^1H NMR of reaction of Cg10062(E114N) (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 0$ h. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

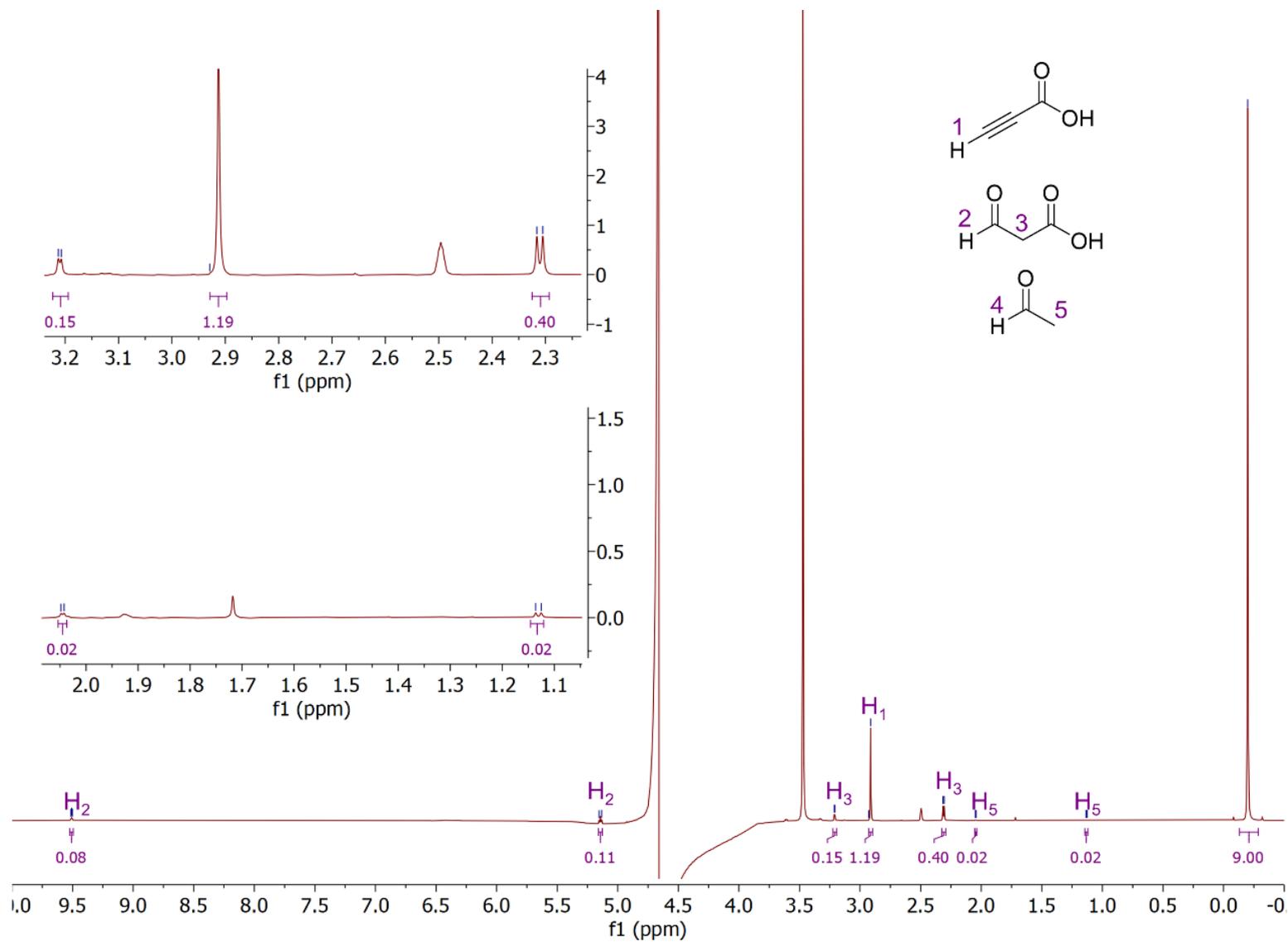


Figure 1.32. ^1H NMR of reaction of Cg10062(E114N) (0.27 mg mL^{-1}) incubated with ACA (111 mM) at $t = 1\text{ h}$. The panels of the left are zoomed in regions indicating the ACA, MSA and acetaldehyde signals.

1.8. Structural analysis of Cg10062 and variants

X-ray crystallography was carried out in collaboration with Prof. James Geiger (Department of Chemistry, MSU). Dr. Hadi Nayebi (Geiger group) and Katelyn Silva (Draths/Geiger groups) carried out preparation of protein crystals, processing of crystallographic diffraction data and modelling. This data has been included in the dissertation to facilitate a more complete understanding of proposed mechanism of Cg10062.

To better understand the enzyme mechanism at atomic resolution, crystals of wild-type Cg10062 and variants (E114N, E114D, H28A, R73A, and Y103F) were grown.¹ All enzymes crystallized contained the TEV site and His₆-tag on the C-terminus. In all cases, the crystals were soaked in ACA 2 in an attempt to capture substrate, products or intermediate in the active site. It was envisioned that the enzymes might capture different intermediates of the reaction and better illuminate the role of each active site residue. Crystallographic information and conditions used for obtaining structures of Cg10062 (PDB ID: 7MS0), E114D (PDB ID: 7MS9), H28A (PDB ID: 7MS1), Y103F (PDB ID: 7MS8), and R73A (PDB ID: 7MS3) generated in this study are available elsewhere.¹ Several attempts were made to obtain apo and soaked crystals of the E114Q variant, but the resulting crystals were unstable and did not produce high quality diffraction data. Crystals of wild type Cg10062 soaked with ACA 2 were unstable, perhaps due to rapid substrate turnover in the crystals. Thus, only an apo structure (1.37 Å resolution) was obtained for wild type Cg10062 (Figure 1.33), which was similar to the lower resolution structure (PDB ID: 3N4G) reported previously.⁶⁵ For Cg10062 and E114D, the C-terminal residues from Leu-119 to the end of the chain were disordered (Figure 1.34). However, the C-termini of E114N, H28A, R73A, and Y103F were ordered, similar to that seen previously in wild type Cg10062.⁶⁵

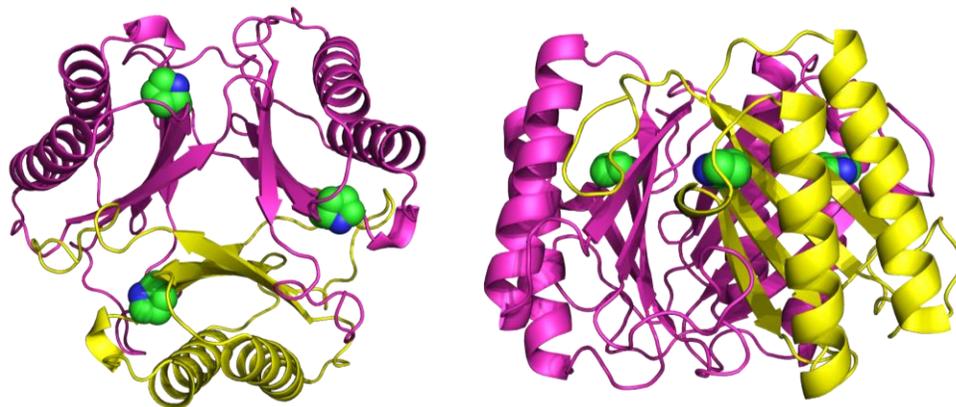


Figure 1.33. Structure of apo-Cg10062 (PDB 7MS0). Single chain in the ASU (yellow), symmetry mates to complete the trimer (pink) and active site Pro-1(green).

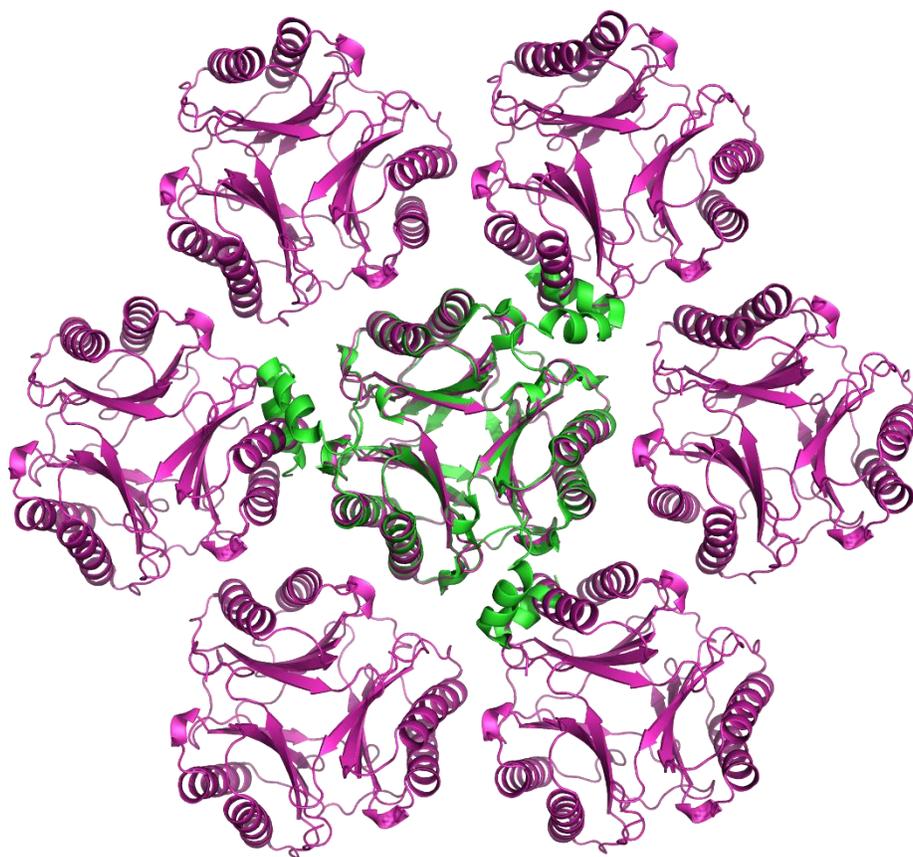


Figure 1.34. Overlay of wild type Cg10062 structures. Structure from PDB (3N4G) (green) and wild type Cg10062 (PDB ID: 7MS0) obtained in this study (pink). Unlike 3N4G, the C-termini of 7MS0 in this study were not ordered since the crystal packing prevents proper folding of the C-termini.

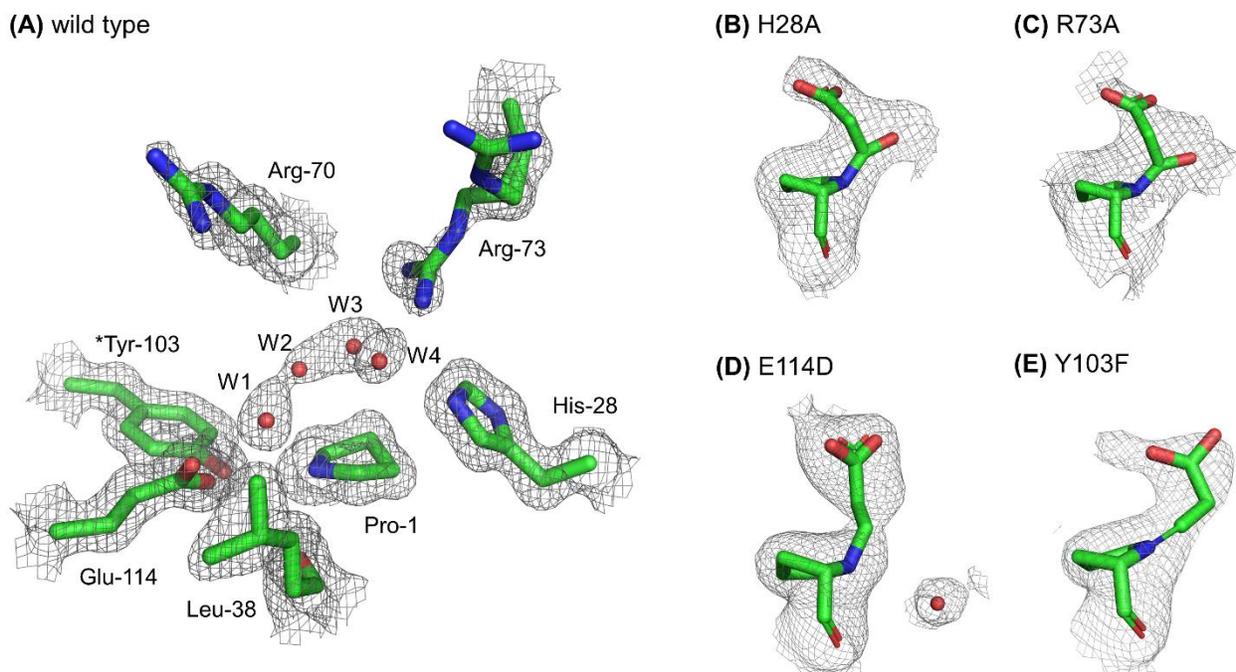


Figure 1.35. The active site of wild type Cg10062 and covalently-bound intermediates captured in Cg10062 variants. (A) Ordered water molecules in the active site of wild type Cg10062 (PDB ID: 7MS0); (B) H28A active site modeled with 3-(*N*-prolyl)3-hydroxypropionate (PDB ID: 7MS1); (C) R73A active site modeled with 3-(*N*-prolyl)-3-hydroxypropionate (PDB ID: 7MS3); (D) E114D active site (chain A) modeled with 3-(*N*-prolyl)-acrylate (enamine) (PDB ID: 7MS9); and (E) Y103F active site modeled with 3-(*N*-prolyl)-acrylate (iminium) (PDB ID: 7MS8). Atoms colored by type: C – green; N – blue; O – red. The colors of all atoms remain the same throughout, except for carbon atoms. *Tyr-103 is from the neighboring monomer.

In *cis*-CaaD, Glu-114 and Tyr-103 have been identified as active site residues involved in water activation.⁵⁵ The active site of wild type Cg10062 contains four ordered water molecules (Figure 1.35A), one (W1) of which is hydrogen bonded to Glu-114 and Pro-1 and interacts with Tyr-103 via the Glu-114 residue. The other three water molecules are arrayed over Pro-1, where the substrate is presumed to bind. Each of the active sites of the five substrate-soaked Cg10062 variants (H28A, R73A, E114D, E114N and Y103F) showed electron density consistent with an intermediate covalently linked to the nitrogen of Pro-1. All panels are oriented similarly to depict the different shapes of the active site electron densities in each variant (Figure 1.35). In a previously proposed mechanism, His-28 and Arg-73 residues were hypothesized to play a role in substrate-binding.⁵⁹ Additionally, the presence of the positively charged Arg-70 and Arg-73 are hypothesized to facilitate binding of the negatively charged ACA **2** substrate.⁵⁹ The lack of

detectable activity in R70A and R73A variants is consistent with this hypothesis. The similar electron densities of the active sites of H28A and R73A indicate the presence of a heteroatom bonded to the C-3 of the substrate, in addition to the covalent bond to the nitrogen of Pro-1 (Figure 1.35B-C). This led to the conclusion that the H28A and R73A mutants contained 3-(*N*-prolyl)-3-hydroxypropionate, the result of water addition to a covalently bound intermediate. The electron density in the active site of the E114D variant occupied the top of the proline ring (left side) with an ordered water molecule on the right side. The shape and geometry of the electron density indicated a covalent intermediate. A 3-(*N*-prolyl)-acrylate intermediate was modelled in the active site of the E114D mutant (Figure 1.35D). Most of the electron density in the active site of Y103F variant is located alongside the Pro-1 and indicated a flatter geometry for the nitrogen of the proline relative to E114D (Figure 1.35E). The electron density on this variant was consistent with an iminium intermediate.

The E114D variant crystallized in a different lattice relative to the wild type and other variants examined, where twelve protein chains (four homotrimers) made up the asymmetric unit of the E114D variant (Figure 1.36). The active site electron density showed significant differences between the twelve chains, most notably in the vicinity of Pro-1 (Figure 1.37). Electron density consistent with a 3-acrylate covalent intermediate bound to the Pro-1 nitrogen was seen in protomers A, B, H, I and K, and the orientation of the intermediate was suggestive of an enamine. In contrast, the electron density in the other protomers could be accounted for by partial occupancy of a sulfate ion, which is a component of the crystallization condition.¹ Though the enamine intermediate was satisfactorily occupied, the additional electron density in molecules A, B, H, I and K may also be due to partial occupancy of sulfate. Neither can it be ruled out that there is partial occupancy of the enamine intermediate in some of the other molecules as well (Figure 1.37).

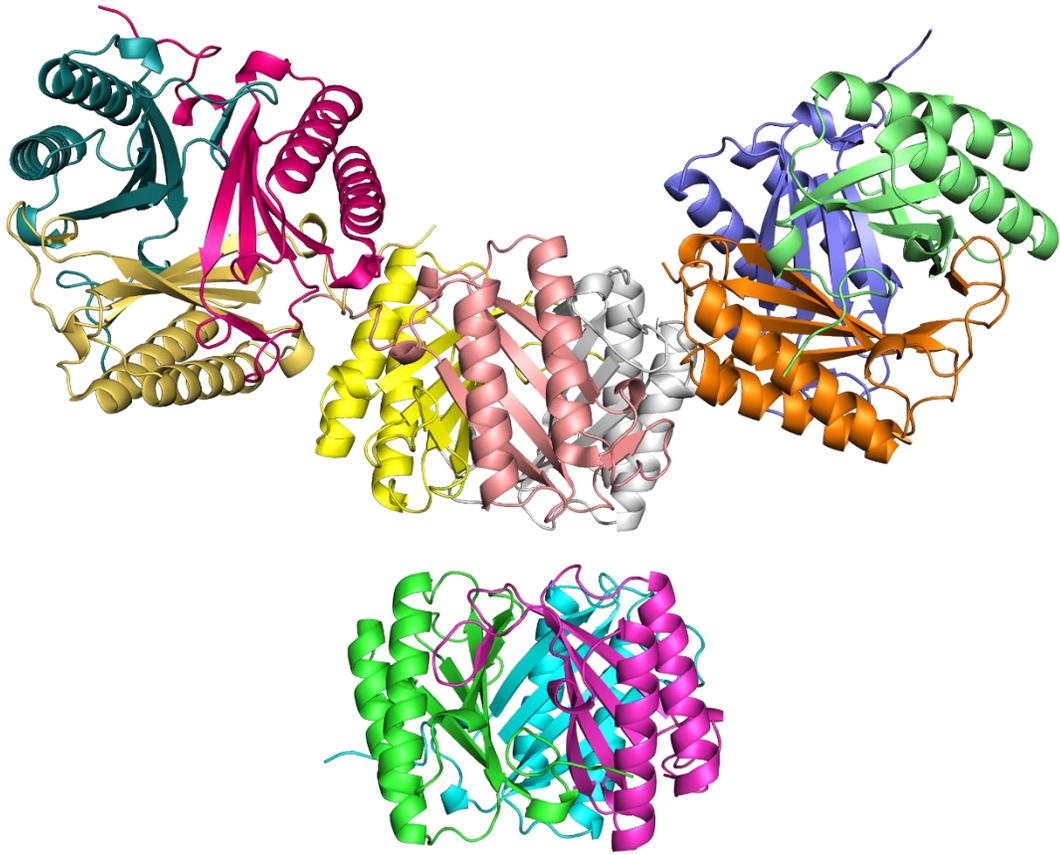


Figure 1.36. Asymmetric unit of Cg10062(E114D) with twelve chains in four trimers.

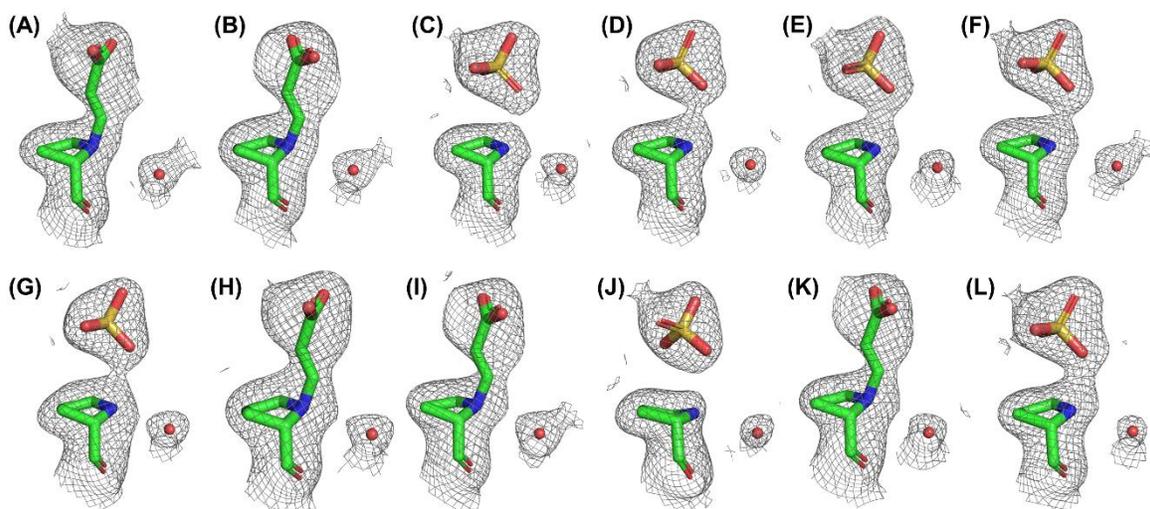


Figure 1.37. 3-(*N*-prolyl)-acrylate (enamine) intermediate in chains A, B, H, I and K of Cg10062(E114D). The remaining chains were modelled with a sulfate ion.

In comparison to the wild type enzyme, the E114D variant structures depict that the shorter Asp-114 residue creates space for a water molecule to occupy a position that affords the formation of four strong hydrogen bonds with the nitrogen of Pro-1, the hydroxyl of Tyr-103, the carboxylate of Asp-114, and the main chain carbonyl of Leu-38 to fully satisfy its coordination (Figure 1.38A). It was hypothesized that this tightly coordinated water molecule plays a crucial role in Cg10062(E114D) acting as a hydratase-only variant. It was presumed that the tetracoordinated water interacts with the nitrogen of Pro-1, preventing formation of an iminium intermediate required for the formation of the decarboxylation product **8**. In order to support this hypothesis, the E114S variant was prepared where the shorter side chain of the Ser-114 residue would prevent a similar interaction with a water molecule in the same position as observed for E114D. However, the expected decarboxylation product was not observed as this mutant was completely inactive (Table 1.2). Furthermore, the absence of the water molecule in wild type Cg10062 led to the conclusion that the shorter Asp-114 creates space for the water molecule, leading to design of the E114N variant with a side chain of similar length. Biochemical characterization and the crystal structure of ACA-soaked E114N, which shows a similar interaction with the water molecule

and the Pro-1 with a covalent intermediate, support the hypothesis that water molecule plays a crucial role in preventing decarboxylation (Figure 1.39B).

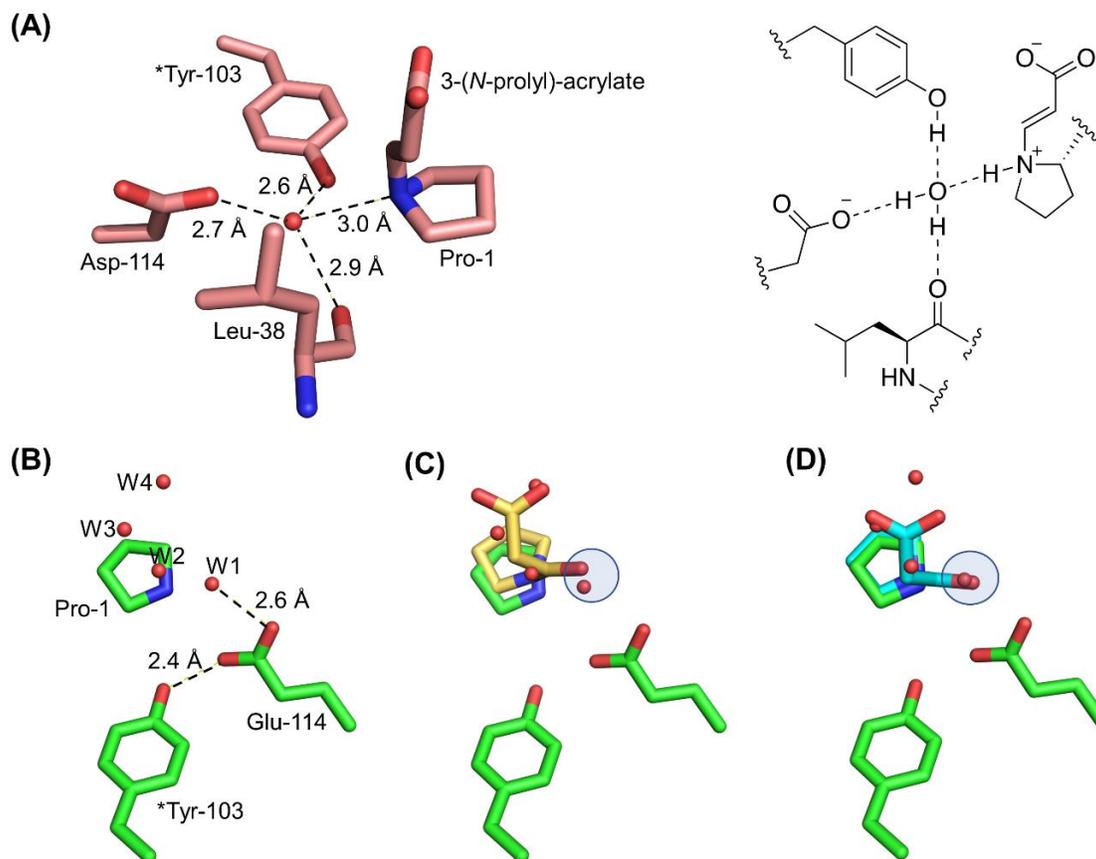


Figure 1.38. Comparison of active sites of Cg10062 variants. (A) Water molecule in the active site of E114D makes four hydrogen bonds with the enzyme. The nitrogen of Pro-1 and the hydroxyl of Tyr-103 serve as hydrogen bond donors, and the carboxylate of Asp-114 and the main chain carbonyl of Leu-38 serve as hydrogen bond acceptors. The active site water molecules (W1-W4) of (B) wild type Cg10062 overlaid with the active site intermediates of (C) H28A and (D) R73A. W1 overlaps with the hydroxyl group of the covalent intermediates formed in H28A and R73A (blue circle).

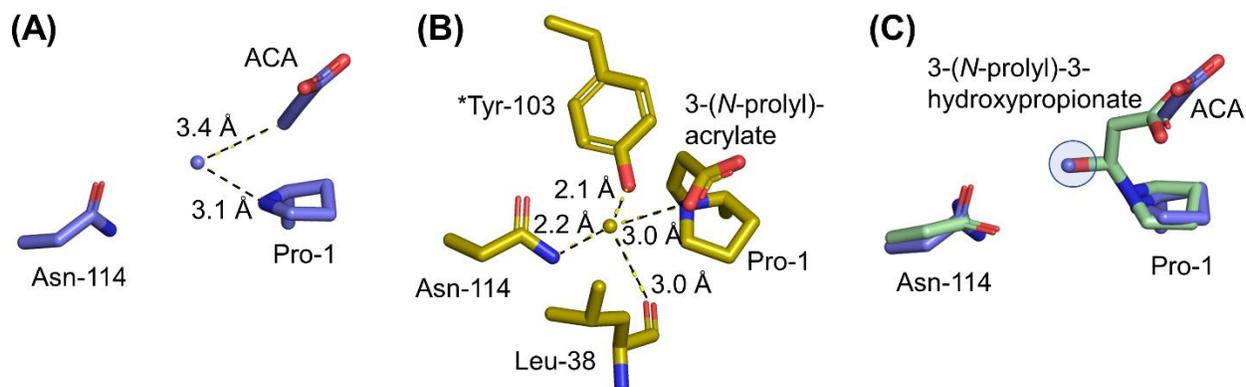


Figure 1.39. Active site intermediates in Cg10062(E114N). (A) Unreacted ACA and water molecule in the active site of E114N. (B) water molecule in the active site of E114N makes four hydrogen bonds with the enzyme similar to that seen with E114D. (C) Overlay of (A) (blue) with a E114N active site containing the 3-(*N*-prolyl)-3-hydroxypropionate intermediate on Pro-1 (green). The water molecule from (A) (blue circle) overlaps with the hydroxyl group of the 3-(*N*-prolyl)-3-hydroxypropionate covalent intermediate on Pro-1 (green).

1.9. Proposed mechanism for hydration of ACA by Cg10062

The combination of kinetic and detailed structural analysis of a set of active site mutants of Cg10062 have unveiled a comprehensive picture of the mechanistic elements of the enzyme and provided structurally rich information regarding the roles of several active site residues.⁵⁹ The crystals of attenuated Cg10062 variants soaked in ACA **2**, reveal the catalytic cycle of both hydration and hydration/decarboxylation of ACA (Figure 1.40). The first step of the reaction appears to be a nucleophilic attack on C-3 of the acetylene substrate to form a covalent, enamine intermediate 3-(*N*-prolyl)-acrylate **15**, as captured in the active sites of E114D (Figure 1.38A) and E114N (Figure 1.39B). This is followed by tautomerization of the enamine intermediate to the iminium form of 3-(*N*-prolyl)-acrylate **16**, as captured in the Y103F active site. The strongly electron-withdrawing iminium species can activate the decarboxylation pathway or water addition can lead to the 3-(*N*-prolyl)-3-hydroxypropionate **17** species. The hydroxylated intermediate **17** is captured in the active sites of both H28A and R73A variants, as well as the new E114N variant (Figure 1.39C). Arg-73 and His-28 in the active sites of H28A and R73A variants, respectively, do not interact with carboxylate group of the 3-(*N*-prolyl)-3-hydroxypropionate **17**,⁶⁶ suggesting that either their substrate-binding role is critical before the hydration step or that they are co-dependent

and mutation of either one results in the loss of their hydrogen-bonding interactions with the intermediate **17**. In the final step of the reaction, the loss of the covalent bond between the nitrogen of Pro-1 and the intermediate **17** occurs resulting in MSA **7** formation. Structural data described herein supports a covalent mechanism similar to one previously proposed by Whitman and coworkers,⁵⁹ although a mechanism that does not go through a covalent intermediate cannot be strictly ruled out for wild type Cg10062.

The nature of the intermediates trapped in the active sites of the variants shed light on the roles these residues play in the hydration of ACA **2**. Both E114D (Figure 1.38A) and E114N (Figure 1.39B), and Y103F (Figure 1.35E) trap the intermediates 3-(*N*-prolyl)-acrylate **15**, and **16**, respectively, that have not been attacked by water. This is consistent with the idea that Glu-114 and Tyr-103 are responsible for activating the attacking water molecule, as previously hypothesized in the case of *cis*-CaaD and Cg10062.^{59,67} Mutation of either residue compromises water addition, potentially making it the rate-determining step in the pathway for these mutants and resulting in slower catalysis relative to wild type Cg10062.

In the case of H28A and R73A, water addition is not compromised, leading to the covalently bound 3-(*N*-prolyl)-3hydroxypropionate **17**, suggesting that proline release and aldehyde formation may be the rate-determining step in these variants. Consistent with this hypothesis, an overlay of either H28A or R73A with the apo wild-type structure (Figure 1.38C-D) show that the water molecule (W1) hydrogen bonded to Glu-114 in the apo wild type structure almost perfectly overlaps with the hydroxyl group of the intermediate (Figure 1.38B), strongly suggesting that W1 ultimately attacks the substrate. Though Tyr-103 does not interact with W1, it forms a tight hydrogen bond with Glu-114, appropriately positioning it for the reaction and potentially activating Glu-114 to deprotonate W1, as it attacks the covalently bound substrate. A similar interaction is observed with ACA-soaked E114N, where the water molecule overlaps perfectly with the hydroxyl group of the 3-(*N*-prolyl)-3hydroxypropionate **23** intermediate observed in the E114N active site (Figure 1.39C).

It is instructive to remember that E114D and E114N are the only two of the five crystallized variants found to have no decarboxylation activity (Figure 1.18, Table 1.2). As discussed previously, the shortening of the 114 sidechain creates a space for a water molecule to be simultaneously hydrogen bonded to four moieties. Since the carboxylate of Asp-114 and the main chain carbonyl of Leu-38 would presumably be the hydrogen bond acceptors, the nitrogen of Pro-1 and the hydroxyl of Tyr-103 would act as hydrogen bond donors, which requires the Pro-1 nitrogen to be protonated. A similar interaction is observed with E114N (Figure 1.39B) where the Asn-114 acts as a proton acceptor. Since a protonated nitrogen of a tertiary amine does not form an iminium, the presence of this fully coordinated water molecule would be expected to suppress iminium formation and therefore inhibit the decarboxylation reaction. In other words, E114D and E114N likely bypass the formation of the iminium and avoid the decarboxylation pathway.

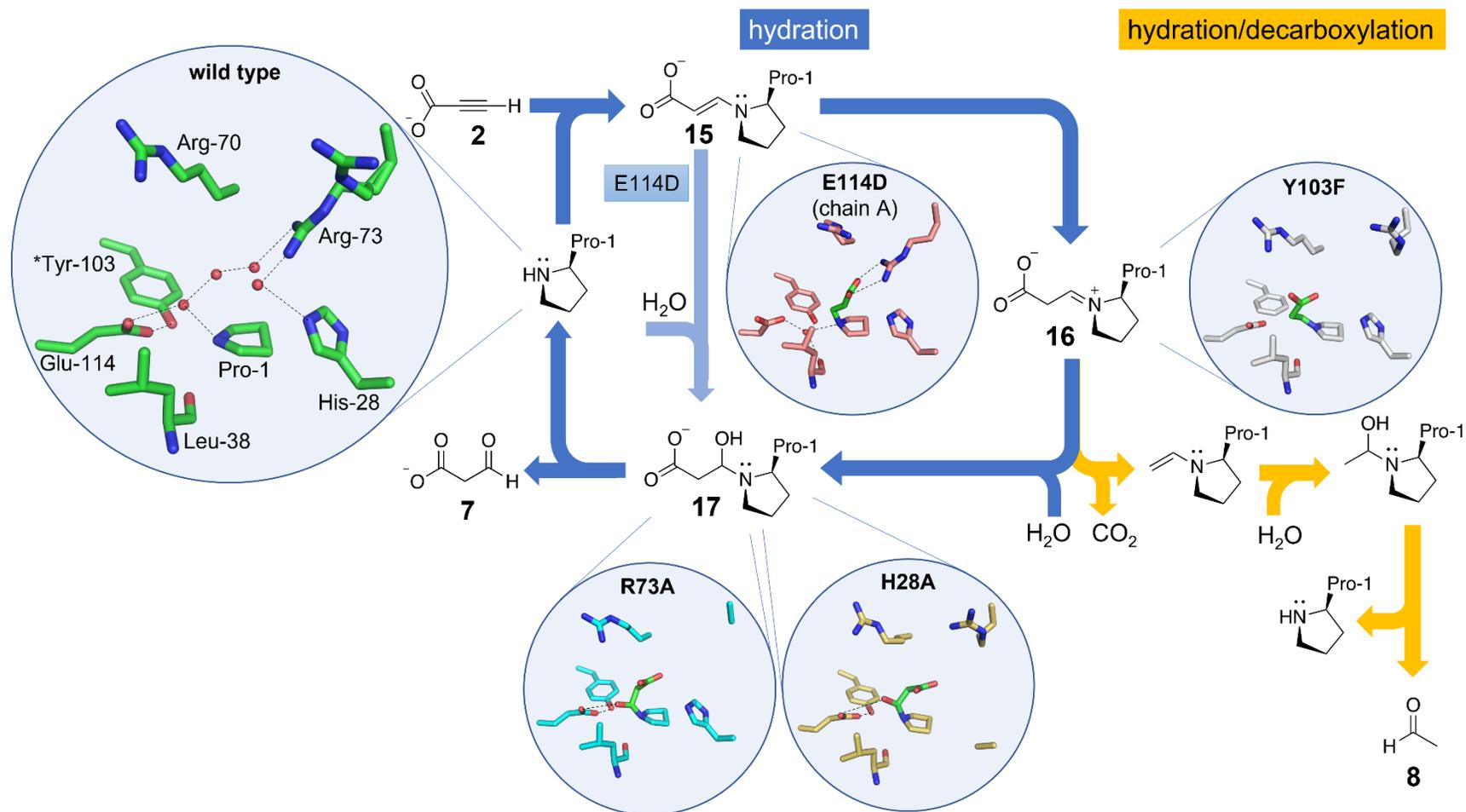


Figure 1.40. Proposed catalytic cycle of Cg10062 hydration and hydration/decarboxylation of ACA. The wild type and the intermediate-bound active sites are shown (circles). In wild-type Cg10062, nucleophilic attack on ACA **2** affords enamine intermediate 3-(*N*-prolyl)-acrylate **15**, which tautomerizes to the iminium intermediate 3-(*N*-prolyl)-acrylate **16**. Decarboxylation and hydration (yellow) resulting in acetaldehyde **8** occurs exclusively via iminium **16**. Hydration of either **15** or **16** (blue) provides 3-(*N*-prolyl)-3-hydroxypropionate **17** preventing further decarboxylation. Release of **7** from the active site regenerates the free Pro-1. The E114D and E114N variants suppresses iminium formation, thus preventing decarboxylation.

1.10. Assessing the growth of *E. coli* expressing Cg10062 or variant on ACA

In order to explore the possibility of *E. coli* utilizing ACA 2 as the sole carbon source for growth and energy, wild type Cg10062 and variants were expressed from a plasmid using a *trc* promoter in *E. coli* RB791 *serA*. *E. coli* RB791 *serA*^{68,69} is unable to express the *serA* gene encoding D-3-phosphoglycerate dehydrogenase, an enzyme required for L-serine biosynthesis. The plasmids pAS1.163, pAS1.240 and pAS1.241 expressing Cg10062 and variants E114Q and E114D, respectively, were prepared (see Chapter Four) from plasmid pKM6.240. The pKM6.240 vector⁷⁰ contains *serA* downstream of *P_{trc}*. Elimination of *serA* from the host organism and inclusion of *serA* on the plasmid forces stable plasmid maintenance in minimal media. The genes expressed by each strain is summarized below (Table 1.4). The experiments were carried out in solid and liquid M9 minimal media containing glucose and/or ACA 2.

Table 1.4. Genes expressed by *E. coli* RB791 *serA* transformants.

Strain #	<i>E. coli</i> Strain	<i>serA</i> gene	<i>cg10062</i> gene
I	RB791 <i>serA</i>	–	–
II	RB791 <i>serA</i> /pKM6.240	+	–
III	RB791 <i>serA</i> /pAS1.163	+	wild type
IV	RB791 <i>serA</i> /pAS1.240	+	E114Q
V	RB791 <i>serA</i> /pAS1.241	+	E114D

Single colonies of RB791 *serA*, RB791 *serA*/pKM6.240, RB791 *serA*/pAS1.163, RB791 *serA*/pAS1.240 and RB791 *serA*/pAS1.241 were inoculated into four different types of freshly prepared M9 media. These included M9 (with no carbon), M9/glu, M9/ACA, and M9/ACA/glu where all cultures were induced with 1 mM IPTG. The cultures were incubated at 37 °C for 48 h and 1 mL of each culture was used to inoculate freshly prepared 5 mL media and incubated further for 48 h at 37 °C. The transfer of the cultures into fresh media was carried out at regular

time intervals rather than at a particular OD₆₀₀, since the growth pattern in some of cultures was unpredictable. The culturing was carried out for 12 cycles.

As expected, no growth was observed in any of the cultures without a source of carbon and all cultures, except RB791 *serA* (with no plasmid) grew to a similar OD₆₀₀ (~1) in the presence of glucose. The lack of growth of RB791 *serA* can be attributed to the absence of the *serA* gene required for L-serine biosynthesis, on the chromosome or via plasmid and this prevented the strain from growing in any of the media prepared for this experiment. Furthermore, growth of all other strains was observed in M9/ACA/glu to an OD₆₀₀ of ~1. These results suggest that ACA 2 did not appear to be toxic to cell growth as the cultures grew to a similar OD₆₀₀ as that observed in the presence of exclusively glucose. Unfortunately, no growth was observed in any of the strains cultured in M9 media containing exclusively ACA 2.

Similar to the liquid media, the RB791 *serA* strains expressing Cg10062 or variants were grown on four different types of plates: M9 (no carbon), M9/ACA, M9/glu, and M9/ACA/glu. Plates were supplemented with 1 mM IPTG to enable expression of Cg10062 and variants from the respective plasmids. Individual colonies of RB791 *serA*, RB791 *serA*/pKM6.240, RB791 *serA*/pAS1.163, RB791 *serA*/pAS1.240 and RB791 *serA*/pAS1.241 were replicate plated onto each plate and incubated at 37 °C for 7 days. The appearance of small white colonies of all strains (excluding RB791 *serA*) was observed on M9/glu and M9/ACA/glu plates within 12 h. No growth was observed on any of the plates lacking glucose. After 18 h, the media of the M9/ACA/glu plate surrounding colonies of RB791 *serA*/pAS1.163 expressing wild type Cg10062 appeared pale yellow (Figure 1.41). This phenomenon was not observed on the M9/glu plate.

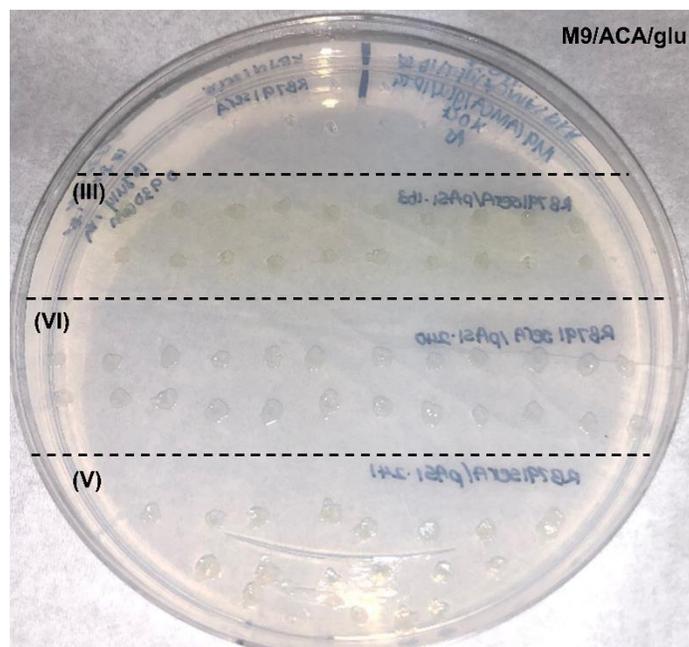


Figure 1.41. Growth of *E. coli* on M9/ACA/glu plates. After 18 h, the colonies and agar surrounding RB791 *serA*/pAS1.163 appeared pale yellow. The dotted lines separate different strains (Table 1.4). (III): RB791 *serA*/pAS1.163; (VI): RB791 *serA*/pAS1.240; (V): RB791 *serA*/pAS1.241.

After 24 h, the entire M9/ACA/glu plate, as well as the colonies themselves appeared yellow (Figure 1.42). Interestingly, further incubation of the plates for a total of four days revealed the formation of thin “film”-like of the Cg10062-expressing strains on M9 plates containing ACA 2 as the sole source of carbon (Figure 1.43), suggesting that ACA may be used by *E. coli* for cell growth, albeit at an extremely slow growth rate. However, the metabolic pathway for ACA utilization by the Cg10062-expressing strains remains unclear.

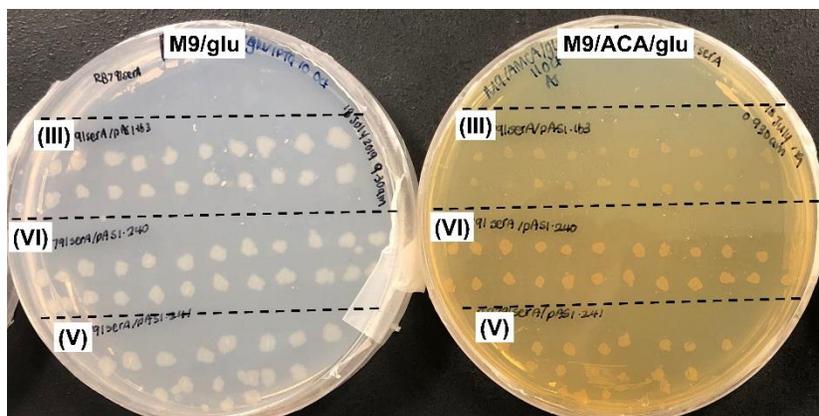


Figure 1.42. A comparison of *E. coli* growth on M9 plates after 24 h. The entire M9/ACA/glu plate appeared dark yellow relative to M9/glu. The dotted lines separate different strains (Table 1.4). (III): RB791 serA/pAS1.163; (VI): RB791 serA/pAS1.240; (V): RB791 serA/pAS1.241

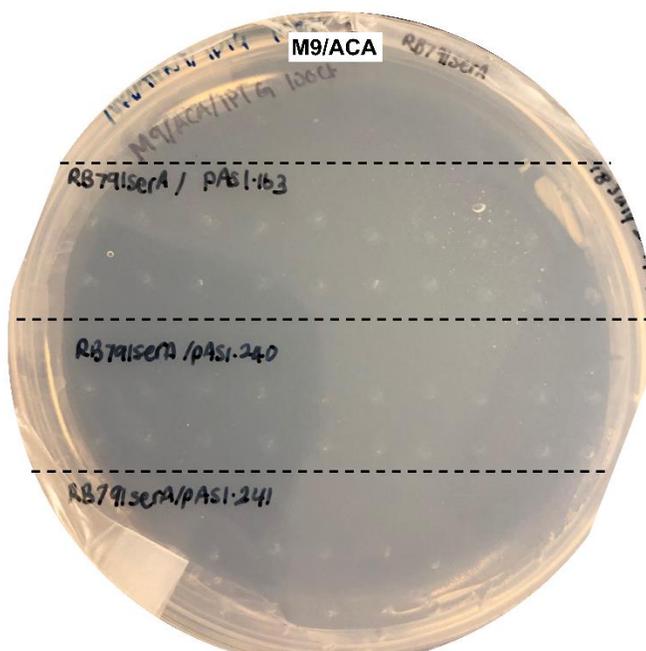


Figure 1.43. “Film”-like growth of *E. coli* expressing Cg10062 and variants on M9/ACA. These observations were made after four days of growth at 37 °C.

Collectively, these observations, where the strains expressing Cg10062 and variants appear yellow, indicate a correlation between catalysis of ACA 2 by Cg10062 and the color change. At the time of this experiment, an explanation for these observations was unclear. However, when Cg10062(E114N) was utilized for 3-HP synthesis (see chapter two), reactions that did not contain excess coupling enzyme (see chapter two) appeared yellow. HPLC analysis

indicated that ACA **2** was not completely converted to 3-HP. The yellow appearance was hypothesized to be due to a mixture of MSA **7** and acetaldehyde **8** accumulation (from spontaneous decarboxylation of MSA **7**), as confirmed by NMR analysis (data not shown). Literature reports that aldehydes can covalently modify side chains of certain amino acid residues in proteins, resulting in a pale yellow color and protein aggregation.^{71,72} This phenomenon was considered for use as a colorimetric method to screen for Cg10062 mutants with faster, hydratase-only activity. However, an alternative high-throughput method was desired for rapid screening of hydratase-only Cg10062 variants.

1.11. Growth of *E. coli* on aldehyde indicator plates

The Schiff test is widely used as a qualitative test for the detection of aldehydes in organic reactions and is also used as a biological stain to detect aldehyde production by various organisms.^{73,74} The Schiff reagent is a dye formulation typically consisting of a mixture of *p*-rosaniline and bisulfite (Figure 1.44). Conway et al. demonstrated the incorporation of the Schiff reagent into minimal media plates to successfully screen for *E. coli* strains expressing a heterogeneously expressed alcohol dehydrogenase.⁷⁵ *p*-Rosaniline is a pink solid that appears colorless upon sulfonation by sodium bisulfite of the central carbon atom, due to disruption of the extended delocalized pi electron system (Figure 1.44).⁷⁵ The uncharged aromatic amino groups can then react with acetaldehyde to produce a Schiff base or aldimine resulting in a deep red color.⁷⁵ In colonies producing large amounts of aldehydes, the aldehydes can diffuse into the surrounding agar resulting in a deep red color.^{73,75}

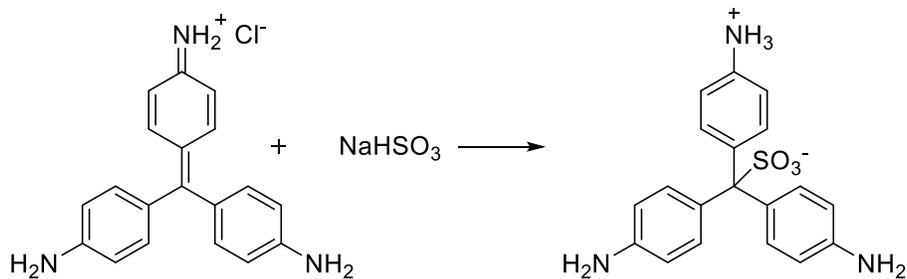


Figure 1.44. *p*-Rosaniline reacts with bisulfite to produce "decolorized" Schiff reagent.

The Schiff test was used as a qualitative method to investigate the diffusion of ACA or ACA-derived metabolites into the surrounding media of *E. coli* expressing Cg10062 and variants. Any MSA **7** or acetaldehyde **8** produced from Cg10062 catalytic activity was expected to react with the Schiff reagent (ros), resulting in a color change. Six types of plates containing 1 mM IPTG were prepared: M9/ACA, M9/ACA/ros, M9/glu, M9/glu/ros, M9/ACA/glu, M9/ACA/glu/ros, and sealed with aluminum foil to reduce potential background color⁷⁵ and stored at 4 °C. Individual colonies of RB791 *serA*, RB791 *serA*/pKM6.240, RB791 *serA*/pAS1.163, RB791 *serA*/pAS1.240 and RB791 *serA*/pAS1.241 were replicate plated onto each type of plate, wrapped in foil and incubated at 37 °C.

RB791 *serA* did not grow on any of the plates due to the lack of L-serine supplementation.^{68,69} After 12 h, growth of all other strains were observed on M9/glu, M9/glu/ros and M9/ACA/glu/ros plates with relatively larger colonies observed on M9/glu and M9/ACA/glu/ros. The agar surrounding RB791 *serA*/pAS1.163 expressing wild type Cg10062 appeared slightly pink on the M9/ACA/glu/ros plate (Figure 1.45). Plates containing ACA **2** and the aldehyde indicator appear slightly pink even in the absence of cell growth (on M9/ACA/ros). It was hypothesized that at 37 °C, ACA is slowly being degraded to acetaldehyde **8** as previous studies by other researchers in the Draths group have indicated the formation of trace amounts

of acetaldehyde **8** from ACA degradation at elevated temperatures, although it remains unclear if it proceeds via MSA **7** intermediate.

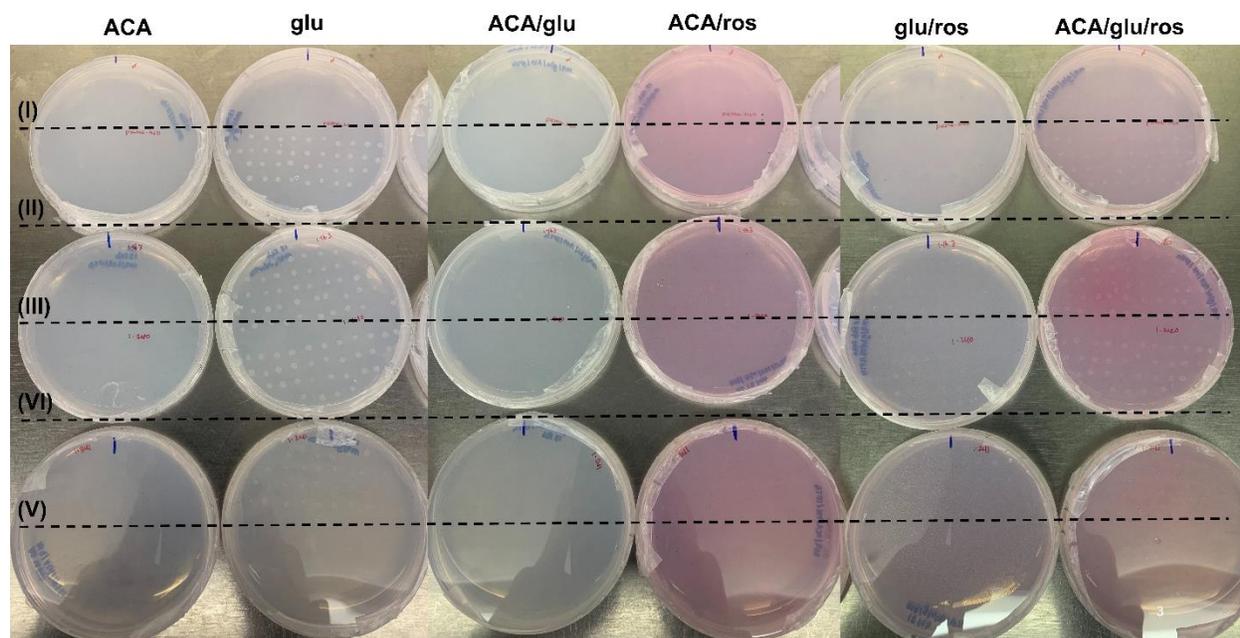


Figure 1.45. Growth of *E. coli* RB791 *serA* on aldehyde indicator plates. Transformants were replicate-plated on to M9 plates containing either ACA, glucose, or a combination of both. The last three plates contain the Schiff reagent (*ros*). The dotted lines separate different strains (Table 1.4). (I): RB791 *serA*; (II): RB791 *serA*/pKM6.240; (III): RB791 *serA*/pAS1.163; (VI): RB791 *serA*/pAS1.240; (V): RB791 *serA*/pAS1.241.

After 48 h of incubation, no growth was observed in any of the plates lacking glucose. White colonies of similar size were now observed on all of the plates containing glucose for all strains except RB791 *serA* (Table 1.4). Over time, all plates containing the aldehyde indicator turned slightly pink. However, the most significant observation was made with M9/ACA/glu/*ros* plates, where different areas of the plates displayed varying shades of red (Figure 1.46). The agar surrounding RB791 *serA* appeared light pink although no growth was observed and accounted for background. The colonies of RB791 *serA*/pKM6.240 as well as the agar were darker relative to the background. This may be caused by aldehydes formed natively by *E. coli* during growth. In contrast, the colonies of RB791 *serA*/pAS1.163 expressing Cg10062 appeared red with a dark red ring formation around the colonies (Figure 1.47). This intense red color can be attributed to

the diffusion of MSA **7** and acetaldehyde **8** accumulated from the hydration and hydration-dependent decarboxylation of ACA **2** by Cg10062, reacting with the Schiff reagent. Similar observations were made with the strains expressing E114Q and E114D variants although the colonies appeared relatively lighter due to the low efficiency of the two hydratase-only variants. Further incubation of the plates resulted in the appearance of the dark red ring around all strains expressing Cg10062 or variants in M9/ACA/glu/ros plates (Figure 1.48).

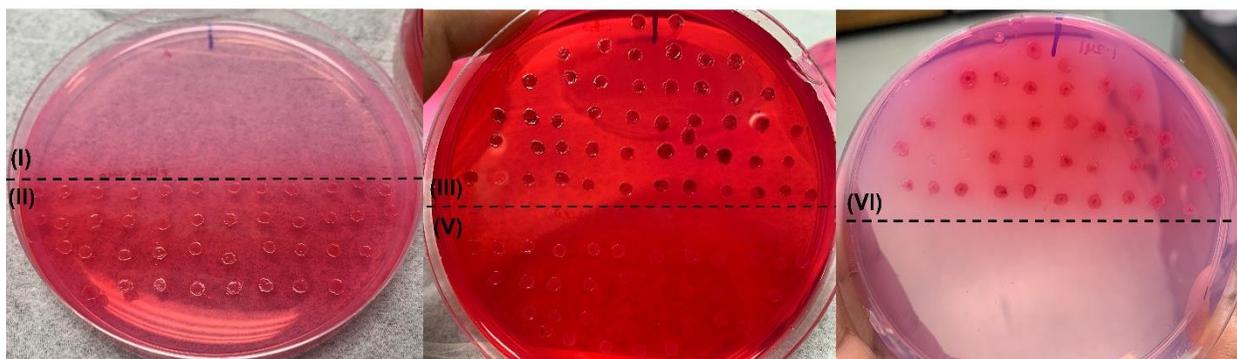


Figure 1.46. Color changes observed on M9/ACA/glu/ros after 48 h of growth. The dotted lines separate different strains (Table 1.4). (I): RB791 *serA*; (II): RB791 *serA/pKM6.240*; (III): RB791 *serA/pAS1.163*; (VI): RB791 *serA/pAS1.240*; (V): RB791 *serA/pAS1.241*.



Figure 1.47. Dark red rings around RB791*serA/pAS1.163*. The color of the outer rings on M9/ACA/glu/ros intensified after 72 h of growth.

A comparison of all strains after 60 h of growth on *p*-rosaniline-containing plates provides a clear picture of the varying degrees of aldehyde production in each strain under different

conditions of growth (Figure 1.48). Any color change in M9/glu/ros was minimal relative to the ACA-containing plates where both M9/ACA/ros and M9/ACA/glu/ros appeared dark red. The M9/ACA/glu/ros plates carrying the Cg10062-expressing strains appeared dark purple, especially around the area of colony formation, with the wild type expressing colonies appearing the darkest. “Film”-like growth was observed by strains III, VI and V on M9/ACA/ros, suggesting that ACA 2 may indeed serve as a potential carbon source for *E. coli* strains expressing Cg10062. No growth was observed with the control strains I and II on M9/ACA/ros supporting this hypothesis. These observations are consistent with the “film”-like growth observed in M9/ACA plates (see Section 1.10.2) that ACA 2 may function as a carbon source for growth, albeit not very effectively. The apparent growth of strains II, VI and V were confirmed after further incubation of the plates for a total of four days, where the colonies on M9/ACA/ros can be characterized by their dark purple color and coloration of the agar surrounding these colonies, particularly around strain III expressing the efficient wild type Cg10062 (Figure 1.49). While the aldehyde indicator does not differentiate between MSA 7 and acetaldehyde 8, the differences in color intensity is like related to the difference in catalytic activity of each variant.

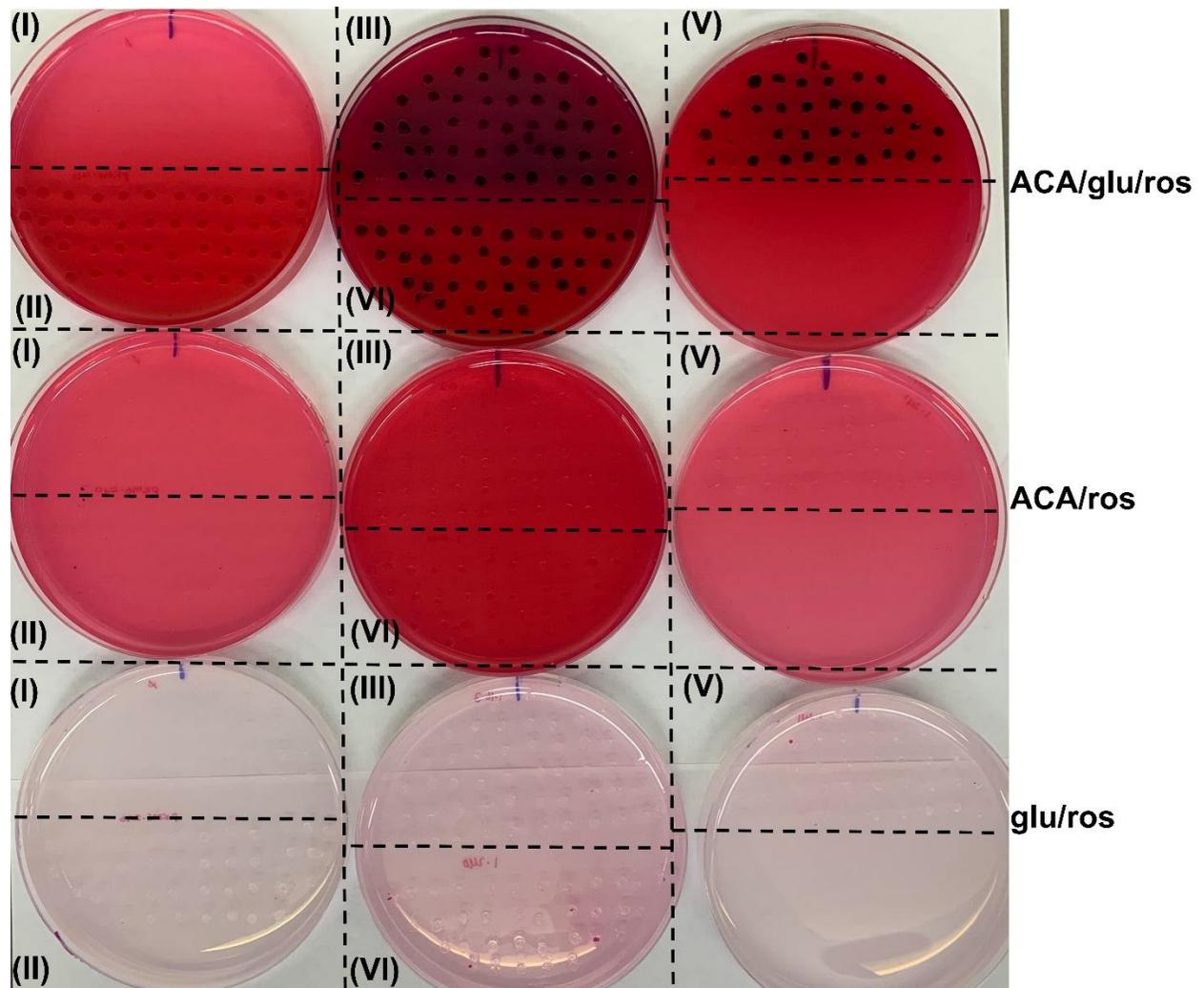


Figure 1.48. Growth of *E. coli* RB791 *serA* on aldehyde indicator plates. After 60 h, strains expressing Cg10062 appeared dark purple. “Film”-like growth of the same strains was observed on M9/ACA/ros plates. The dotted lines separate different strains (Table 1.4). (I): RB791 *serA*; (II): RB791 *serA*/pKM6.240; (III): RB791 *serA*/pAS1.163; (VI): RB791 *serA*/pAS1.240; (V): RB791 *serA*/pAS1.241.

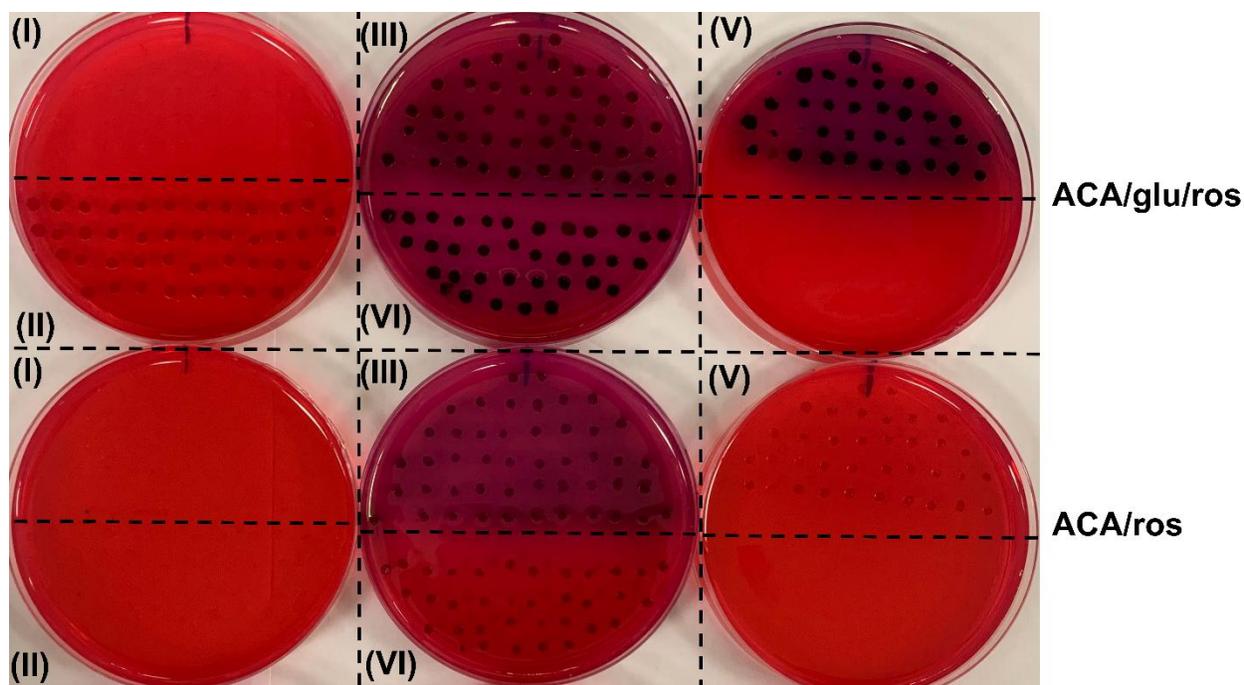


Figure 1.49. Growth of *E. coli* RB791 *serA* on M9/ACA/ros. After four days, the aldehyde indicator resulted in color changes that confirmed the growth of strains expressing Cg10062 on ACA as the sole carbon source. The dotted lines separate different strains (Table 1.4). (I): RB791 *serA*; (II): RB791 *serA*/pKM6.240; (III): RB791 *serA*/pAS1.163; (VI): RB791 *serA*/pAS1.240; (V): RB791 *serA*/pAS1.241.

1.12. Conclusion

Motivated by the efforts to reduce dependence on petroleum-derived carbon, ACA was proposed as a unique carbon feedstock for microbial biocatalysis. While an industrial-scale sustainable route to ACA does not exist at present, it was proposed that C1-derived ACA may be produced from CH₄ dehydrodimerization,³⁸ followed by direct carboxylation of acetylene.⁴¹ In this study, tautomerase enzyme Cg10062 was selected as a suitable candidate for the incorporation of ACA into bacterial metabolism. However, wild type Cg10062 produces a mixture of MSA and acetaldehyde.⁵⁹ In the pursuit of accessing C3 chemicals via ACA, the loss of a carbon from the decarboxylation of MSA is undesirable. E114Q and E114D variants were previously identified by others⁵⁹ as hydratase-only variants but the low catalytic efficiency exhibited by both variants prompted deeper investigation into the mechanism of ACA hydration and hydration-dependent decarboxylation of MSA to form acetaldehyde. Using a combination of site-directed mutagenesis,

kinetic characterization and protein crystallography, several covalent intermediates were captured in ACA-soaked crystals of Cg10062 variants, including a novel hydratase-only E114N variant. Together, these results indicate that ACA hydration proceeds via covalent catalysis, where the Pro-1 intermediate plays a crucial role in forming a covalent intermediate with the substrate. A tetracoordinated water molecule observed in the active sites of two non-decarboxylating variants, E114D and E114N, implied that the decarboxylation proceeds via iminium formation, which is not possible in these variants due to the presence of the water. Rational mutagenesis of the active site provided a deeper understanding of the catalytic mechanism and sets the stage for further optimization of Cg10062(E114N) hydratase activity using a random mutagenesis approach. Furthermore, the growth of *E. coli* expressing Cg10062 and its E114Q and E114D variants, indicated that ACA transport into *E. coli* allows the cells to utilize the acid as a sole source of carbon, albeit not very efficiently. Further investigation of the *E. coli* transport mechanisms and metabolic pathways incorporating the MSA and acetaldehyde intermediates are necessary to allow the establishment of ACA as a novel carbon feedstock for biocatalysis.

REFERENCES

REFERENCES

- (1) Mathes Hewage, A.; Nayebi Gavvani, H.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochemistry* 2021, 60 (51), 3879–3886.
- (2) Yadav, V. G.; Yadav, G. D.; Patankar, S. C. The Production of Fuels and Chemicals in the New World: Critical Analysis of the Choice between Crude Oil and Biomass Vis-à-Vis Sustainability and the Environment. *Clean Technol. Environ. Policy* 2020, 22 (9), 1757–1774.
- (3) Bardhan, S. K.; Gupta, S.; Gorman, M. E.; Haider, M. A. Biorenewable Chemicals: Feedstocks, Technologies and the Conflict with Food Production. *Renew. Sustain. Energy Rev.* 2015, 51, 506–520.
- (4) IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
- (5) Correa, D. F.; Beyer, H. L.; Fargione, J. E.; Hill, J. D.; Possingham, H. P.; Thomas-Hall, S. R.; Schenk, P. M. Towards the Implementation of Sustainable Biofuel Production Systems. *Renew. Sustain. Energy Rev.* 2019, 107, 250–263.
- (6) US EPA, O. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2020. <https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2020> (accessed 2022-07-08).
- (7) Miller, R. G.; Sorrell, S. R. The Future of Oil Supply. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* 2014, 372 (2006), 20130179.
- (8) Zimmerman, J. B.; Anastas, P. T.; Erythropel, H. C.; Leitner, W. Designing for a Green Chemistry Future. *Science* 2020, 367 (6476), 397–400.
- (9) National Research Council. Sustainability in the Chemical Industry: Grand Challenges and Research Needs; National Academies Press: Washington, DC, 2006. <https://doi.org/10.17226/11437>.
- (10) Takkellapati, S.; Li, T.; Gonzalez, M. A. An Overview of Biorefinery-Derived Platform Chemicals from a Cellulose and Hemicellulose Biorefinery. *Clean Technol. Environ. Policy* 2018, 20 (7), 1615–1630.
- (11) Frisvold, G. B.; Moss, S. M.; Hodgson, A.; Maxon, M. E. Understanding the U.S. Bioeconomy: A New Definition and Landscape. *Sustainability* 2021, 13 (4), 1627.
- (12) Guo, M.; Song, W. The Growing U.S. Bioeconomy: Drivers, Development and Constraints. *New Biotechnol.* 2019, 49, 48–57.

- (13) Cavani, F.; Albonetti, S.; Basile, F.; Gandini, A. *Chemicals and Fuels from Bio-Based Building Blocks*; John Wiley & Sons, 2016.
- (14) Sun, H.; Zhang, H.; Ang, E. L.; Zhao, H. Biocatalysis for the Synthesis of Pharmaceuticals and Pharmaceutical Intermediates. *Bioorg. Med. Chem.* 2018, 26 (7), 1275–1284. <https://doi.org/10.1016/j.bmc.2017.06.043>.
- (15) Patel, R. N. Biocatalysis for Synthesis of Pharmaceuticals. *Bioorg. Med. Chem.* 2018, 26 (7), 1252–1274.
- (16) Johnson, C. W.; Salvachúa, D.; Rorrer, N. A.; Black, B. A.; Vardon, D. R.; St. John, P. C.; Cleveland, N. S.; Dominick, G.; Elmore, J. R.; Grundl, N.; Khanna, P.; Martinez, C. R.; Michener, W. E.; Peterson, D. J.; Ramirez, K. J.; Singh, P.; VanderWall, T. A.; Wilson, A. N.; Yi, X.; Bidy, M. J.; Bomble, Y. J.; Guss, A. M.; Beckham, G. T. Innovative Chemicals and Materials from Bacterial Aromatic Catabolic Pathways. *Joule* 2019, 3 (6), 1523–1537.
- (17) Bidy, M. J.; Scarlata, C.; Kinchin, C. *Chemicals from Biomass: A Market Assessment of Bioproducts with Near-Term Potential*; NREL/TP--5100-65509, 1244312; 2016; p NREL/TP--5100-65509, 1244312. <https://doi.org/10.2172/1244312>.
- (18) Schwartz, T. J.; O'Neill, B. J.; Shanks, B. H.; Dumesic, J. A. Bridging the Chemical and Biological Catalysis Gap: Challenges and Outlooks for Producing Sustainable Chemicals. *ACS Catal.* 2014, 4 (6), 2060–2069.
- (19) J. Bozell, J.; R. Petersen, G. Technology Development for the Production of Biobased Products from Biorefinery Carbohydrates—the US Department of Energy's "Top 10" Revisited. *Green Chem.* 2010, 12 (4), 539–554.
- (20) Saini, R.; Osorio-Gonzalez, C. S.; Hegde, K.; Brar, S. K.; Magdoui, S.; Vezina, P.; Avalos-Ramirez, A. Lignocellulosic Biomass-Based Biorefinery: An Insight into Commercialization and Economic Standout. *Curr. Sustain. Energy Rep.* 2020, 7 (4), 122–136.
- (21) Senatore, A.; Dalena, F.; Sola, A.; Marino, A.; Valletta, V.; Basile, A. Chapter 2 - First-Generation Feedstock for Bioenergy Production. In *Second and Third Generation of Feedstocks*; Basile, A., Dalena, F., Eds.; Elsevier, 2019; pp 35–57. <https://doi.org/10.1016/B978-0-12-815162-4.00002-1>.
- (22) Rosegrant, M. W.; Msangi, S. Consensus and Contention in the Food-Versus-Fuel Debate. *Annu. Rev. Environ. Resour.* 2014, 39 (1), 271–294.
- (23) Searchinger, T.; Heimlich, R. *Avoiding Bioenergy Competition for Food Crops and Land*, Working Paper, Installment 9 of Creating a Sustainable Food Future, World Resources Institute: Washington DC, 2015.
- (24) Liu, Z.; Wang, K.; Chen, Y.; Tan, T.; Nielsen, J. Third-Generation Biorefineries as the Means to Produce Fuels and Chemicals from CO₂. *Nat. Catal.* 2020, 3 (3), 274–288.
- (25) Xu, L.; Xiu, Y.; Liu, F.; Liang, Y.; Wang, S. Research Progress in Conversion of CO₂ to Valuable Fuels. *Molecules* 2020, 25 (16).

- (26) Zhao, T.-T.; Feng, G.-H.; Chen, W.; Song, Y.-F.; Dong, X.; Li, G.-H.; Zhang, H.-J.; Wei, W. Artificial Bioconversion of Carbon Dioxide. *Chin. J. Catal.* 2019, *40* (10), 1421–1437.
- (27) Nguyen, A. D.; Lee, E. Y. Engineered Methanotrophy: A Sustainable Solution for Methane-Based Industrial Biomanufacturing. *Trends Biotechnol.* 2021, *39* (4), 381–396.
- (28) Cantera, S.; Bordel, S.; Lebrero, R.; Gancedo, J.; García-Encina, P. A.; Muñoz, R. Bio-Conversion of Methane into High Profit Margin Compounds: An Innovative, Environmentally Friendly and Cost-Effective Platform for Methane Abatement. *World J. Microbiol. Biotechnol.* 2019, *35* (1), 16.
- (29) Kasprzycka, A.; Lalak-Kańczugowska, J.; Walkiewicz, A.; Bulak, P.; Proc, K.; Stępień, Ł. Biocatalytic Conversion of Methane – Selected Aspects. *Curr. Opin. Chem. Eng.* 2019, *26*, 28–32.
- (30) A. Henard, C.; R. Akberdin, I.; G. Kalyuzhnaya, M.; T. Guarnieri, M. Muconic Acid Production from Methane Using Rationally-Engineered Methanotrophic Biocatalysts. *Green Chem.* 2019, *21* (24), 6731–6737.
- (31) Wang, Y.; Sun, T.; Gao, X.; Shi, M.; Wu, L.; Chen, L.; Zhang, W. Biosynthesis of Platform Chemical 3-Hydroxypropionic Acid (3-HP) Directly from CO₂ in *Cyanobacterium synechocystis* sp. PCC 6803. *Metab. Eng.* 2016, *34*, 60–70.
- (32) Saur, G.; Milbrandt, A. *Renewable Hydrogen Potential from Biogas in the United States*; NREL/TP-5400-60283; National Renewable Energy Lab. (NREL), Golden, CO (United States), 2014. <https://doi.org/10.2172/1149657>.
- (33) Sahoo, K. K.; Goswami, G.; Das, D. Biotransformation of Methane and Carbon Dioxide into High-Value Products by Methanotrophs: Current State of Art and Future Prospects. *Front. Microbiol.* 2021, *12*.
- (34) Steckhan, E. Electrochemistry, 3. Organic Electrochemistry. In *Ullmann's Encyclopedia of Industrial Chemistry*; John Wiley & Sons, Ltd, 2011.
- (35) Fry, A. J. Electrochemical Processing, Organic. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Ltd, 2004.
- (36) Hort, E. V.; Taylor, P. Acetylene-Derived Chemicals. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Ltd, 2000.
- (37) Nohe, H.; Jaeger, P.; Aktiengesellschaft, B. Manufacture of Propiolic Acid. US3969200A, July 13, 1976.
- (38) Dinh, D. K.; Lee, D. H.; Song, Y.-H.; Jo, S.; Kim, K.-T.; Iqbal, M.; Kang, H. Efficient Methane-to-Acetylene Conversion Using Low-Current Arcs. *RSC Adv.* 2019, *9* (56), 32403–32413.
- (39) Arndt, M.; Risto, E.; Krause, T.; Gooßen, L. J. C-H Carboxylation of Terminal Alkynes Catalyzed by Low Loadings of Silver(I)/DMSO at Ambient CO₂ Pressure. *ChemCatChem* 2012, *4* (4), 484–487.

- (40) Díaz Velázquez, H.; Wu, Z.-X.; Vandichel, M.; Verpoort, F. Inserting CO₂ into Terminal Alkynes via Bis-(NHC)-Metal Complexes. *Catal. Lett.* 2017, 147 (2), 463–471.
- (41) Wang, X.; Lim, Y. N.; Lee, C.; Jang, H.-Y.; Lee, B. Y. 1,5,7-Triazabicyclo[4.4.0]Dec-1-Ene-Mediated Acetylene Dicarboxylation and Alkyne Carboxylation Using Carbon Dioxide. *Eur. J. Org. Chem.* 2013, 2013 (10), 1867–1871.
- (42) Cheng, H.; Zhao, B.; Yao, Y.; Lu, C. Carboxylation of Terminal Alkynes with CO₂ Catalyzed by Bis(Amidate) Rare-Earth Metal Amides. *Green Chem.* 2015, 17 (3), 1675–1682.
- (43) Dingyi, Y.; Yugen, Z. The Direct Carboxylation of Terminal Alkynes with Carbon Dioxide. *Green Chem.* 2011, 13 (5), 1275–1279.
- (44) Hong, J.; Li, M.; Zhang, J.; Sun, B.; Mo, F. C–H Bond Carboxylation with Carbon Dioxide. *ChemSusChem* 2019, 12 (1), 6–39.
- (45) Manjolinho, F.; Arndt, M.; Gooßen, K.; Gooßen, L. J. Catalytic C–H Carboxylation of Terminal Alkynes with Carbon Dioxide. *ACS Catal.* 2012, 2 (9), 2014–2021.
- (46) Shi, J.-B.; Bu, Q.; Liu, B.-Y.; Dai, B.; Liu, N. Organocatalytic Strategy for the Fixation of CO₂ via Carboxylation of Terminal Alkynes. *J. Org. Chem.* 2021, 86 (2), 1850–1860.
- (47) Wang, W.-H.; Jia, L.; Feng, X.; Fang, D.; Guo, H.; Bao, M. Efficient Carboxylation of Terminal Alkynes with Carbon Dioxide Catalyzed by Ligand-Free Copper Catalyst under Ambient Conditions. *Asian J. Org. Chem.* 2019, 8 (8), 1501–1505.
- (48) Gooßen, L. J.; Rodríguez, N.; Manjolinho, F.; Lange, P. P. Synthesis of Propiolic Acids via Copper-Catalyzed Insertion of Carbon Dioxide into the C-H Bond of Terminal Alkynes. *Adv. Synth. Catal.* 2010, 352 (17), 2913–2917.
- (49) Yamada, E. W.; Jakoby, W. B. Enzymatic Utilization of Acetylenic Compounds. II. Acetylenemonocarboxylic Acid Hydrase. *J. Biol. Chem.* 1959, 234 (4), 941–945.
- (50) Yamada, E. W.; Jakoby, W. B. Enzymatic Utilization of Acetylenic Compounds: I. An Enzyme Converting Acetylenedicarboxylic Acid to Pyruvate. *J. Biol. Chem.* 1958, 233 (3), 706–711.
- (51) Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, W. H.; Whitman, C. P. Characterization of Cg10062 from *Corynebacterium glutamicum*: Implications for the Evolution of *cis*-3-Chloroacrylic Acid Dehalogenase Activity in the Tautomerase Superfamily. *Biochemistry* 2008, 47 (31), 8139–8147.
- (52) Davidson, R.; Baas, B.-J.; Akiva, E.; Holliday, G. L.; Polacco, B. J.; LeVieux, J. A.; Pullara, C. R.; Zhang, Y. J.; Whitman, C. P.; Babbitt, P. C. A Global View of Structure–Function Relationships in the Tautomerase Superfamily. *J. Biol. Chem.* 2018, 293 (7), 2342–2357.
- (53) Poelarends, G. J.; Veetil, V. P.; Whitman, C. P. The Chemical Versatility of the β – α – β Fold: Catalytic Promiscuity and Divergent Evolution in the Tautomerase Superfamily. *Cell. Mol. Life Sci.* 2008, 65 (22), 3606–3618.

- (54) Huddleston, J. P.; Johnson, W. H.; Schroeder, G. K.; Whitman, C. P. The Accidental Assignment of Function in the Tautomerase Superfamily. *Perspect. Sci.* 2015, 4, 38–45.
- (55) Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, William H.; Murzin, A. G.; Whitman, C. P. Cloning, Expression, and Characterization of a *cis*-3-Chloroacrylic Acid Dehalogenase: Insights into the Mechanistic, Structural, and Evolutionary Relationship between Isomer-Specific 3-Chloroacrylic Acid Dehalogenases. *Biochemistry* 2004, 43 (3), 759–772.
- (56) Wang, S. C.; Person, M. D.; Johnson, William H.; Whitman, C. P. Reactions of *trans*-3-Chloroacrylic Acid Dehalogenase with Acetylene Substrates: Consequences of and Evidence for a Hydration Reaction. *Biochemistry* 2003, 42 (29), 8762–8773.
- (57) Poelarends, G. J.; Johnson, W. H.; Murzin, A. G.; Whitman, C. P. Mechanistic Characterization of a Bacterial Malonate Semialdehyde Decarboxylase: Identification of a New Activity in The Tautomerase Superfamily. *J. Biol. Chem.* 2003, 278 (49), 48674–48683.
- (58) Guo, Y.; Serrano, H.; Poelarends, G. J.; Johnson, W. H.; Hackert, M. L.; Whitman, C. P. Kinetic, Mutational, and Structural Analysis of Malonate Semialdehyde Decarboxylase from *Coryneform* Bacterium Strain FG41: Mechanistic Implications for the Decarboxylase and Hydratase Activities. *Biochemistry* 2013, 52 (28).
- (59) Huddleston, J. P.; Johnson, W. H.; Schroeder, G. K.; Whitman, C. P. Reactions of Cg10062, a *cis*-3-Chloroacrylic Acid Dehalogenase Homologue, with Acetylene and Allene Substrates: Evidence for a Hydration-Dependent Decarboxylation. *Biochemistry* 2015, 54 (19), 3009–3023.
- (60) Poelarends, G. J.; Serrano, H.; Johnson, W. H.; Whitman, C. P. Stereospecific Alkylation of *cis*-3-Chloroacrylic Acid Dehalogenase by (R)-Oxirane-2-Carboxylate: Analysis and Mechanistic Implications. *Biochemistry* 2004, 43 (22), 7187–7196.
- (61) Kapust, R. B.; Tözsér, J.; Copeland, T. D.; Waugh, D. S. The P1' Specificity of Tobacco Etch Virus Protease. *Biochem. Biophys. Res. Commun.* 2002, 294 (5), 949–955.
- (62) Wood, D. W. New Trends and Affinity Tag Designs for Recombinant Protein Purification. *Curr. Opin. Struct. Biol.* 2014, 26, 54–61.
- (63) Kapust, R. B.; Waugh, D. S. Controlled Intracellular Processing of Fusion Proteins by TEV Protease. *Protein Expr. Purif.* 2000, 19 (2), 312–318.
- (64) Kim, K.-S.; Pelton, J. G.; Inwood, W. B.; Andersen, U.; Kustu, S.; Wemmer, D. E. The Rut Pathway for Pyrimidine Degradation: Novel Chemistry and Toxicity Problems. *J. Bacteriol.* 2010, 192 (16), 4089–4102.
- (65) Guo, Y.; Robertson, B. A.; Hackert, M. L.; Whitman, C. P. Crystal structure of native Cg10062 2011, DOI: 10.2210/pdb3N4G/pdb.
- (66) Nayebi Gavgani, H. *Structural Enzymology Investigating the Mechanism of Rice Branching Enzyme I, Rice Granule Bound Starch Synthase, and CG10062, a Member of the Tautomerase Superfamily*. Ph.D., Michigan State University, United States.

- (67) Sevastik, R.; Whitman, C. P.; Himo, F. Reaction Mechanism of *cis*-3-Chloroacrylic Acid Dehalogenase: A Theoretical Study. *Biochemistry* 2009, 48 (40), 9641–9649.
- (68) Frost, J. W.; Hansen, C. A. Synthesis of 1,2,3,4-Tetrahydroxybenzenes and 1,2,3-Trihydroxybenzenes Using Myo-Inositol-1-Phosphate Synthase and Myo-Inositol 2-Dehydrogenase. US6750049B1, June 15, 2004.
- (69) Li, K.; Mikola, M. R.; Draths, K. M.; Worden, R. M.; Frost, J. W. Fed-Batch Fermentor Synthesis of 3-Dehydroshikimic Acid Using Recombinant *Escherichia coli*. *Biotechnol. Bioeng.* 1999, 64 (1), 61–73.
- (70) Miller, Kelly. *Semi-Synthesis of Aromatic Diacids and Biosynthesis of Kanosamine in Escherichia coli*. Ph.D., Michigan State University, United States.
- (71) Di Lorenzo, M.; Hidalgo, A.; Molina, R.; Hermoso, J. A.; Pirozzi, D.; Bornscheuer, U. T. Enhancement of the Stability of a Prolipase from *Rhizopus oryzae* toward Aldehydes by Saturation Mutagenesis. *Appl. Environ. Microbiol.* 2007, 73 (22), 7291–7299.
- (72) Franken, B.; Eggert, T.; Jaeger, K. E.; Pohl, M. Mechanism of Acetaldehyde-Induced Deactivation of Microbial Lipases. *BMC Biochem.* 2011, 12 (1), 10.
- (73) Dennis, R. T.; Young, T. W. A Simple, Rapid Method for the Detection of Subspecies of *Zymomonas mobilis*. *J. Inst. Brew.* 1982, 88 (1), 25–29.
- (74) Lillie, R. D.; Conn, H. J.; Biological Stain Commission. *H.J. Conn's Biological Stains: A Handbook on the Nature and Uses of the Dyes Employed in the Biological Laboratory*; Williams & Wilkins: Baltimore, 1977.
- (75) Conway, T.; Sewell, G. W.; Osman, Y. A.; Ingram, L. O. Cloning and Sequencing of the Alcohol Dehydrogenase II Gene from *Zymomonas mobilis*. *J. Bacteriol.* 1987, 169 (6), 2591–2597.

CHAPTER TWO: SYNTHESIS OF 3-HYDROXYPROPIONIC ACID

Reprinted with permission from Draths, K.; Geiger, J.; Mathes Hewage, A. N. S.; Gavgani, H. N. Synthesis of 3-Hydroxypropionic Acid via Hydration of Acetylenecarboxylic Acid. Patent pending, filed on April 21, 2022. Manuscript in preparation.

2.1. Introduction

In pursuit of a suitable C3 target molecule that can be produced from C1-derived ACA, 3-hydroxypropionic acid (3-HP) emerged as the top candidate. 3-HP is an achiral, 3-carbon molecule consisting of a β -hydroxyl moiety and carboxylic acid group. The functionality of this molecule makes it a versatile building block as it opens up numerous routes to a variety of chemicals that have well-established markets and are already being produced on an industrial scale. Since 2004, the number of studies highlighting the synthesis of 3-HP via different biocatalytic syntheses have been on the rise.¹ This surge was driven by a comprehensive report authored by a U.S. DOE-funded collaboration between the Pacific Northwest National Laboratory (PNNL), National Renewable Energy Laboratory (NREL) and the Office of Biomass Program (EERE).² The report consists of a thorough examination of the top value-added chemicals to be targeted for a two-step process: production from biomass via chemical or biological means, and subsequent conversion to a variety of high-value biobased commodity chemicals. The motivation behind this collaborative program's extensive study was to promote and examine the feasibility of integrated biorefineries in which biomass would be processed and converted to the target molecules and their downstream conversion to commodity chemicals or materials would be carried out within the same facility.² Starting with over 300 potential candidates (C1-C6 molecules), the final selection of target molecules were narrowed down to a list of twelve, in which 3-HP ranked third.² Several screening criteria were considered when compiling the list, most of which are associated with the economic feasibility of implementing new and competitive pathways to the targets relative to the existing petrochemical-based counterparts. Some of these criteria included the cost of the raw materials and the processing costs associated with obtaining

feedstocks necessary for microbial catalysis and the market potential for each molecule with respect to existing products on the market.²

3-HP stands out as a platform chemical with a diverse portfolio of secondary commodity chemicals it can be converted to including (but not limited to) acrylic acid, 1,3-propanediol, methyl acrylate and acrylamide, all of which currently have commercial uses and are produced industrially on a commodity-scale (Figure 2.1).² These secondary chemicals are used for the synthesis of a variety of polymers, resins, adhesives and coatings. 3-HP itself can also serve as a monomer for polymerization reactions. The resulting polymer from the self-condensation of 3-HP displays biodegradable properties and is biocompatible, which may allow its use in the manufacture of surgical materials as well as for applications in *in vivo* drug-release.^{3,4} Perhaps the most significant commodity chemical target from 3-HP is acrylic acid, with a global demand that continues to grow each year, with approximately 6.2 million metric tons consumed in 2020.⁵ Industrially, acrylic acid is primarily obtained from gas-phase catalytic oxidation of petroleum-based propylene via an acrolein intermediate.⁶ The oxidation of propylene is a two-step process that requires operation at different temperatures in order to obtain high efficiency and yields.⁶ Acrylic acid is primarily converted to acrylate esters which are readily polymerized for applications in industrial coatings, paints, adhesives and binders for the leather, paper and textile industries.⁶ Additionally, super absorbent polymers derived from the acrylate family are used in diapers and hygiene products.² Furthermore, methyl acrylate is used as a co-polymer for the production of polyacrylonitrile.⁶ Despite being an economically competitive process, the production of acrylic acid via propylene is not without its drawbacks. In addition to use of a non-renewable precursor, the acrolein intermediate in acrylic acid synthesis is toxic, begging the need for an alternative route.⁷ Due to significant efforts being made to transition from the current fossil-based economy to a greener, more sustainable bioeconomy, 3-HP has garnered the attention of several companies and academic researchers as a more suitable bio-based precursor for acrylic acid. In 2008, Cargill and Novozymes announced a joint venture for the development and

commercialization of bio-based acrylic acid via dehydration of 3-HP obtained from fermentation of renewable sugars.⁸ The microbe genetically engineered for the production of 3-HP from glucose was initially developed by OPX Biotechnologies and later acquired by Cargill.⁹ Chemical giant BASF, the world's largest producer of acrylic acid, joined the DOE-funded project in 2012 to contribute towards the downstream manufacture of superabsorbent polymers from the bio-based acrylic acid.¹⁰ In 2013, the industrial partners demonstrated successful pilot scale production of 3-HP via microbial catalysis and production of superabsorbent polymers the following year.^{10,11} Although BASF exited the R&D collaboration shortly after due to undisclosed reasons, the partnership between Novozymes and Cargill continues.^{11,12} More recently, Proctor & Gamble developed new technology to produce bio-based acrylic acid from lactic acid (2-hydroxypropionic acid) and partnered with Cargill for commercialization of this technology.¹³ Another notable collaboration between NREL and Cargill highlights the manufacture of renewable carbon fiber from biomass.^{14,15} The DOE-funded program focused on the production of bio-based acrylonitrile from 3-HP.¹⁵ The resulting acrylonitrile is polymerized and can be used to produce carbon fiber, a high-strength, lightweight material that has applications in the automotive, aerospace, and renewable energy industries.¹⁶⁻¹⁸ It can be used to replace steel and heavier metals and improve overall costs and fuel-efficiency of vehicles or be used for the manufacture of renewable energy components such as wind-turbine blades.¹⁶⁻²⁰

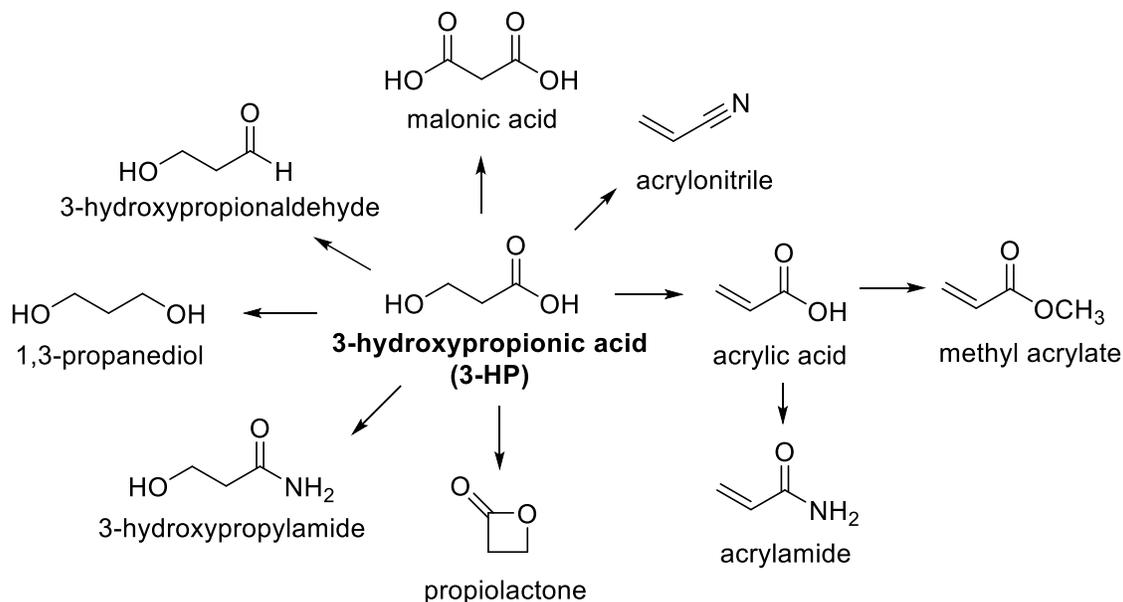


Figure 2.1. Secondary chemicals that can be produced from the 3-HP platform chemical

2.2. Synthetic routes to 3-HP

Despite its high potential for commercialization in diverse markets, an industrially-relevant technology for the synthesis of 3-HP has not been established thus far. 3-HP is an expensive chemical that can only be acquired as an aqueous solution (30% w/v) from a limited number of manufacturers globally.²¹ The hydrophilicity of 3-HP makes its extraction from aqueous solution challenging and distillation of 3-HP results in dehydration of 3-HP to form acrylic acid.^{22–24} Traditionally, 3-HP is prepared by the hydration of acrylic acid under alkaline conditions or by hydrolysis of β -propiolactone, which is readily available.^{21,25} Other notable chemical routes to 3-HP include Pd-catalyzed oxidation of 1,3-PDO to 3-HP under alkaline conditions, patented by Henkel.^{21,26} All currently existing chemical syntheses are stoichiometric conversions or metal-catalyzed transformations of value-added chemicals to 3-HP, begging the need for industrially viable syntheses that can be executed on a large scale that are both economically feasible and environmental-friendly. Driven by the list of the top 12 key building block chemicals published by the U.S. DOE,² countless research efforts and resources that have been allocated towards 3-HP biosynthesis using naturally existing metabolic pathways or genetically modified organisms using

a variety of carbon feedstocks have been reviewed.^{4,27–32} Across scientific literature examining the production of 3-HP via biocatalysis, the two most common renewable feedstocks utilized are glycerol and glucose.^{4,7,27,29,33–36}

2.2.1. 3-HP production from glycerol

Glycerol is an abundant and renewable carbon feedstock that is a by-product in the biodiesel industry, with approximately 100 kg of crude glycerol generated per ton of biodiesel produced.³⁷ It is also a main by-product in bioethanol and soap manufacturing processes.^{28,38} Glycerol is considered a waste product and its non-toxic nature and availability makes it an attractive substrate for the biosynthesis of 3-HP as it can be used to achieve high titers.²⁹ 3-HP can be accessed from glycerol via two different pathways: CoA-dependent and CoA-independent pathways.

The oxidation of glycerol **25** for 3-HP production via the CoA-dependent pathway (Figure 2.2) is based on the metabolic pathway for 1,2-propanediol degradation native to *Salmonella*, *Klebsiella* and *Lactobacillus* species and is governed by enzymes in the *pdu* (propanediol utilization) operon, allowing them to produce 3-HP **24** naturally.^{39,40} In the first step of the pathway, the formation of 3-hydroxypropionaldehyde (3-HPA) **26** from glycerol is catalyzed by coenzyme B₁₂-dependent glycerol dehydratase (GDHt).⁴¹ The 3-HPA **26** is then converted to 3-hydroxypropionyl-CoA (3-HP-CoA) **27** by propionaldehyde dehydrogenase (PduP).⁴⁰ The resulting 3-HP-CoA **27** is phosphorylated by phosphate propanoyltransferase (PduL) and subsequently dephosphorylated by propionate kinase (PduW) to form 3-HP **24**.⁴² 3-HP **24** synthesis has been demonstrated in several strains such as *L. reuteri* and *K. pneumoniae* using the Pdu pathway.^{7,43,44} These strains are economically attractive due to their native ability to synthesize coenzyme B₁₂, as opposed to recombinant *E. coli* strains which require exogenous addition of the coenzyme.^{28,45,46} Nevertheless, the production of high concentrations of 3-HP **24** via this pathway is limited by the accumulation of 3-HPA **26**, which can be toxic to cells.¹ To address this limitation, a microbial consortium of *L. reuteri* producing 3-HPA **26** was co-cultured

with recombinant *E. coli* overexpressing an aldehyde dehydrogenase (ALDH) for reduction of 3-HPA **26** to 3-HP **24** was demonstrated, achieving titers of 125 g L⁻¹ 3-HP **24** and 88 g L⁻¹ 1,3-PDO **29**.⁴⁶ Although certain species carrying the native *pdu* operon are able to produce 3-HP **24** naturally, the 3-HPA **26** intermediate can be reduced to 1,3-propanediol (1,3-PDO) **29** by 1,3-propanediol reductase, resulting in equimolar quantities of 3-HP **24** and 1,3-PDO **29**, thus reducing the maximum theoretical yield of 3-HP **24** from glycerol **25**.^{7,28} On the other hand, the co-production of 1,3-PDO **29** and 3-HP **24** has been promoted as a means of addressing the requirement of NAD⁺ for the conversion of 3-HPA **26** to 3-HP-CoA **27** in the Pdu pathway, since NAD⁺ can be regenerated by 1,3-PDO reductase (Figure 2.2).^{7,47}

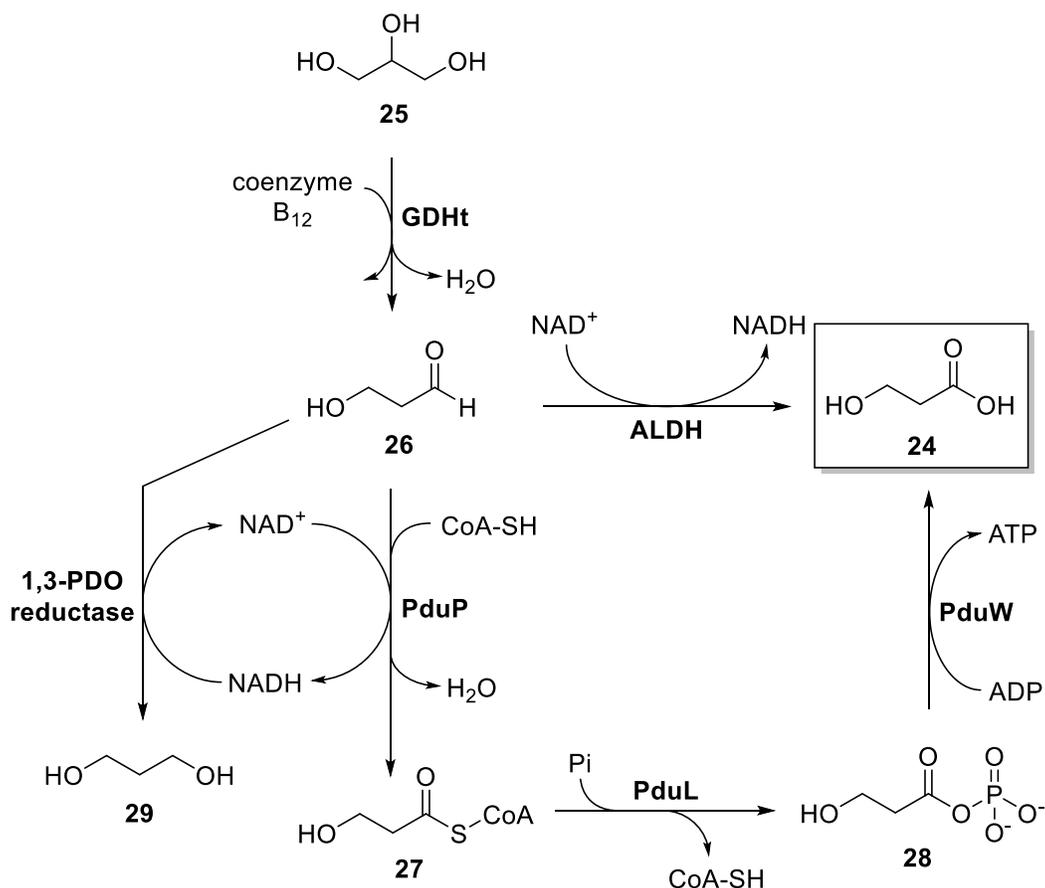


Figure 2.2. 3-HP synthesis from glycerol via CoA-dependent and CoA-independent pathways. GDHt, glycerol dehydratase; PduP, propionaldehyde dehydrogenase; PduL, phosphate propanoyltransferase; PduW, propionate kinase; ALDH, aldehyde dehydrogenase.

Alternatively, 3-HP **24** production has also been demonstrated via the relatively simpler CoA-independent pathway (Figure 2.2). In this two-step pathway that has been studied more extensively, glycerol **25** is converted by 3-HPA **26** via coenzyme B₁₂-dependent glycerol dehydratase (GDHt) and subsequently oxidized to 3-HP **24** by an aldehyde dehydrogenase (ALDH), a ubiquitous enzyme present in most microbes.^{4,31,44,48,49} 3-HP **24** production of up to 102 g L⁻¹ was achieved using *K. pneumoniae*, a microbe capable of naturally synthesizing coenzyme B₁₂, thereby contributing to lowered production costs.⁴⁴ This pathway is favored over the CoA-dependent pathway since it allows the production of 3-HP **24** using overexpression of ALDH and does not require participation of CoA.⁴ In both CoA-dependent and CoA-independent pathways, maintaining a redox-balance and ensuring an adequate supply of NAD⁺ within the cells are important considerations for efficient *in vivo* synthesis of 3-HP **24**. NAD⁺ can be regenerated by oxidative metabolism of glucose or in the electron transport chain, by increased aeration of cell cultures. However, delicate optimization strategies are required due to inactivation of coenzyme B₁₂ by oxygen.^{50,51} An additional hurdle of utilizing glycerol **25** as a carbon source for 3-HP **24** production is the inactivation of GDHt by its substrate glycerol **25**.^{51,52} The suicide-based inactivation of GDHt occurs during the catalytic cycle, where coenzyme B₁₂ is not regenerated due to the formation of a catalytically inactive form of the coenzyme that remains bound to the enzyme, preventing any further interaction with free coenzyme B₁₂.⁵³ Nevertheless, studies published by Toraya and coworkers show that reactivation is achieved in the presence of ATP and Mg²⁺.^{51,53-55}

2.2.2. 3-HP production from glucose

Although 3-HP synthesis via the abundant and inexpensive substrate glycerol is the most extensively studied, glucose can also be used as a starting material for 3-HP synthesis via two main pathways: malonyl-CoA and β -alanine pathways.³¹ In the malonyl-CoA pathway, acetyl-CoA **30** is produced from the transformation of glucose via glycolysis. Instead of entering the tricarboxylic acid (TCA) cycle, acetyl-CoA **30** is converted to malonyl-CoA **31** by acetyl-CoA

carboxylase (ACC). Malonyl-CoA **31** is subsequently converted to 3-HP **24** via intermediate MSA **13** by NADPH-dependent malonyl-CoA reductase (MCR) (Figure 2.3).³¹ The microbial strain capable of 3-HP **24** production from glucose, developed by OPX Biotechnologies and now owned by Cargill, utilizes the malonyl-CoA pathway.^{4,31} ACC is commonly found in microorganisms and is involved in fatty acid metabolism.³¹ MCR, first discovered in *Chloroflexus aurantiacus*, has two catalytic domains that allow it to convert malonyl-CoA **31** to 3-HP **24** via two steps. The C-terminal functional domain catalyzes the formation of MSA **13**, while the N-terminal domain catalyzes reduction of MSA **13** to 3-HP **24**.^{1,56} Fragmentation of MCR yielded two dissected proteins with improved kinetics for 3-HP **24** synthesis.^{57,58} However, the requirement of two equivalents of NADPH for every mole of 3-HP **24** consumed has been identified as a significant limitation of using this pathway.^{4,31}

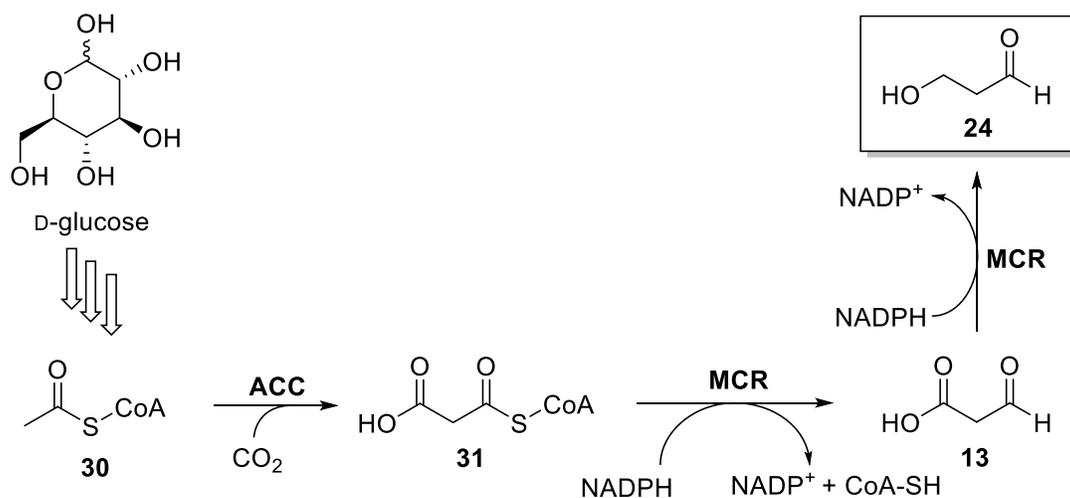


Figure 2.3. 3-HP synthesis from glucose via the malonyl-CoA pathway. ACC, acetyl-CoA carboxylase; MCR, malonyl-CoA reductase.

The β -alanine pathway shares the same immediate precursor to 3-HP, MSA, as the malonyl-CoA pathway. However, in contrast, the MSA **13** in the β -alanine pathway is derived from the TCA cycle.^{4,27,31} The main precursor in the pathway, aspartate **33** is formed from the aspartate aminotransferase-catalyzed conversion of fumarate **32** (Figure 2.4). β -Alanine **34** is then formed

from the aspartate decarboxylase-catalyzed decarboxylation. The resulting β -alanine **33** is transformed into MSA **13** by β -alanine pyruvate transaminase. The final step for the formation of 3-HP **24** from MSA **13** is catalyzed by MCR or a malonic semialdehyde reductase.^{27,31,36} Song et. al reported the highest production of 3-HP **24** (31.1 g L⁻¹) using this pathway in a recombinant strain of *E. coli*.³⁶

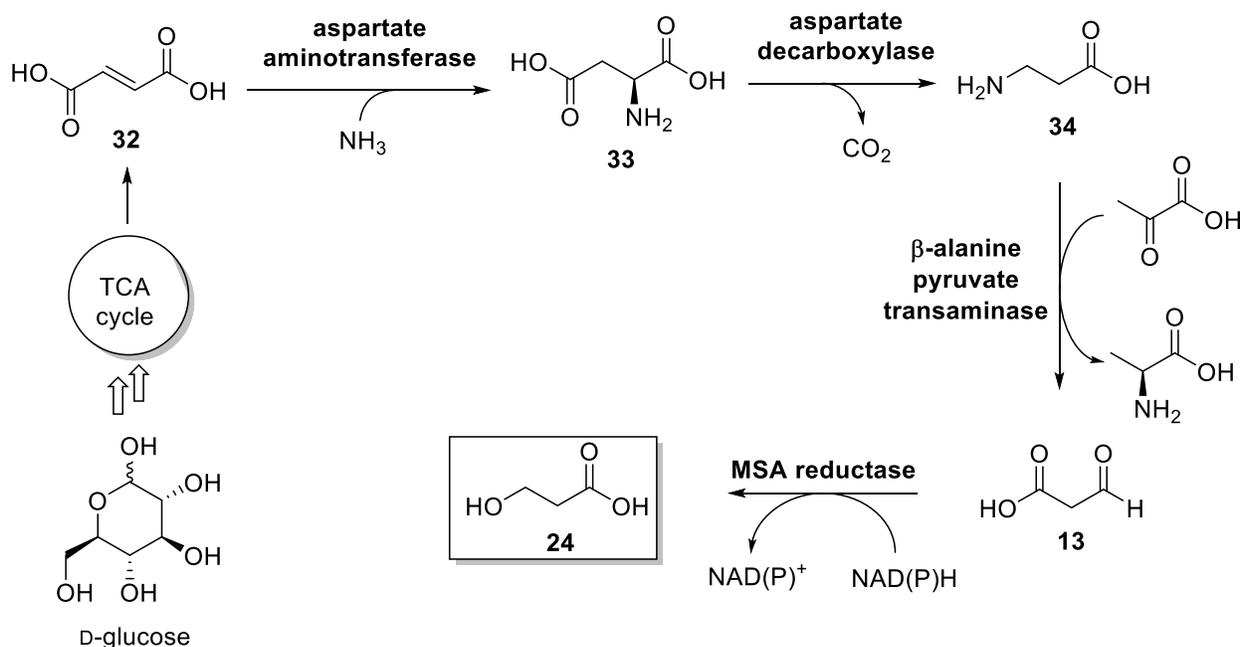


Figure 2.4. Synthesis of 3-HP from glucose via the β -alanine pathway

2.2.3. Other biocatalytic routes to 3-HP

Though most metabolic engineering for 3-HP production is carried out using glycerol or glucose substrates, other less common pathways have also been explored for 3-HP biosynthesis. Some autotrophic producers such as *Chloroflexus aurantiacus* produce 3-HP natively from CO₂ fixation, via the 3-HP cycle, where acetyl-CoA **30** is transformed to 3-HP **24** via the malonyl-CoA.⁵⁹ However, this organism has not been engineered for 3-HP production. Heterologous expression of the malonyl-CoA reductase from *C. aurantiacus* in a *Synechococcus* sp. enabled

photosynthetic production of low levels of 3-HP (0.84 g L^{-1}) production from CO_2 .⁶⁰ In 2021, Jiang et. al reported a titer of 154 g L^{-1} 3-HP from the fermentation of *Halomonas bluephagenesis* using 1,3-PDO as a substrate.⁶¹ While the high titers of 3-HP is commendable, the use of a value-added chemical such as 1,3-PDO, where its own convenient industrial synthesis remains a challenge,²¹ to produce a building block chemical such as 3-HP takes away from its usefulness in an industrial setting. The highest titer of 184 g L^{-1} 3-HP to date was reported by Yu et. al using immobilized cells of recombinant *E. coli* and substrate 3-hydroxypropionitrile.⁶² A comprehensive list of 3-HP syntheses using various substrates as well as their titers, productivities and yields are discussed in recent reviews.^{4,27,31,32}

2.3. Synthesis of 3-HP from ACA

Previously, 3-HP biosynthesis has been demonstrated using two main substrates: glycerol and glucose. Herein, ACA is presented as a novel substrate for 3-HP synthesis. As described in Chapter One, ACA can be derived via CH_4 dehydrodimerization and subsequent carboxylation of acetylene. To our knowledge, this research is the first example of ACA utilization for 3-HP synthesis and has previously not been used as a substrate for microbial catalysis. The enzyme evolution efforts with Cg10062 revealed a new variant Cg10062(E114N), which within the levels of detection produces exclusively MSA from ACA. Relative to the existing biocatalytic pathways described earlier in this chapter, the combination of ACA and Cg10062(E114N) is an attractive pair for 3-HP synthesis given that ACA exists as a stable, liquid that can be derived from C1 greenhouse gases (see chapter one) and Cg10062(E114N) does not require any metal cofactors or coenzymes for catalysis. This research demonstrated efficient conversion of ACA to MSA *in vitro* and coupling with a second enzyme such as an MSA reductase would result in a simpler, efficient two-step pathway to 3-HP that has not been described previously.

In order to be able to utilize ACA for 3-HP synthesis, it was necessary to identify a suitable enzyme for the reduction of the unstable MSA intermediate. YdfG, a short-chain dehydrogenase/reductase from *E. coli*, previously used by others for *in vivo* synthesis of 3-HP via

the β -alanine route³⁶ in *E. coli* was identified as the first candidate. YdfG is a homo-tetrameric protein consisting of a 27 kDa monomeric subunit (248 amino acids).⁶³ YdfG was biochemically characterized by Fujisawa et. al for the oxidation of a broad number of 3-hydroxy acids including L-serine, L-threonine, and 3-HP.⁶³ The group also studied a second enzyme YMR226C, a homo-tetramer with a 29 kDa monomeric subunit (267 amino acids), from *Saccharomyces cerevisiae*, for dehydrogenase activity with the same set of 3-hydroxy acid substrates.⁶³ Both enzymes displayed similar activity for the oxidation of L-serine and the relative oxidative activities of both enzymes with all substrates tested are shown below (Figure 2.5). YdfG demonstrated superiority in the oxidation of 3-HP to MSA relative to YMR226C; 23% vs. 13 % relative activity, respectively, where 100% relative activity was observed with L-serine for both enzymes.⁶³

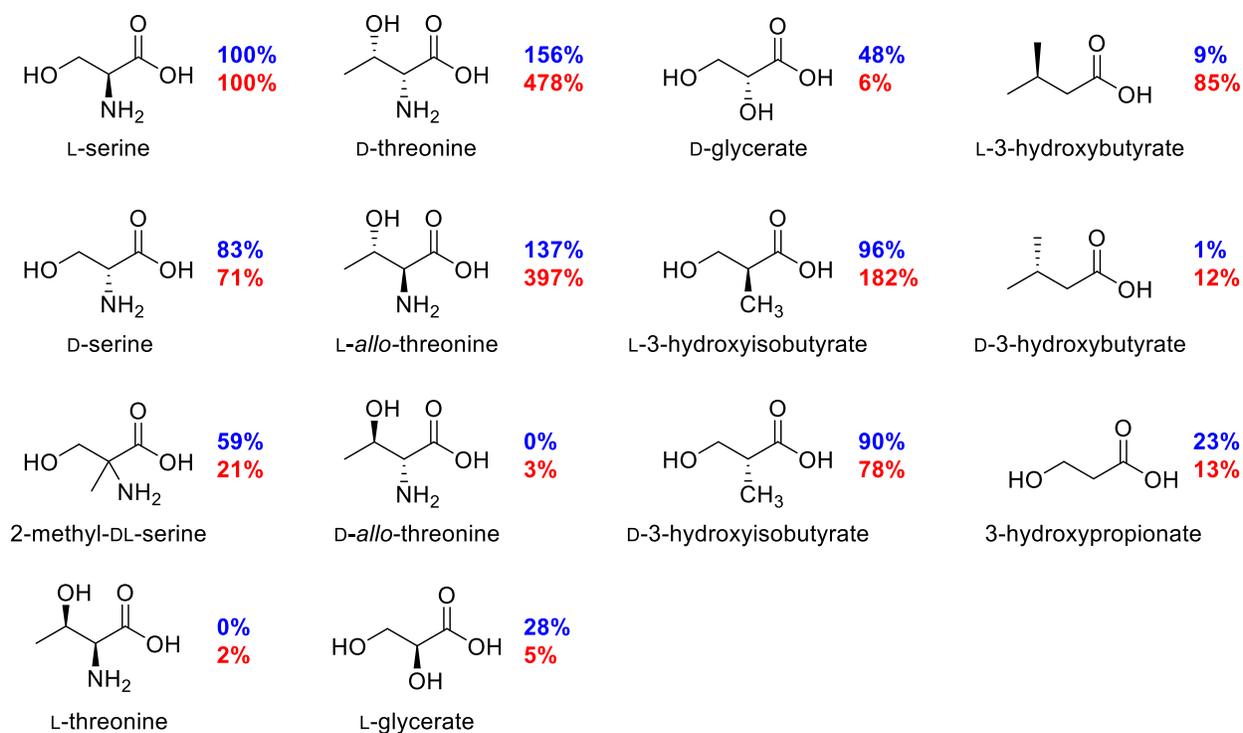


Figure 2.5. Substrate specificity of YdfG and YMR226C. Activities of YdfG (blue) and YMR226C (red) are indicated relative to the activity measured for L-serine with both enzymes.

A second potential candidate, RutE (196 amino acids, 21.6 kDa) was also examined for reduction of MSA. The Rut pathway in *E. coli* consists of seven proteins required for growth on uracil as the sole source of nitrogen.⁶⁴ In 2006, Loh et al. utilized ¹⁴C-labelled uridine to demonstrate that a C3 compound from the pyrimidine ring was excreted by *E. coli* K-12 using the Rut pathway. The identity of the final waste product of this pyrimidine catabolic pathway was confirmed to be 3-HP by mass spectrometry.⁶⁴ Kim et al. conducted an extensive study to characterize each enzyme in the Rut pathway and identified MSA as an intermediate in the Rut pathway, prior to its reduction to 3-HP, after which 3-HP is excreted from the cells.⁶⁵ In the same study, strains with mutations in the *ydfG* gene, encoding the known short-chain dehydrogenase YdfG, failed to grow on uridine as the sole source of nitrogen, indicating that the accumulation of MSA may be toxic to *E. coli*.⁶⁵ Furthermore, *E. coli* strains with mutations in the *rutE* locus initially failed to grow on uridine as the sole nitrogen source. However, whole-genome sequencing was used to identify mutations in the genome that suppressed the loss of RutE and restored growth on uridine. These mutations led to increased production of the flavoprotein NemR, an *N*-ethylmaleimide reductase and presumably acted as a substitute for RutE (in the strain with *rutE* mutations), restoring growth on uridine. When the same mutations (that led to increased NemR) were introduced into a $\Delta ydfG$ strain, growth on uridine was restored, suggesting that NemR also substituted for YdfG.⁶⁵ Therefore, it was postulated that RutE shares the same catalytic function as the known 3-hydroxy acid dehydrogenase YdfG. However, RutE was not further biochemically characterized to confirm its reductase activity on MSA. Therefore, in this study, in addition to YdfG, RutE was also examined for the reduction of MSA to 3-HP.

2.3.1. Characterization of RutE and YdfG for MSA reduction to 3-HP

Both *ydfG* and *rutE* were amplified from *E. coli* W3110 and expressed as His₆-tagged proteins using a T7 promoter system. Overexpression enabled detailed examination of YdfG and RutE for their ability to catalyze the reduction of MSA to 3-HP. The canonical Tobacco Etch Virus (TEV) protease recognition sequence ENLYFQG is recognized and cleaved by TEV protease

between the glutamate and glycine residues and is used extensively for the generation of the target protein with the native amino acid sequence.⁶⁶ A lack of stringency for the glycine residue allowed the use of a N-terminal His₆-tag followed by the recognition sequence ENLYFQM (Figure 2.6). Incubation of His₆-tagged YdfG and RutE with TEV protease enabled cleavage between the glutamate and methionine residues of the recognition sequence resulting in the target proteins with the native amino acid sequences. The plasmid pMHTΔ238 expresses a His₇-tagged TEV protease.⁶⁷ This provided the flexibility of a second round of affinity purification using a nickel affinity column to separate the native YdfG and RutE from the TEV protease mixture, where the YdfG and RutE were isolated from the flowthrough.

(A) 5' – TAATACGACTCACTATAGG GGAATTGTGAGCGGATAACAATTCCCTCTAGA
 AATAATTTTGTTTAACTTTAAGAAGGAG ATATACCATGGGCAGCAGC CATCATCAT
 CATCATCAC AGCAGCGGCCTGGTGCCGCGCGGCAGCCAA GAAAACCTGTATTT
 TCAGATGATCGTTTTAGTAACTGGAGCAACGGCAGGTTTTGGTGAATGCATTAC
 TCGTCGTTTTATTCAACAAGGGCATAAAGTTATCGCCACTGGCCGTCGCCAGGA
 ACGGTTGCAGGAGTTAAAAGACGAACTGGGAGATAATCTGTATATCGCCAACT
 GGACGTTGCAACCGCGCCGCTATTGAAGAGATGCTGGCATCGCTTCTGCCG
 AGTGGTGCAATATTGATACCTGGTAAATAATGCCGGCCTGGCGTTGGGCATGG
 AGCCTGCGCATAAAGCCAGCGTTGAAGACTGGGAAACGATGATTGATACCAAC
 AACAAAGGCCTGGTATATATGACGCGCGCCGCTTACCGGGTATGGTTGAACGT
 AATCATGGTCATATTATTAACATTGGCTCAACGGCAGGTAGCTGGCCGTATGCC
 GGTGGTAACGTTTACGGTGCGACGAAAGCGTTTGTTCGTGAGTTTAGCCTGAA
 TCTGCGTACGGATCTGCATGGTACGGCGGTGCGCGTCACCGACATCGAACCG
 GGTCTGGTGGGTGGTACCGAGTTTTCCAATGTCCGCTTTAAAGGCGATGACGG
 TAAAGCAGAAAAAACCTATCAAAAATACCGTTGCATTGACGCCAGAAGATGTCAG
 CGAAGCCGTCTGGTGGGTGTCAACGCTGCCTGCTCACGTCAATATCAATACCC
 TGGAAATGATGCCGGTTACCCAAAGCTATGCCGGACTGAATGTCCACCGTCAG
 TAA– 3'

(B) 5' – TAATACGACTCACTATAGG GGAATTGTGAGCGGATAACAATTCCCTCTAGA
 AATAATTTTGTTTAACTTTAAGAAGGAG ATATACCATGGGCAGCAGC CATCATCAT
 CATCATCAC AGCAGCGGCCTGGTGCCGCGCGGCAGCCAA GAAAACCTGTATTT
 TCAGATGAACGAAGCCGTTAGCCCAGGTGCGCTTAGCACCCCTGTTACCGGATG
 CCCGCACTCACAACGGCTGGCGGGAGACACCCGTCAGCGATGAGACGTTACG
 GGAGATTTATGCCCTGATGAAATGGGGGCCGACATCAGCTAACTGTTCTCCGG
 CACGGATCGTGTTTACCCGCACGGCAGAAGGAAAAGAACGTCTGCGCCCGGC
 ACTTTCCAGCGGCAATCTGCAAAAAACCCTGACCGCGCCCGTCACCGCTATCG
 TCGCCTGGGACAGTGAATTTTATGAACGGTTACCACTACTGTTTCCCCACGGTG
 ATGCCCGCAGTTGGTTTACCTCCAGCCCACTTGCCGAAGAAACAGCGTTT
 CGCAACAGTTCCATGCAGGCGCCTATCTGATCGTCGCCTGCCGGCGCTGG
 GACTGGATAACCGCCCGATGTGCGGGCTTTGACCGTCAACACGTGGACGACGC
 CTTTTTACGGGCAGCACGCTGAAGAGCAATCTGCTGATTAATATCGGCTATGG
 CGATAGCAGCAAGCTTTATGCGCGCCTGCCACGTCTGTCTTTGAAGAAGCCT
 GCGGGCTGTTGTAA-3'

Figure 2.6. Nucleotide sequences of (A) YdfG and (B) RutE. Various elements have been highlighted: P_{T7} (blue); lac operator (light grey); RBS (yellow); His₆-tag (green); modified TEV site (magenta); Met-1 of native protein (dark grey).

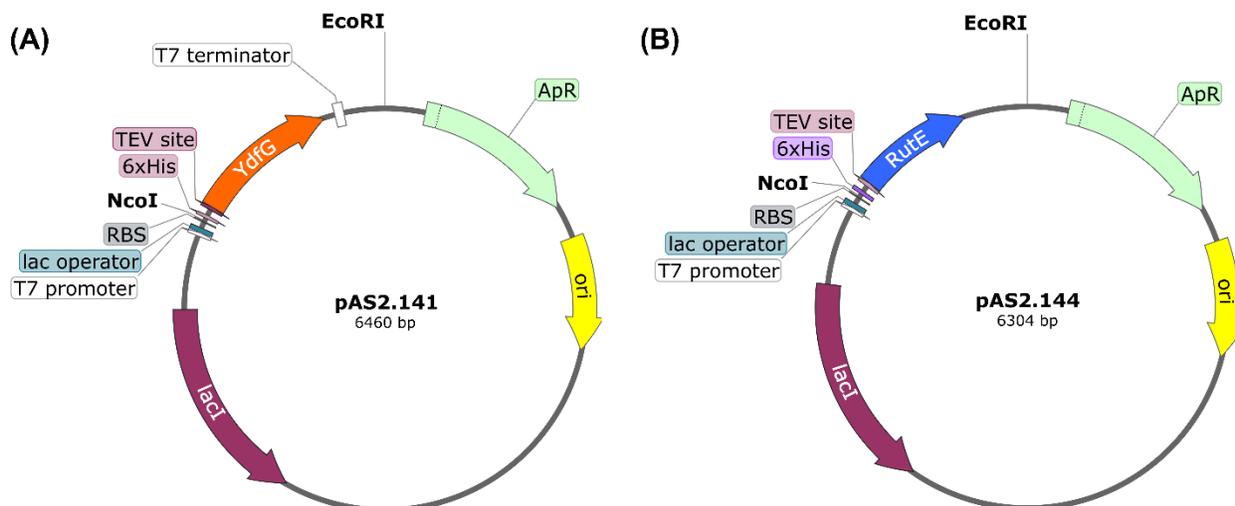


Figure 2.7. Plasmids used for the expression of His₆-tagged (A) YdfG and (B) RutE.

His₆-tagged YdfG and RutE (Figures 2.8–2.9) were purified using affinity chromatography from cell lysates of BL21(DE3)/pAS2.141 and BL21(DE3)/pAS2.144 (Figure 2.7) cultures induced with 1 mM IPTG at 37 °C. The protein fractions were pooled and concentrated in 50 mM sodium phosphate pH 8 with 0.5 mM EDTA and 1 mM DTT and incubated with His₇-tagged TEV protease (YdfG/RutE: TEV molar ratio = 1:50) at 4 °C for 12 h. The mixture of proteins was re-purified using nickel affinity chromatography to isolate the native YdfG and RutE proteins from His₇-tagged TEV protease (Figures 2.10–2.11). The flowthrough fractions containing YdfG and RutE were concentrated in 100 mM sodium phosphate pH 8. The amount of each protein and their specific activities from the purification of a 1 L of BL21(DE3)/pAS2.141 and BL21(DE3)/pAS2.144 with IPTG induction (1 mM) at 37 °C is summarized in Table 2.1. It must also be noted that the yield of RutE was significantly lower than that of YdfG, despite utilizing the strong T7 promoter system for expression of both enzymes. Previous studies on the Rut pathway demonstrated that the rut operon allows growth on pyrimidines at room temperature but not at the optimum temperature for *E. coli* growth (37 °C).^{64,65} Lowering the temperature to 25 °C for growth and IPTG induction did not result in improvements in RutE expression as indicated by chromatogram during purification.

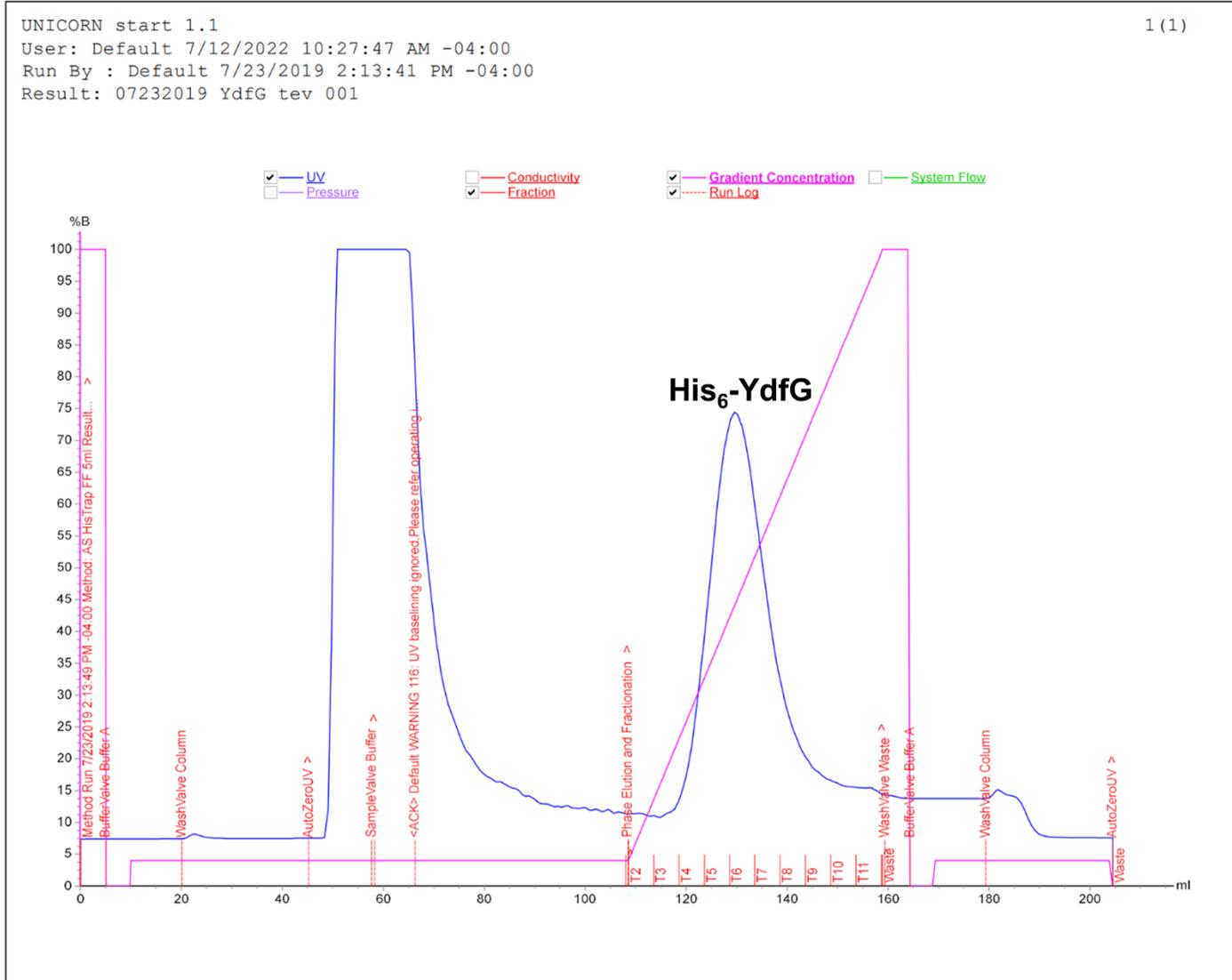


Figure 2.8. Purification of His₆-tagged YdfG from BL21(DE3)/pAS2.141 (1 L) by affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

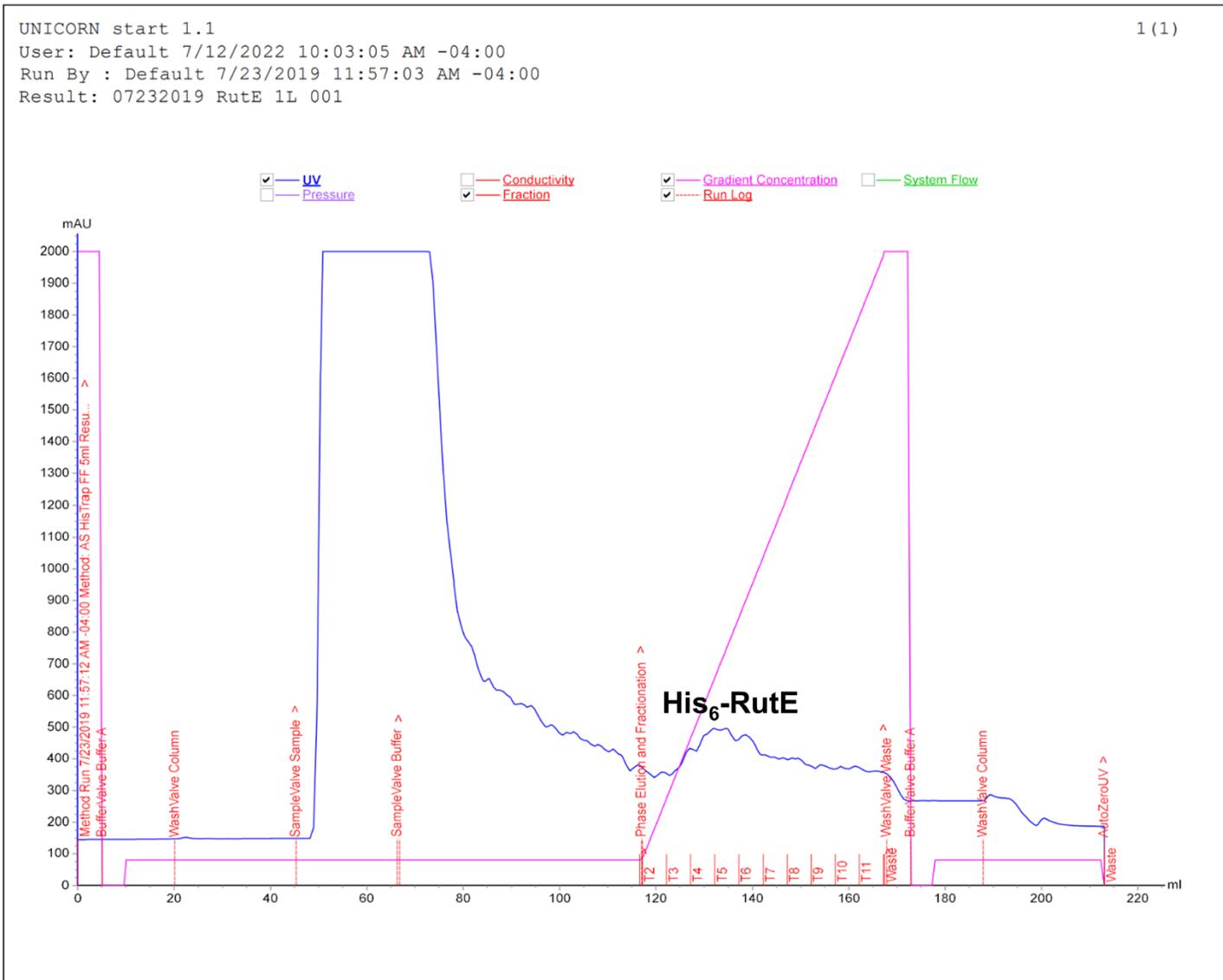


Figure 2.9. Purification of His₆-tagged RutE from BL21(DE3)/pAS2.144 (1 L) by affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

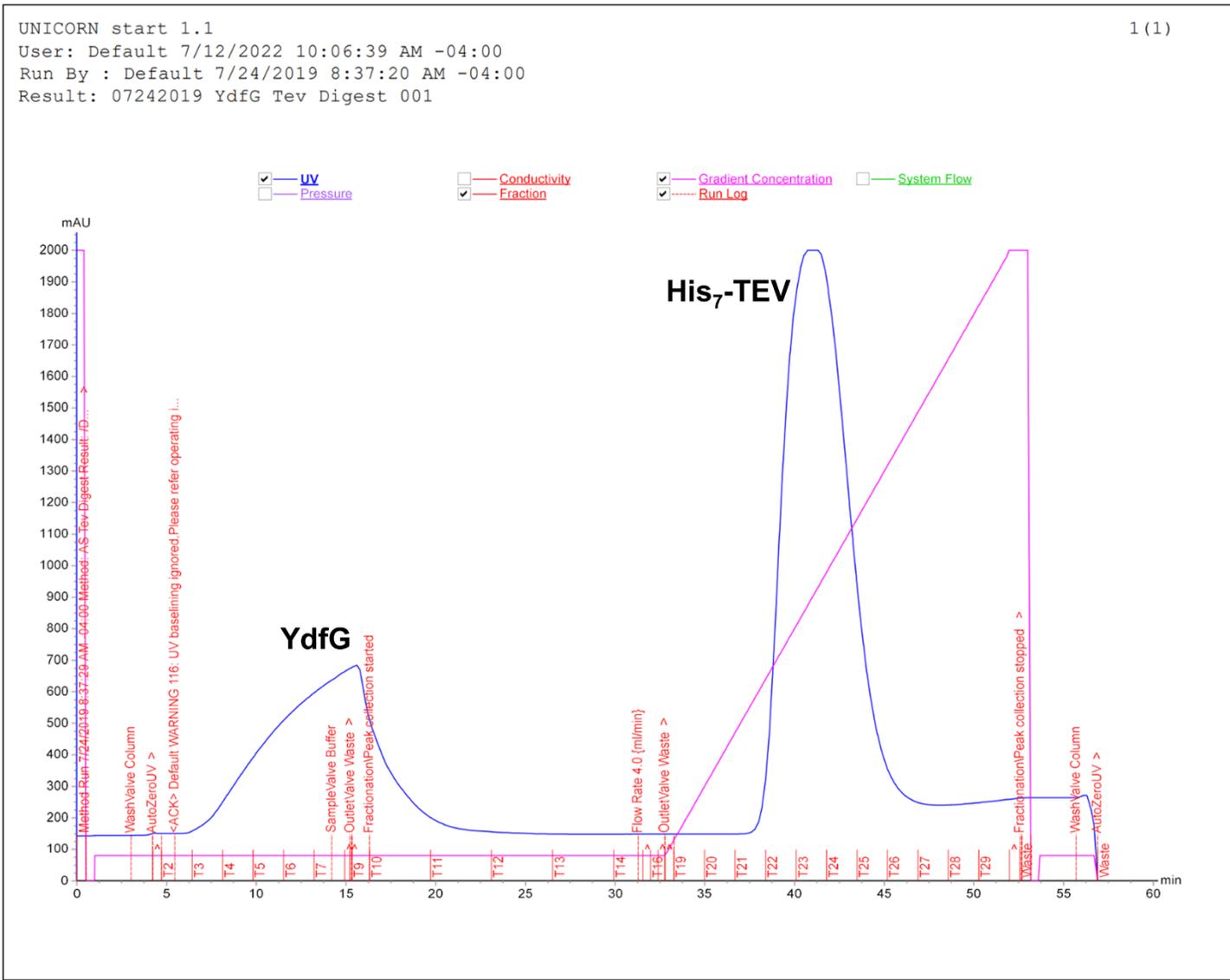


Figure 2.10. Isolation of YdfG (flowthrough) from His₇-tagged TEV protease via affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

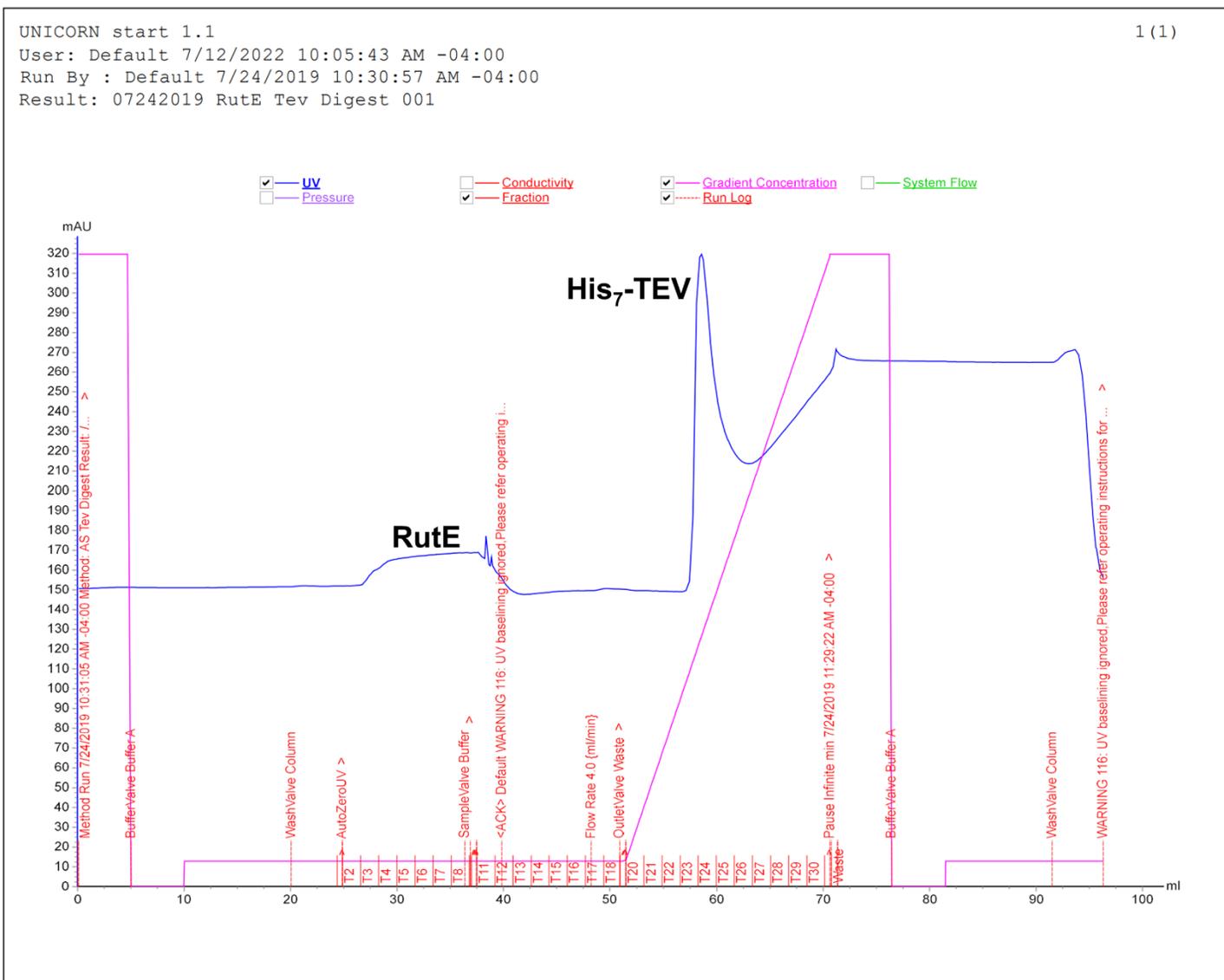


Figure 2.11. Isolation of RutE (flowthrough) from His₇-tagged TEV protease via affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

Furthermore, subjecting both enzymes to a TEV digest led to a significant decrease in yields after the second affinity purification. The activity of each enzyme was monitored using the coupled enzyme assay shown below (Figure 2.12). To circumvent difficulties associated with MSA instability, the substrate was generated *in situ* from *cis*-3-chloroacrylate (cCA) using *cis*-CaaD (see chapter one), where YdfG and RutE were each expected to produce 3-HP using an equivalent of NADPH. The decrease in absorbance at 340 nm, corresponding to NADPH oxidation was monitored to quantify enzyme activity in 100 mM sodium phosphate, pH 8. However, under these conditions, NADPH oxidation was only observed in assays containing YdfG and no activity was detected in the presence of RutE (Table 2.1). A 10-fold increase in the RutE concentration (0.005 mg mL⁻¹ to 0.05 mg mL⁻¹) did not result in detectable activity. Due to the lack of observable activity and poor expression of RutE, the decision was made to continue with YdfG utilization for 3-HP synthesis.

Table 2.1. Biochemical characterization of YdfG and RutE for reduction of MSA

Enzyme	Total Protein (mg) ^a	Specific Activity (U mg ⁻¹) ^b
His ₆ -YdfG	87	-
YdfG	34	30
His ₆ -RutE	3	-
RutE	1	ND ^c

^aThe total protein purified from 1 L of BL21(DE3)/pAS1.241 and BL21(DE3)/pAS2.144 with 1 mM IPTG induction at 37 °C. Each protein was subjected to two rounds of nickel affinity purification (before and after incubation with His7-tagged TEV protease). ^bThe specific activity of each enzyme using the coupled enzyme assay (Figure 2.8) was calculated using equation 4.1. ^cND: not detected.

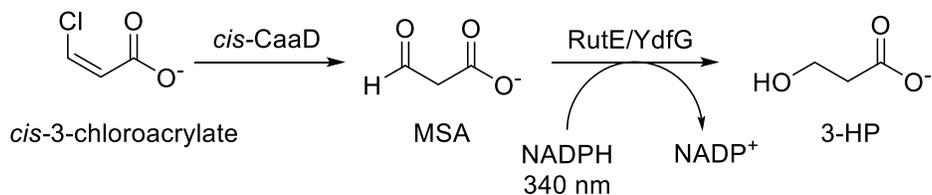


Figure 2.12. Coupled enzyme assay used to determine activity of RutE/YdfG. MSA was generated *in situ* using *cis*-3-chloroacrylate (*cCA*) and *cis*-CaaD. The assays were carried out in 100 mM sodium phosphate pH 8.

2.3.2. Steady-state kinetics and product confirmation of YdfG

Plasmid pAS2.084 was constructed for the expression and purification of YdfG in BL21(DE3), where the gene was cloned into pET-21a(+) resulting in a C-terminal His₆-tag. Plasmid pAS2.084 was constructed prior to pAS1.241 (expressing YdfG with a N-terminal His₆-tag) which was used for activity comparison with RutE. After the decision to continue utilizing YdfG, plasmid pAS2.084 was used for expression since it resulted in a C-terminal His₆-tagged YdfG with a shorter affinity tag that did not require removal by TEV digest. Protein expression was carried out under standard conditions described in chapter four. Protein purification from 1 L of BL21(DE3)/pAS2.084 culture using nickel affinity chromatography resulted in 80–140 mg of YdfG (Figure 2.13). The purified enzyme was stored in 100 mM sodium phosphate, pH 8, 20% glycerol. Initial biochemical characterization to determine the specific activity of YdfG was carried out using the coupled enzyme assay in Figure 2.12. Following the discovery of novel variant, Cg10062(E114N), the generation of MSA *in situ* for Michaelis Menten kinetics was achieved using ACA and Cg10062(E114N) instead of *cis*-CaaD and *cis*-3-chloroacrylate. Although YdfG had been previously characterized by others to determine the relative activity with substrates (serine, threonine, 3-HP),⁶³ this constitutes the first account of its kinetic characterization for MSA reduction to 3-HP. Furthermore, the K_m and k_{cat} values were not previously determined for the YdfG and 3-HP enzyme-substrate pair.⁶⁵ In this study, the specific activity of YdfG for MSA was

UNICORN start 1.1
User: Default 7/12/2022 10:04:10 AM -04:00
Run By : Default 6/4/2019 2:11:13 PM -04:00
Result: 06042019 YdfG 1L 001

1 (1)

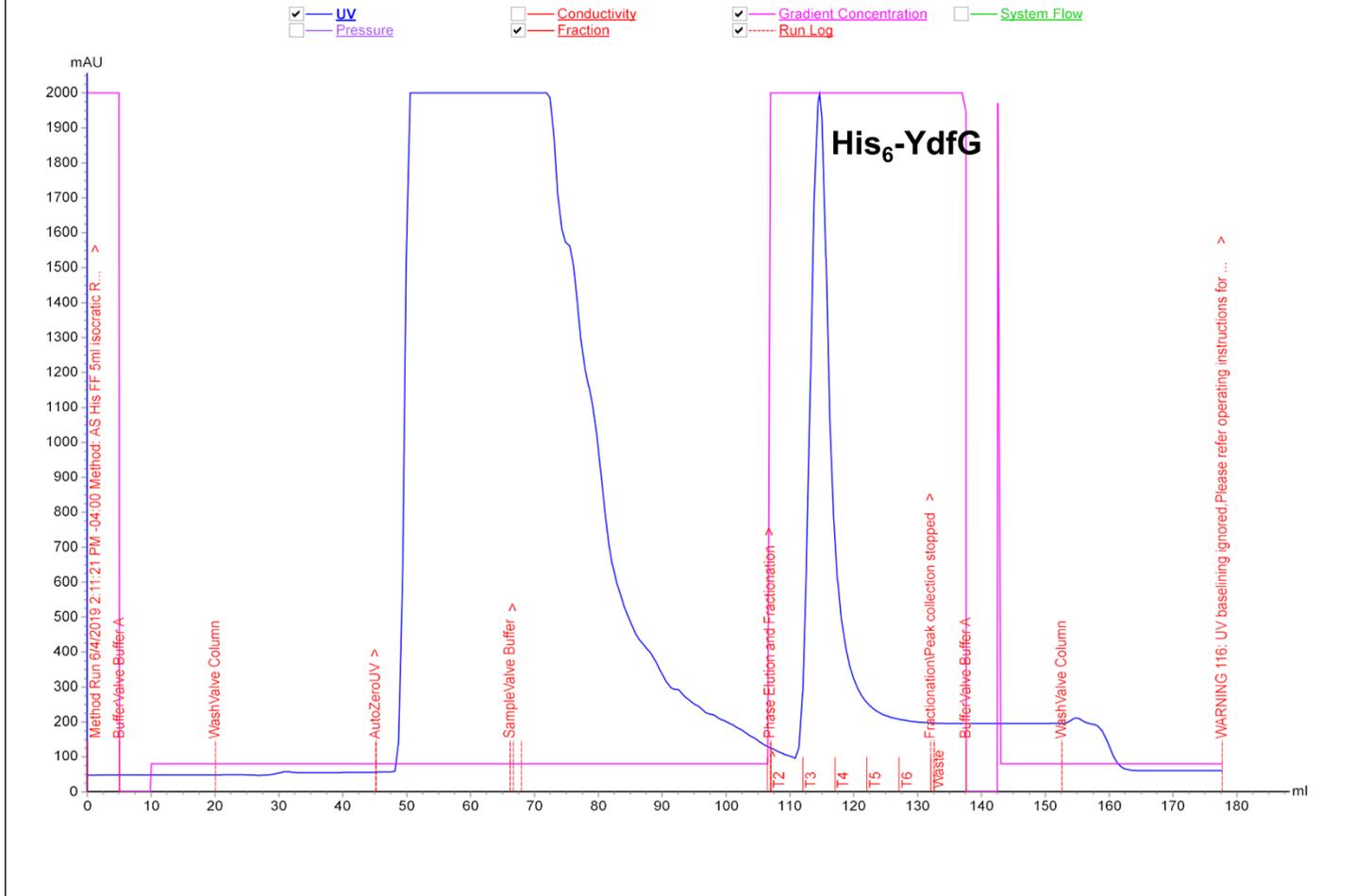


Figure 2.13. Purification of YdfG from BL21(DE3)/pAS2.084 (1 L) via affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

in the range of 20–30 U mg⁻¹. In determining Michaelis Menten kinetic parameters, to ensure YdfG was the rate-limiting enzyme in this assay, Cg10062(E114N) was used in excess. All assay components (except ACA and YdfG) were combined in microcentrifuge tubes and equilibrated for 15 mins at 25 °C. The assays were incubated with ACA for 30 s and initiated by addition of YdfG, assuming complete hydration of ACA to MSA within the 30 s delay. The time taken for complete hydration of ACA was measured prior using the Cg10062 coupled enzyme assay (Figure 4.7). This protocol warranted that the rates measured by the assay were reflective of true YdfG reductase activity at a particular concentration of MSA. The Michaelis Menten kinetic parameters were determined using a MSA concentration range of 10 – 500 μM (Figure 2.14) (ACA stock solutions were used assuming that final [MSA] = initial [ACA] in each assay). The kinetic parameters for YdfG are summarized in Table 2.3. For substrate MSA, a K_m of 137 μM and a k_{cat} of 6.6 s⁻¹ were observed (Table 2.3).

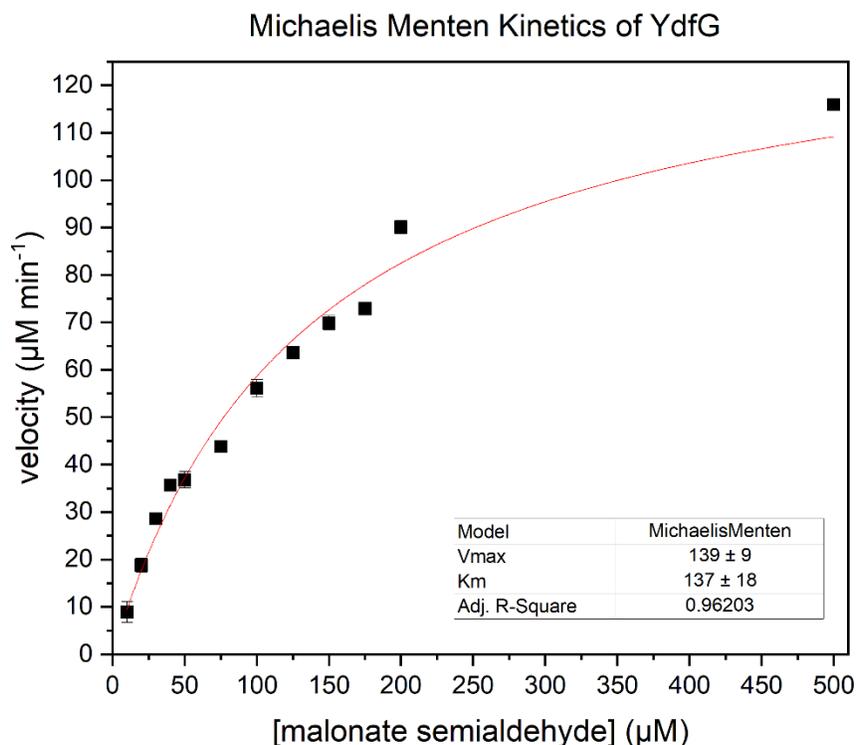


Figure 2.14. Michaelis Menten kinetic parameters of YdfG. Cg10062(E114N) and ACA were used to generate MSA *in situ* to monitor YdfG activity.

In addition to determining the kinetic parameters of YdfG using the UV-based coupled enzyme assay, ^1H NMR was used to confirm the synthesis of 3-HP by YdfG. After 1 h of incubation of ACA (0.05 mM) with Cg10062(E114N) (0.5 U), YdfG (3 U) and NADPH (0.02 mM) in 100 mM sodium phosphate, pH 8, the formation of 3-HP was accounted for by the peaks observed at δ 3.58 (t, 2H) and δ 2.23 (t, 2H) (Figure 2.15). The 3-HP peak at δ 3.58 was obscured by the large glycerol peaks at δ 3.57 (m, 1H). Unreacted ACA was also observed at δ 2.91 (s, 1H) 1 h after the reaction was initiated. Only trace amounts of MSA hydrate were observed at δ 5.13 (t, 1H) and δ 2.30 (d, 2H). Acetaldehyde and its hydrate were also observed at δ 9.47 (q, 1H), δ 2.03 (d, 3H), and δ 5.05 (q, 1H), δ 1.12 (d, 3H), respectively.

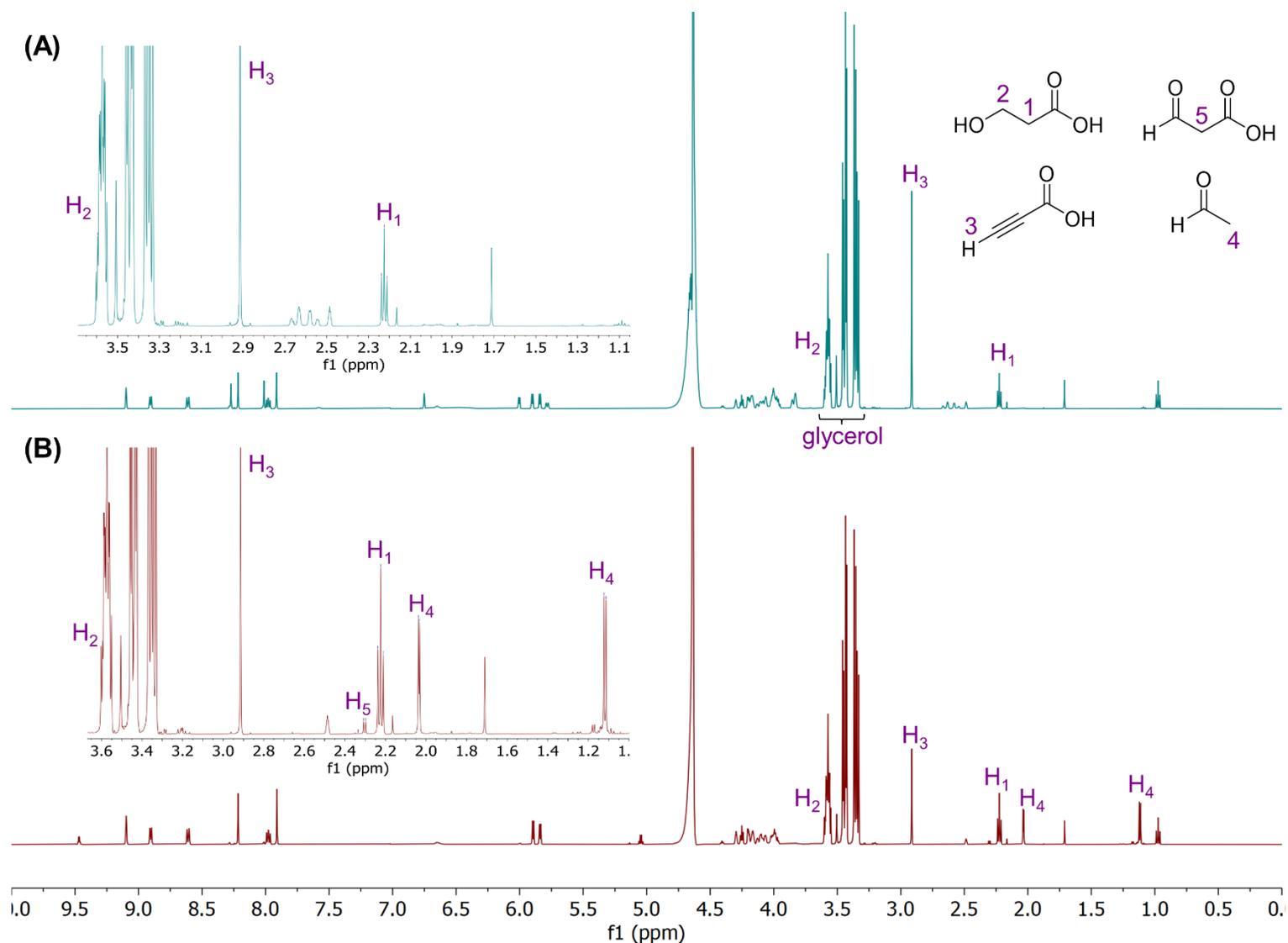


Figure 2.15. ^1H NMR characterization of YdfG. ACA and Cg10062(E114N) were used to generate MSA *in situ* for YdfG reduction. ^1H NMR were obtained (A) immediately and (B) 1 h after addition of ACA. The panels on the left of each spectrum is a zoomed in region of each stacked spectra.

2.4. Cofactor regeneration for 3-HP synthesis

Cg10062(E114N) and YdfG present a unique biocatalytic route via ACA that has not been explored previously for 3-HP production, where ACA may be derived from CH₄ and CO₂ (see chapter one). However, the conversion of 1 mol of ACA to 1 mol of 3-HP requires an equivalent of NADPH. As of 2022, the market prices of NADH and NADPH are listed at \$100 mmol⁻¹ and \$1500 mmol⁻¹ (MilliporeSigma 2022 catalog), respectively. Maintaining a stoichiometric supply of the NADPH equivalent required for industrial production of 3-HP using Cg10062(E114N) and YdfG is not economically feasible.

Continuous regeneration of catalytic amounts of cofactor can be achieved using an additional enzyme that utilizes a sacrificial substrate which is readily available and inexpensive.⁶⁸ In 2001, Costas et al. reported the characterization of a novel phosphorus-oxidizing enzyme, NAD⁺-dependent phosphite dehydrogenase (PTDH) from *Pseudomonas stutzeri* WM88.⁶⁹ PTDH is a homodimer with a 36 kDa monomeric subunit (336 aa) and catalyzes the NAD⁺-dependent oxidation of phosphite (+3 valence) to phosphate (+5 valence) with a K_m of 53 μ M and 54 μ M for phosphite and NAD⁺, respectively and a k_{cat} of 440 min⁻¹.⁶⁹ The enzyme was able to utilize NADP⁺ as a cofactor albeit poorly (data not reported) and the reverse reaction (reduction of phosphate) was not observed.⁶⁹ Zhao and coworkers carried out an extensive study involving several rounds of rational mutagenesis and random mutagenesis to engineer native PTDH for altered cofactor specificity.⁷⁰⁻⁷² The resulting variant, designated 12x PTDH, displayed improved catalysis with NADP⁺ and improved thermostability. The 12x PTDH variant exhibited a K_m of 49 μ M and 75 μ M for NADP⁺ and phosphite, respectively and a k_{cat} of 80 min⁻¹. In contrast, wild type PTDH had a K_m of 2510 μ M and 1880 μ M for NADP⁺ and phosphite, respectively, and k_{cat} of 85 min⁻¹.⁷³ Overall, the improved 12x PTDH variant had a 47-fold increase in k_{cat}/K_m with NADP⁺ relative to wild type PTDH.⁷³ Sequencing of the thermostable 12x PTDH variant revealed a total of 17 amino acid mutations: D13E, M26I, V71I, E130K, Q132R, Q137R, I150F, E175A, Q215L, R275Q, L276Q, I313L, V315A, A319E, A325V, E332N and C336D.^{70,71,73} The 12x PTDH variant was

selected for cofactor recycling in 3-HP synthesis due to its improved K_m and high turnover rate relative to other engineered NADP(H) recycling enzymes such as mutant formate dehydrogenase (FDH) from *Pseudomonas* sp. 101⁷⁴ ($k_{cat} = 2.5 \text{ min}^{-1}$ and $K_m = 150,000 \text{ } \mu\text{M}$) or mutant FDH from *Burkholderia* sp.⁷⁵ ($k_{cat} = 11 \text{ min}^{-1}$ and $K_m = 3200 \text{ } \mu\text{M}$). Furthermore, the low cost and stability of phosphite is an added advantage for its use for 3-HP synthesis.

2.4.1. Steady-state kinetics of PTDH

Plasmid pET15b-12x expressing the NADP⁺-dependent 12x PTDH variant used in our study was available through Addgene (see Table 4.5). The His₆-tagged PTDH was expressed in BL21(DE3) and purified via nickel affinity chromatography (Figure 2.17). Purified enzyme was initially stored in 50 mM MOPS, pH 7.25, as described in literature.^{70,71,73} However, no change in activity was observed when the enzyme was stored in 100 mM sodium phosphate pH 8 (storage and activity buffer for Cg10062(E114N) and YdfG) and consequently, purified PTDH was stored in 100 mM sodium phosphate, pH 8, containing 20% ethylene glycol as a cryoprotectant. Biochemical characterization of PTDH was carried out using an assay measuring the increase in absorbance at 340 nm due to reduction of NADP⁺ (Figure 2.12). Additionally, PTDH activity was also tested with NAD⁺, its native cofactor.

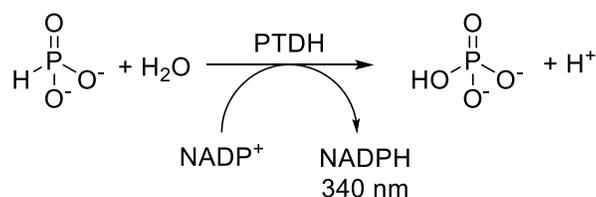


Figure 2.16. Enzyme assay used to determine measure PTDH activity. All assays were carried out in 100 mM sodium phosphate, pH 8.

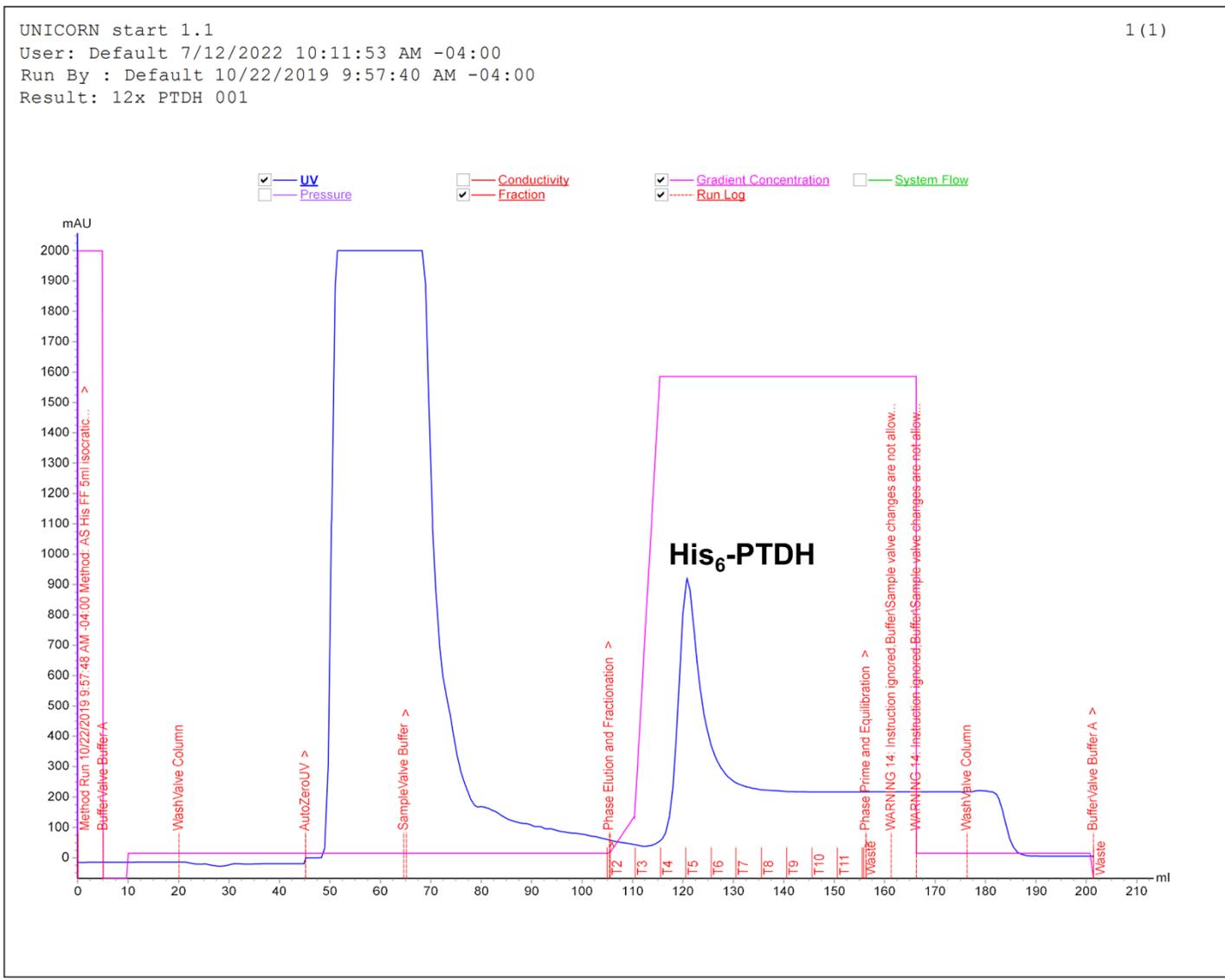


Figure 2.17. Purification of PTDH from BL21(DE3)/pET15b-12x (1 L) via affinity chromatography. UV trace (blue), imidazole gradient concentration (pink).

To determine the specific activity of PTDH, the assays were initiated by the addition of saturating concentrations of sodium phosphite (Table 2.2). Although PTDH displayed considerable activity with NADP⁺, a significantly higher specific activity was observed with NAD⁺. The Michaelis Menten kinetic parameters were determined using a sodium phosphite concentration range of 50 – 2000 μM when NADP⁺ was supplied as cofactor (Figure 2.18). With the engineered PTDH variant, a K_m of 824 μM and an overall catalytic efficiency of $0.15 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ was observed for phosphite (Table 2.3). The k_{cat} of 1.28 s^{-1} obtained in this study was consistent with previously reported values. However, despite several attempts, detectable activity was not observed at lower concentrations of phosphite (< 50 μM) and the experimental K_m obtained in this study was nearly 11-fold higher than previously reported (75 μM).⁷³

Table 2.2. Specific activity of PTDH.

Cofactor	Phosphite (mM)	Specific Activity (U mg ⁻¹)
NADP ⁺	25	1.0
NAD ⁺	25	4.7

Table 2.3. Michaelis Menten kinetic parameters of enzymes used for 3-HP synthesis.

Enzyme	K_m (μM)	k_{cat} (s ⁻¹)	$k_{cat} / K_m \times 10^4$ (M ⁻¹ s ⁻¹)
Cg10062(E114N)	45 ± 1	1.84 ± 0.03	4.08 ± 0.11
YdfG	137 ± 18	6.62 ± 0.43	4.83 ± 0.71
PTDH	824 ± 114	1.28 ± 0.12	0.15 ± 0.03

The kinetic parameters were determined using the respective enzyme assays described for each enzyme in 100 mM sodium phosphate pH 8 at 25 °C.

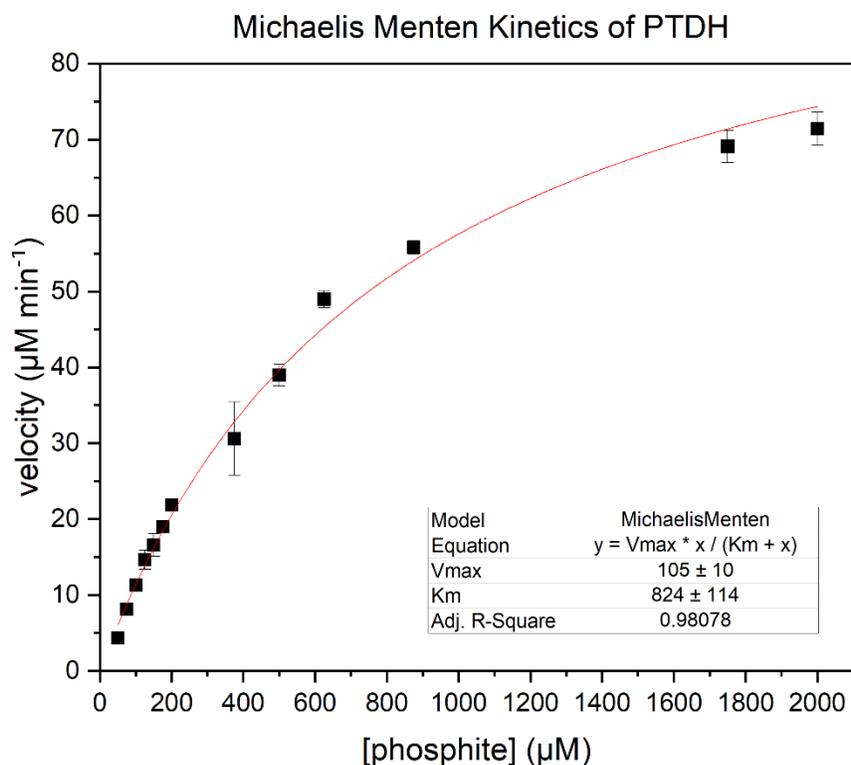


Figure 2.18. Michaelis Menten kinetics of PTDH. The PTDH enzyme assays were carried out in 100 mM sodium phosphate pH 8.

2.5. Buffer dependence of Cg10062(E114N), YdfG and PTDH

In order to optimize conditions for 3-HP synthesis, the activity of each enzyme was determined over a broad pH range using different buffers. Four different buffers were used: 100 mM citrate-phosphate, 100 mM sodium phosphate, 50 mM tris-HCl and 100 mM sodium carbonate/bicarbonate buffers for pH ranges of 3.6–5.6, 6.0–8.0, 7.6–9.2 and 9.2–9.6, respectively. The enzyme activities were determined using the respective enzyme assays described previously for each enzyme. All assays were carried out in triplicate (1 mL) on a Shimadzu UV2600 spectrophotometer at 25 °C to ensure that the final pH of each assay remained unchanged with the addition of each assay component. All stock solutions were prepared in 100 mM sodium phosphate, pH 8, and the assays were carried out in the respective buffers for each pH, where each buffer constituted > 90% of the total assay volume.

For Cg10062(E114N), at each pH, Cg10062(E114N) (0.05 U), MSAD (1.2 U), ADH (12 U) and NADH (1.2 mM) were combined with 920 μ L of buffer. The assays were initiated with the addition of ACA (1 mM). The initial velocities of Cg10062(E114N) across different buffers are shown in Figure 2.19. The highest specific activity of 5 U mg^{-1} was observed in 100 mM sodium phosphate pH 8. All other specific activities at different pH values were reported relative to this specific activity. The activity of Cg10062(E114N) in Tris-HCl followed a similar trend to that observed in sodium phosphate buffer at pH 7.6 and 8.0. However, the overall specific activity of the enzyme was approximately 20% lower at these two pH values in Tris-HCl. Cg10062(E114N) retained 40% enzyme activity at a high pH of 9.2 in tris-HCl buffer but not in carbonate-bicarbonate. Collectively, optimum activity of Cg10062(E114N) was observed at pH 8 in both sodium phosphate and tris-Cl buffers.

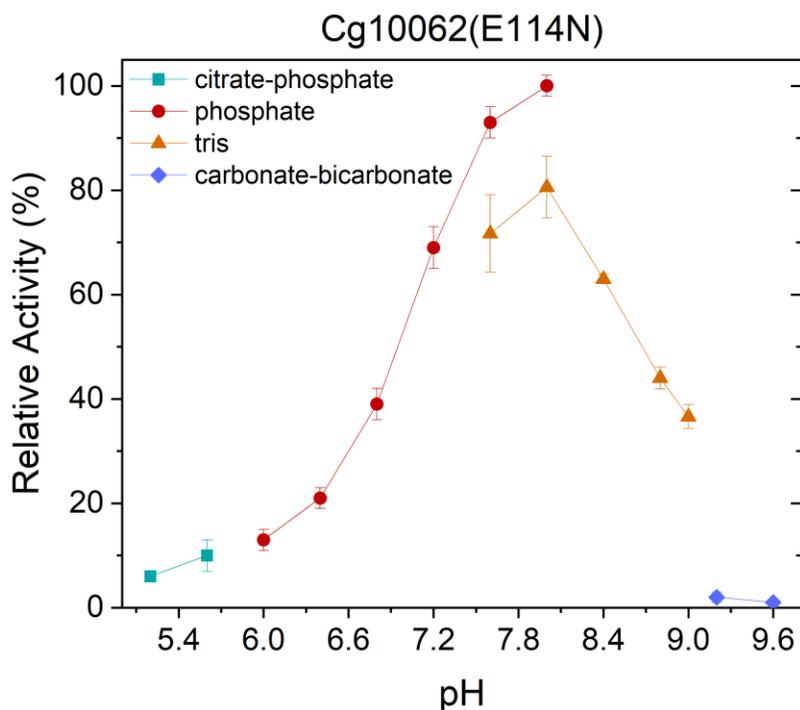


Figure 2.19. Relative activity of Cg10062(E114N) in various buffers.

For YdfG assays, each assay contained Cg10062(E114N) (1.5 U), NADPH (1.2 mM), YdfG (1.5 U) and ACA (1 mM) with 960 μ L of each buffer. It must be noted that the coupled enzyme assay used to measure YdfG activity relies on Cg10062(E114N) to produce the unstable MSA substrate *in situ*. Therefore, the buffer and pH dependence of Cg10062(E114N) was expected to have a direct influence on the activity of YdfG. Thus, in order to obtain the true activity of YdfG, the ACA was first added to Cg10062(E114N) in 100 mM sodium phosphate buffer pH 8 and incubated for 30 s to generate MSA and this mixture was then added to the remaining components in each assay. Using 1.5 U of Cg10062(E114N) ensured rapid hydration of ACA to MSA within 30 s (determined experimentally). Unlike Cg10062(E114N), the highest specific activity for YdfG was recorded in 100 mM sodium phosphate pH 7.6 (36 U mg^{-1}) (Figure 2.20). Using the same buffer at pH 8.0 resulted in a 16% decrease in specific activity. As with Cg10062(E114N), the overall activity of YdfG in tris-HCl buffer was reduced relative to the observed activity in sodium phosphate buffer. Furthermore, at lower and higher pH ranges, using citrate-phosphate and carbonate-bicarbonate buffers, respectively, YdfG displayed minimal activity.

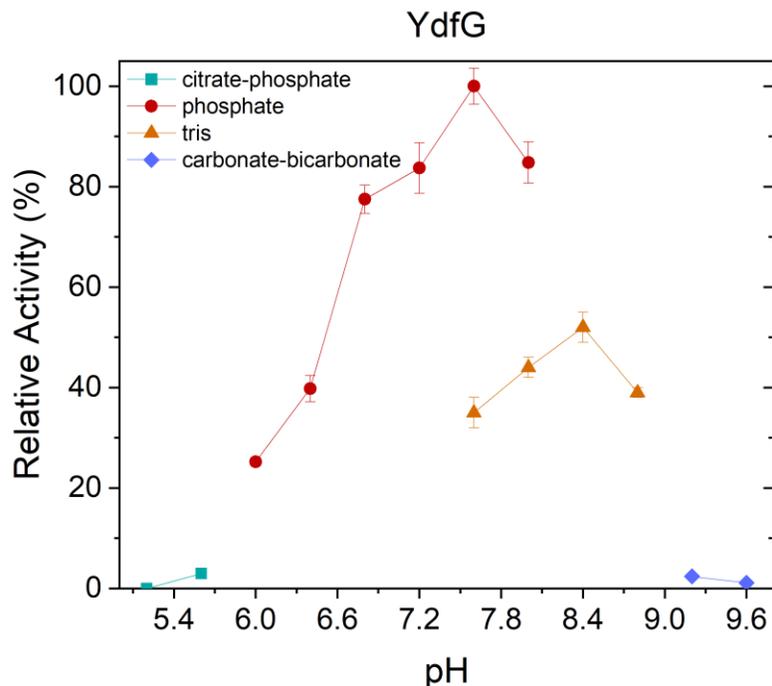


Figure 2.20. Relative activity of YdfG in various buffers.

The buffer dependence of PTDH was determined using the same methods as with Cg10062(E114N) and YdfG. For each assay, PTDH (0.05 U) and NADP⁺ (1.2 mM) were combined with 970 μ L of each buffer. The assays were initiated by the addition of sodium phosphite (10 mM). In line with previously reported literature,⁶⁹ PTDH maintained enzyme activity over a significantly wider pH range relative to the other two enzymes tested (Figure 2.21). An optimal specific activity of 1.2 U mg⁻¹ was observed at pH values of 6.0 and 6.4. In contrast with Cg10062(E114N) and YdfG, PTDH retained relatively high activity at a lower pH range with citrate-phosphate buffer while a reduction in overall activity was observed with Tris-HCl buffer. Overall, PTDH activity was less sensitive to change in pH and salts relative to Cg10062(E114N) and YdfG.

Considering the specific activities of all three enzymes at a range of pH values, 100 mM sodium phosphate pH 8 buffer remained suitable for the synthesis of 3-HP from ACA. Optimal hydratase activity by Cg10062(E114N) for the conversion of ACA to MSA is highly desirable. Under the same conditions, YdfG and PTDH were able to retain approximately 80% and 90%

catalytic activity, respectively. The loss of activity under these conditions can be compensated for by the addition of excess enzyme to allow efficient and quantitative conversion of ACA to 3-HP

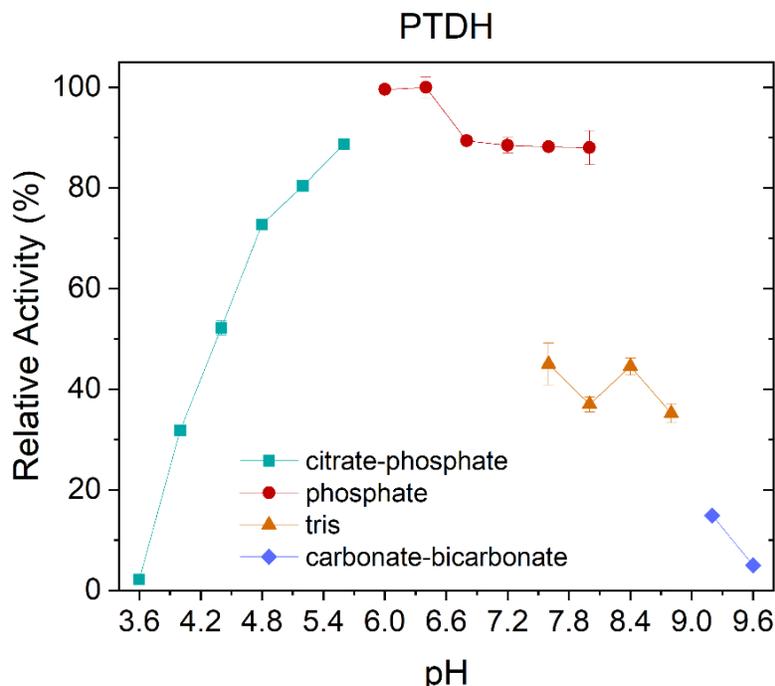


Figure 2.21. Relative activity of PTDH in various buffers.

2.6. *In vitro* synthesis of 3-HP from ACA with cofactor recycling

The *in vitro* syntheses of 3-HP from ACA were carried out in 100 mM sodium phosphate pH 8 (1 mL) using the route shown in Figure 2.22. Cg10062(E114N), YdfG and PTDH were expressed using BL21(DE3)/pAS2.100, BL21(DE3)/pAS2.084 and BL21(DE3)/pET15b-12x using 1 mM IPTG. All three enzymes used for 3-HP synthesis were freshly purified via nickel affinity chromatography using standard conditions described previously (Figure 2.23) and the kinetic parameters of all three enzymes are summarized in Table 2.3. The expression of all three enzymes were observed in cultures that were not induced with IPTG, indicative of leaky expression commonly associated with the high strength T7 promoter system.⁷⁶ Induction with 1 mM IPTG led to significantly higher expression of all three enzymes as indicated by SDS-PAGE.

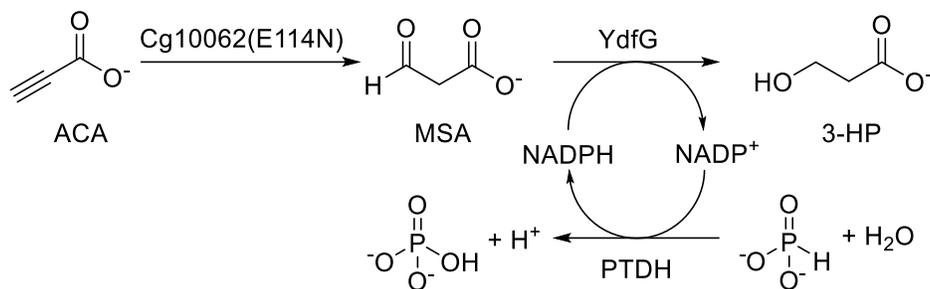


Figure 2.22. Biocatalytic route to 3-HP via ACA

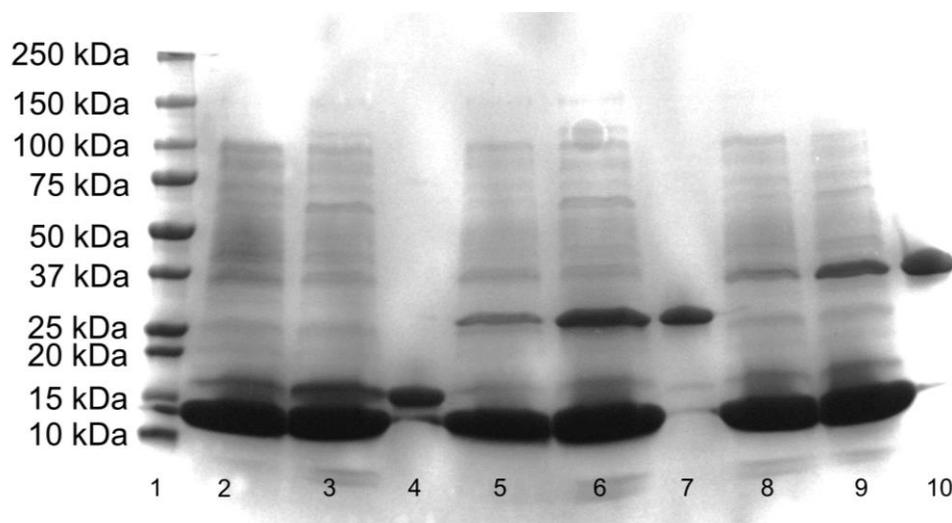


Figure 2.23. SDS-PAGE of enzymes used for 3-HP synthesis. Lanes 1: 10 kDa ladder; 2: lysate of uninduced BL21(DE3)/pAS2.100; 3: lysate of IPTG-induced BL21(DE3)/pAS2.100; 4: purified Cg10062(E114N) (19 kDa); 5: lysate of uninduced BL21(DE3)/pAS2.084; 6: lysate of IPTG-induced BL21(DE3)/pAS2.084; 7: purified YdfG (28 kDa); 8: lysate of uninduced BL21(DE3)/pET15b-12x; 9: lysate of IPTG-induced BL21(DE3)/pET15b-12x and 10: purified PTDH (36 kDa). The large band present in the lanes containing lysate correspond to lysozyme (14 kDa) used for cell lysis.

HPLC was used to monitor the consumption of ACA and formation of 3-HP (Figure 2.24) in each reaction using Aminex HPX-87H column (0.6 mL min⁻¹ with 0.01 N H₂SO₄). The concentrations of each compound were determined using calibration curves constructed using ACA (Figure 2.19) and 3-HP (Figure 2.20) standard solutions. During early attempts to transform ACA to 3-HP using the system described above, 20% glycerol was used as a cryoprotectant for enzyme storage. However, glycerol co-eluted with 3-HP (RT=13.5 min) under the elution

conditions utilized, rendering the quantification of any 3-HP formed difficult. Instead, ethylene glycol (RT = 16 min) was selected as a suitable replacement using the same elution conditions (Figure 2.21).

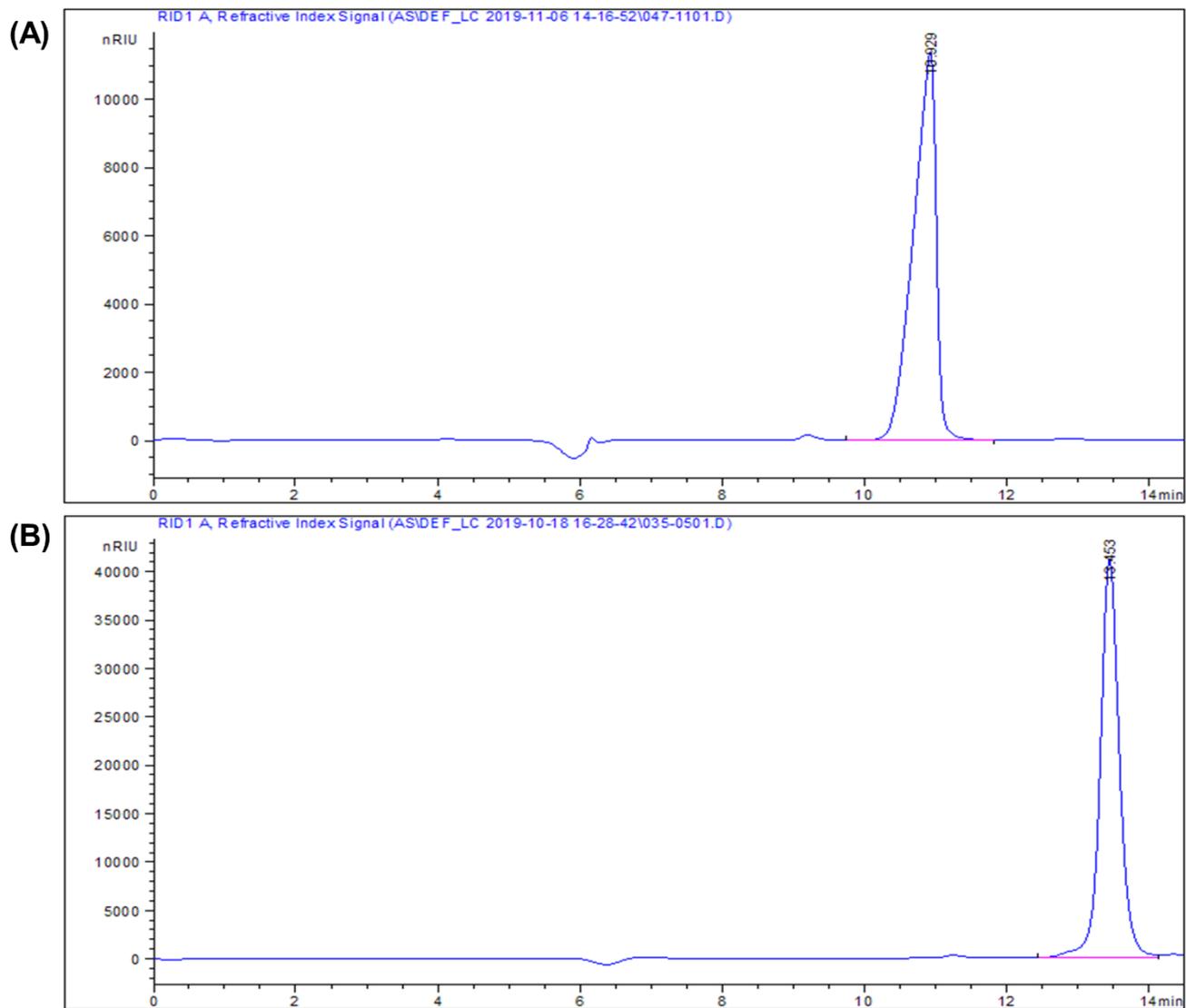


Figure 2.24. HPLC chromatogram of (A) ACA and (B) 3-HP. Elution conditions: Aminex HPX-87H column with 0.01 N H₂SO₄ (0.6 mL min⁻¹).

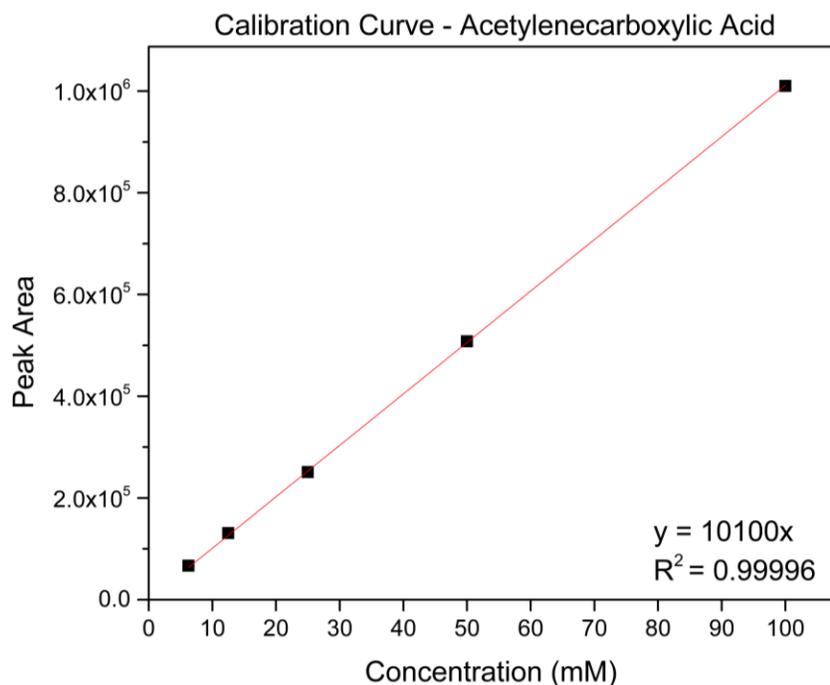


Figure 2.25. HPLC calibration curve of ACA. All standard solutions were prepared in triplicate and analyzed on an Aminex HPX-87H column with 0.01 N H₂SO₄ (0.6 mL min⁻¹).

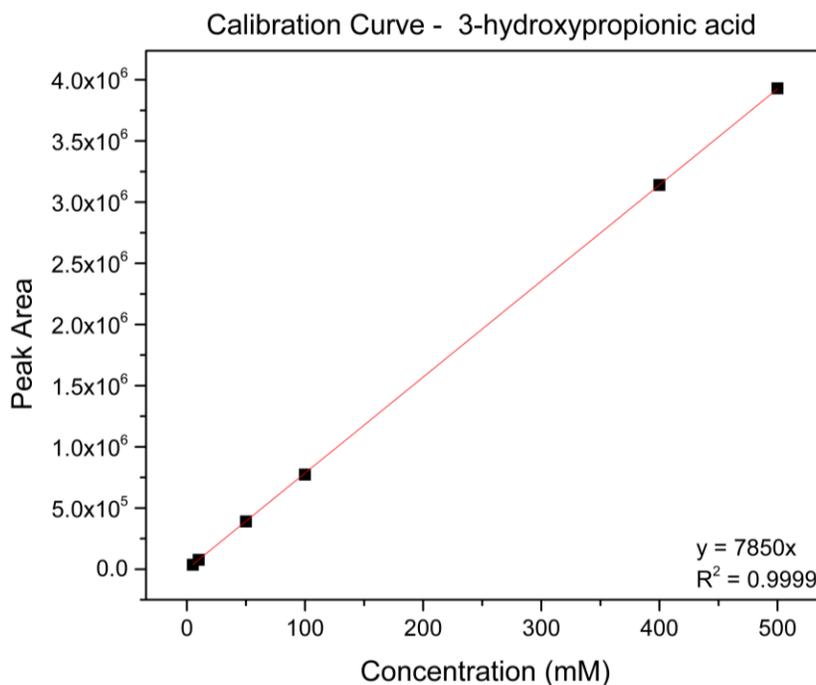


Figure 2.26. HPLC calibration curve of 3-HP. All standard solutions were prepared in triplicate and analyzed on an Aminex HPX-87H column with 0.01 N H₂SO₄ (0.6 mL min⁻¹).

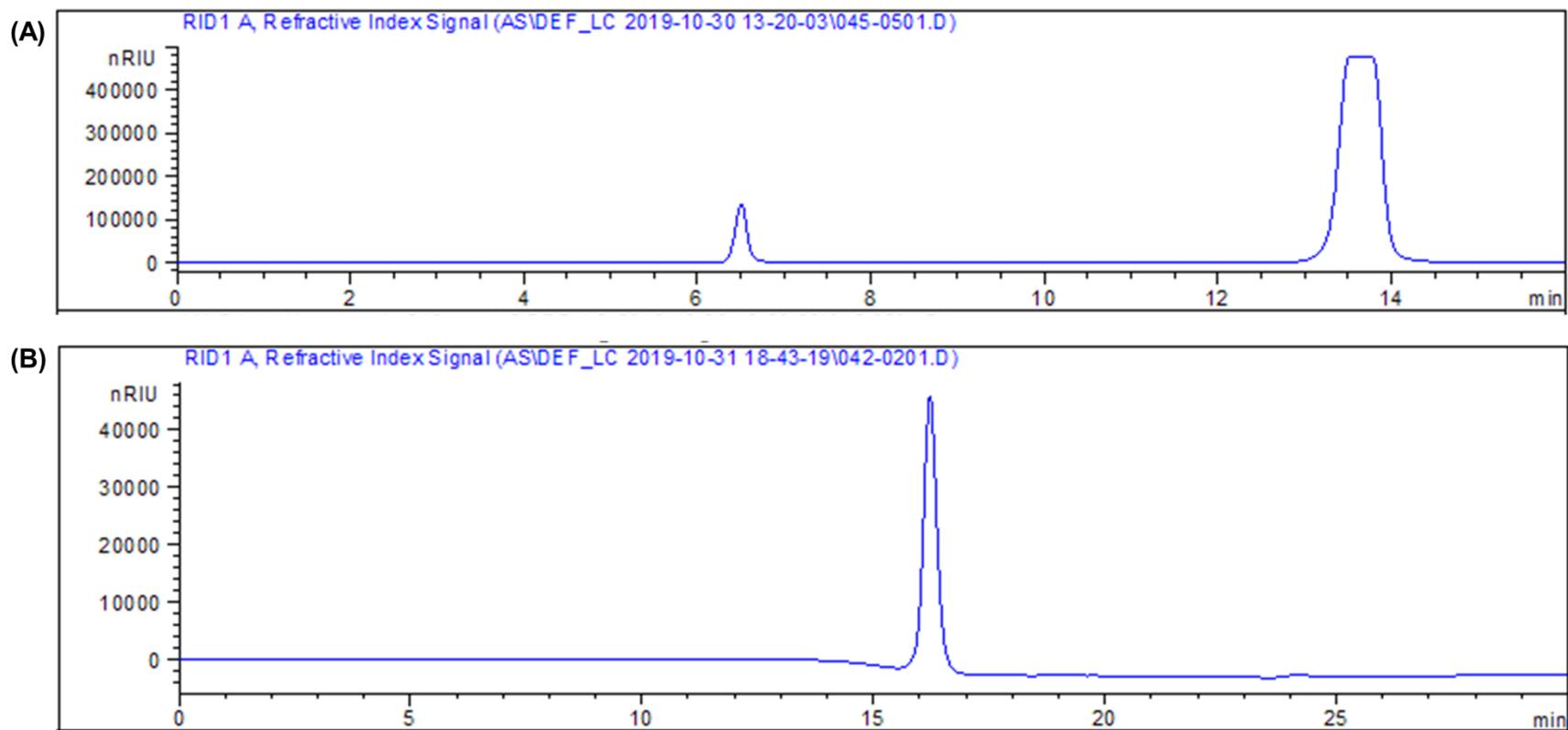


Figure 2.27. HPLC chromatogram of (A) glycerol (RT = 13.5 min) and (B) ethylene glycol (RT = 16 min). Elution conditions: Aminex HPX-87H column with 0.01 N H₂SO₄ (0.6 mL min⁻¹).

To reduce the cost associated with the 3-HP synthesis from ACA, NADP⁺ was used in place of NADPH, since the reduced form of the nicotinamide cofactor is more expensive (\$350 mmol⁻¹ NADP⁺ vs. \$1500 mmol⁻¹ NADPH: MilliporeSigma 2022). The first attempt to synthesize 3-HP via ACA was carried out using four different molar equivalents of NADP⁺ relative to 100 mM ACA starting material: 100 mM (1 eq.), 10 mM (0.1 eq.), 5 mM (0.05 eq.) and 1 mM (0.01 eq.). Each reaction was carried contained Cg10062(E114N) (0.2 U), PTDH (0.3 U), YdfG (3 U), sodium phosphite (400 mM) and was initiated with the addition of ACA (100 mM). Samples were analyzed by HPLC every hour for a period of 5 h (Figure 2.28). In each reaction, an increase in 3-HP proportional to the disappearance of ACA was observed (Figure 2.29). The rates of ACA hydration and 3-HP formation appeared to decrease with increasing concentrations of cofactor, suggesting that enzyme activity may be affected by NADP(H). Additionally, none of the reactions reached completion, as indicated by unreacted ACA. Given 0.2 U of Cg10062(E114N) with a specific activity of 5 U mg⁻¹ and starting with 100 mM ACA (100 μmol), complete hydration of ACA was expected in approximately 8 h, assuming that the Cg10062(E114N) maintains 100% hydratase activity over the course of the reaction. To estimate the time required for complete hydration of ACA by Cg10062(E114N), a control experiment with PTDH and its substrate was carried out with ACA (100 mM), NADP⁺ (100 mM), Cg10062(E114N) (0.2 U) and YdfG (0.3 U). The reaction was monitored over 24 h using HPLC (Figure 2.30). Although the formation of 3-HP increased with the corresponding disappearance of ACA, using a stoichiometric amount of NADP⁺ did not result in quantitative conversion of ACA to 3-HP. As observed in the first experiment (Figure 2.21), it was hypothesized that the high concentrations of NADPH inhibited either Cg10062(E114N) or YdfG, or both. The slow rate of ACA hydration and the slow formation of 3-HP (which is dependent on the rate of ACA hydration) suggested that Cg10062(E114N) was most likely inhibited by NADPH. As a result, in order to improve ACA hydration, the units of Cg10062(E114N) would have to be increased.

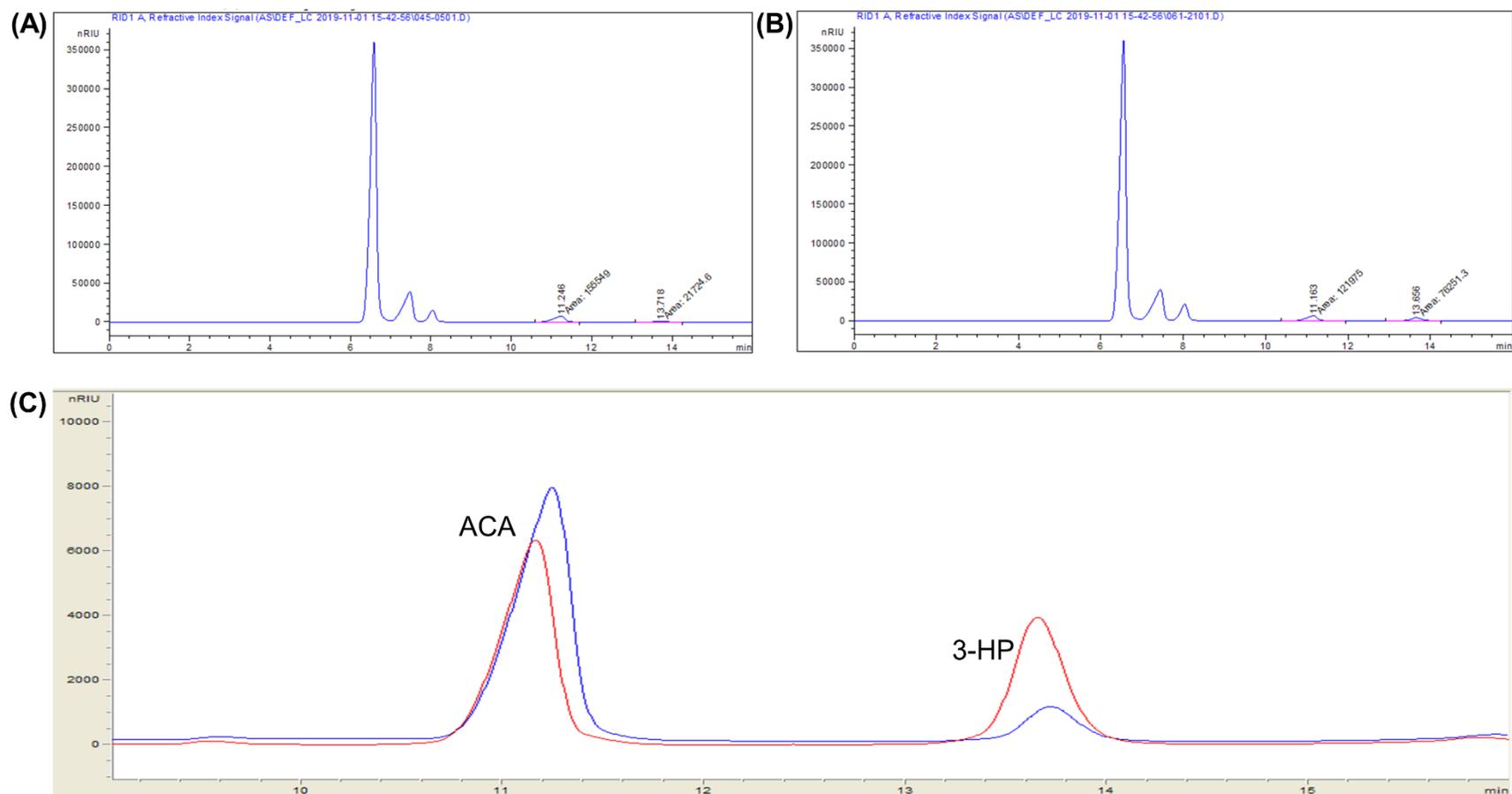


Figure 2.28. HPLC chromatogram of 3-HP synthesis from 100 mM ACA with 1 mM NADP(H) at (A) 1 h and (B) 5 h after initiation. (C) overlay of (A) (blue) and (B) (red) shows disappearance of ACA and formation of 3-HP.

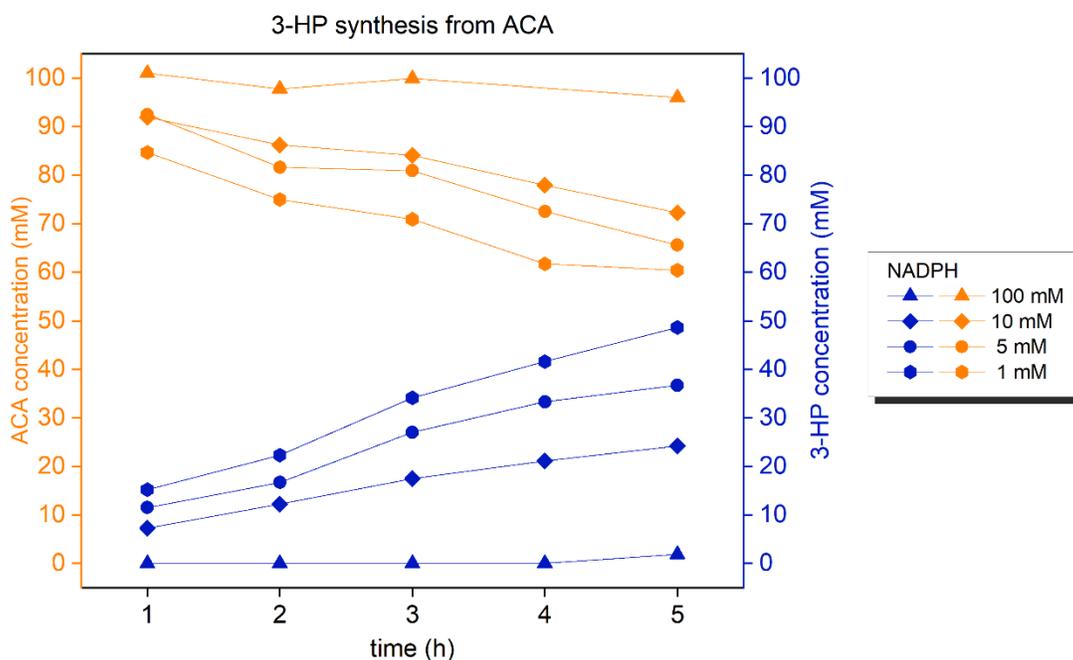


Figure 2.29. *In vitro* synthesis of 3-HP from ACA. Conversion of ACA to 3-HP was carried out using at 25 °C using four different molar equivalents of NADP(H) relative to ACA: 100 mM (triangle), 10 mM (diamond), 5 mM (circle) and 1 mM (hexagon). ACA (orange) and 3-HP (blue) was quantified by HPLC.

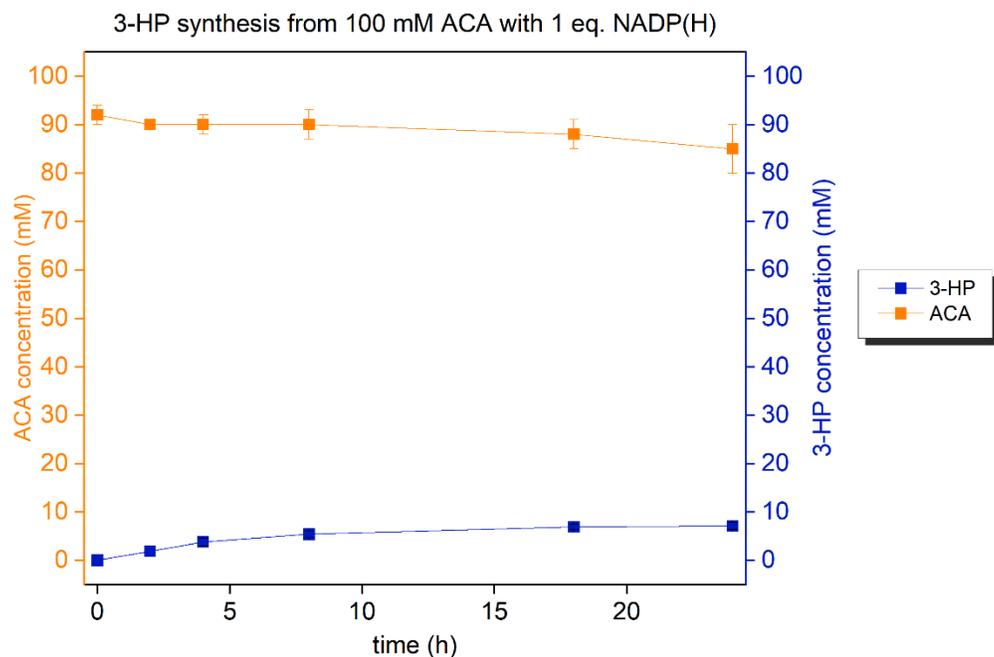


Figure 2.30. *In vitro* synthesis of 3-HP from ACA without cofactor regeneration. Conversion of ACA to 3-HP was carried out using at 25 °C using a stoichiometric amount of NADP(H) (100 mM) relative to ACA. ACA (orange) and 3-HP (blue) was quantified by HPLC.

In subsequent experiments, the transformation of 100 mM ACA to 3-HP were carried out with three concentrations of NADP⁺: 10 mM, 1 mM and 0.1 mM (Table 2.4). In order to improve ACA hydration, 2 U of Cg10062(E114N) were used, with YdfG (15 U) and PTDH (15 U). Each reaction contained an excess of sodium phosphite (150 mM) relative to ACA. An excess of YdfG and PTDH was used to minimize the accumulation of MSA formed by Cg10062(E114N)-catalyzed hydration of ACA. Excess PTDH and phosphite was used to allow uninterrupted recycling of NADP⁺ to NADPH to allow continuous YdfG-catalyzed reduction of MSA to 3-HP. HPLC was used to monitor reaction progress over 30 h (Figure 2.31) and ¹H NMR was used as a second line of evidence to confirm the identity of components present in each reaction (not used for quantification).

Table 2.4. Synthesis of 3-HP from 100 mM ACA with cofactor regeneration.

Sample	ACA (mM)	NADP ⁺ (mM)	Na ₂ HPO ₃ (mM)	Cg10062(E114N) (U)	YdfG (U)	PTDH (U)
0.1 eq.	100	10	150	2	15	15
0.01 eq.	100	1	150	2	15	15
0.001 eq.	100	0.1	150	2	15	15

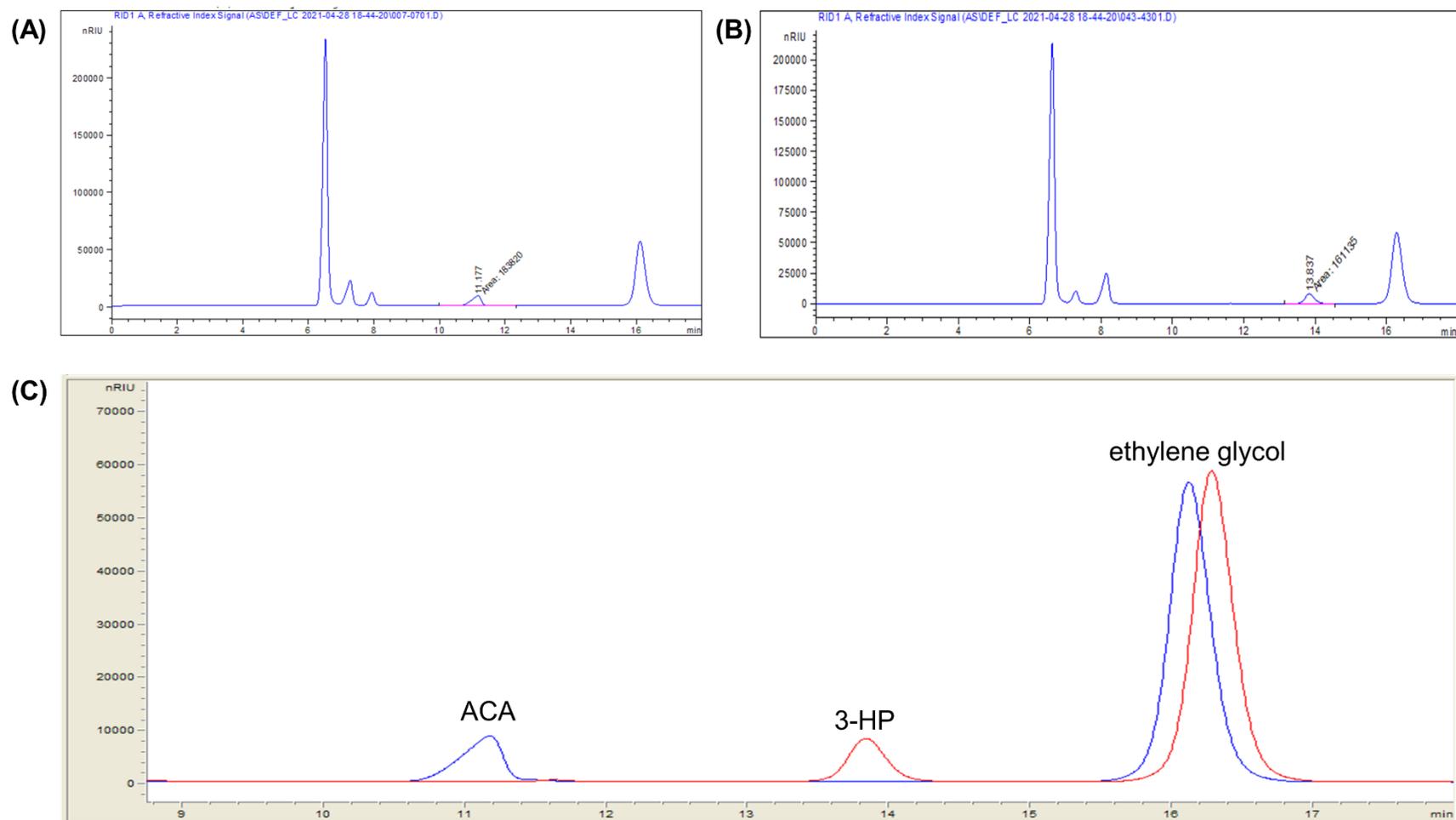


Figure 2.31. HPLC chromatogram of 3-HP synthesis from 100 mM ACA with 0.1 mM NADP(H) (A) immediately after initiation and (B) 30 h after initiation. (C) overlay of (A) (blue) and (B) (red) shows complete consumption of ACA and formation of 3-HP.

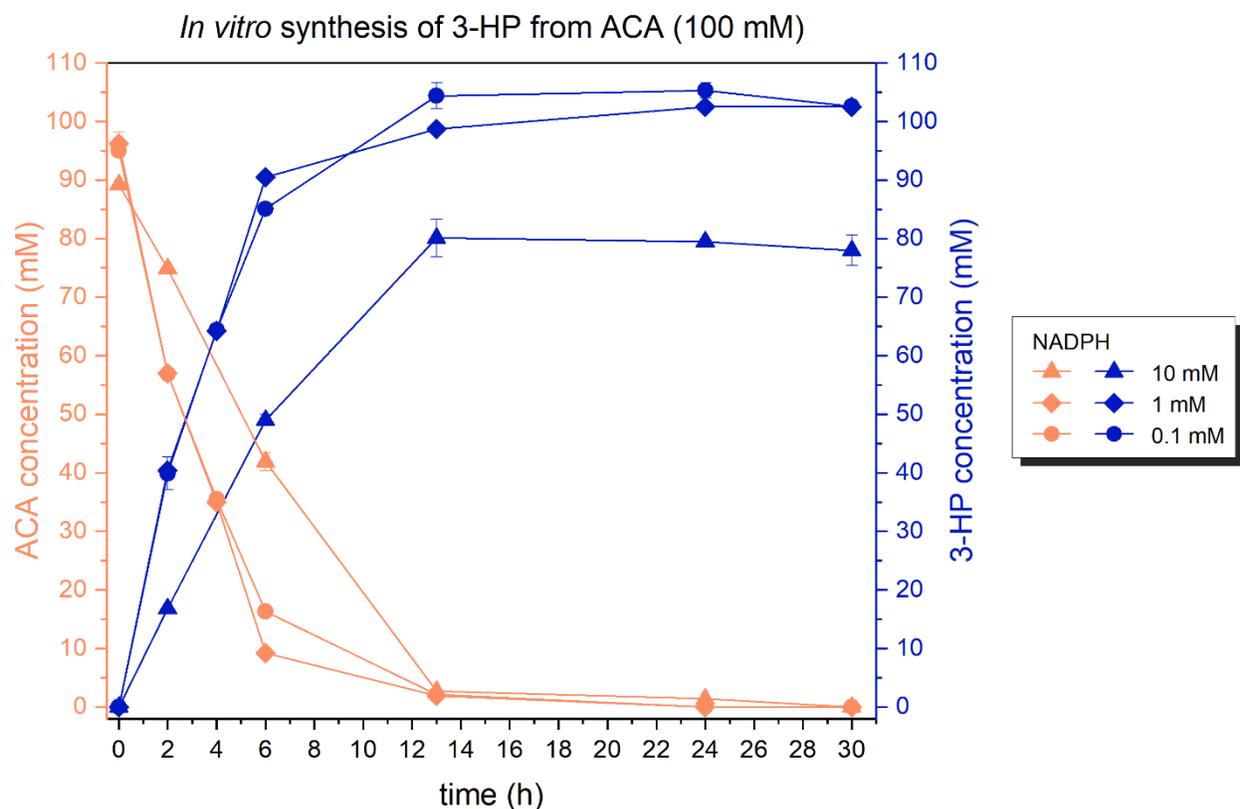


Figure 2.32. *In vitro* synthesis of 3-HP from 100 mM ACA with cofactor regeneration. Conversion of ACA to 3-HP was carried out using three different molar equivalents of NADP(H) relative to ACA: 10 mM (triangles), 1 mM (diamonds), and 0.1 mM (circles). ACA (orange) and 3-HP (blue) were quantified by HPLC.

Starting with an initial concentration of 100 mM ACA, quantitative conversion of ACA to 3-HP was observed with as low as 0.001 eq (0.1 mM) NADP(H). Given the use of 2 U of Cg10062(E114N), complete hydration of 100 mM ACA (100 μ mol) would occur within 1 h. However, these calculations assume that Cg10062(E114N) maintains 100% hydratase activity and is not inhibited over the course of the reaction. In all three reactions, the ACA was completely consumed by 13 h (Figure 2.32) indicating efficient hydration of ACA by Cg10062(E114N). An increase in 3-HP proportional to the rate of ACA hydration was also observed in all reactions. However, the reaction containing the highest concentration of NADP(H) (10 mM) only resulted 80% conversion to 3-HP, consistent with previous observations where the rate of conversion decreased at higher equivalents of cofactor.

It must be noted that ^1H NMR methods used for 3-HP synthesis experiments differ from methods used for characterization of Cg10062 and variants in chapter one since 5 M H_2SO_4 was used to quench an aliquot from each reaction at particular timepoints. As a result, the final pH of each sample was lower than 2, resulting in a downfield shift of the compound signals (see Section 4.15.10). Under these conditions, 3-HP was indicated by the presence of resonances at δ 3.66 (t, 2H) and δ 2.42 (t, 2H). The resonance at δ 3.27 (s, 1H) corresponds to ACA, which was only observed at initiation of the reactions (Figures 2.33–2.35).

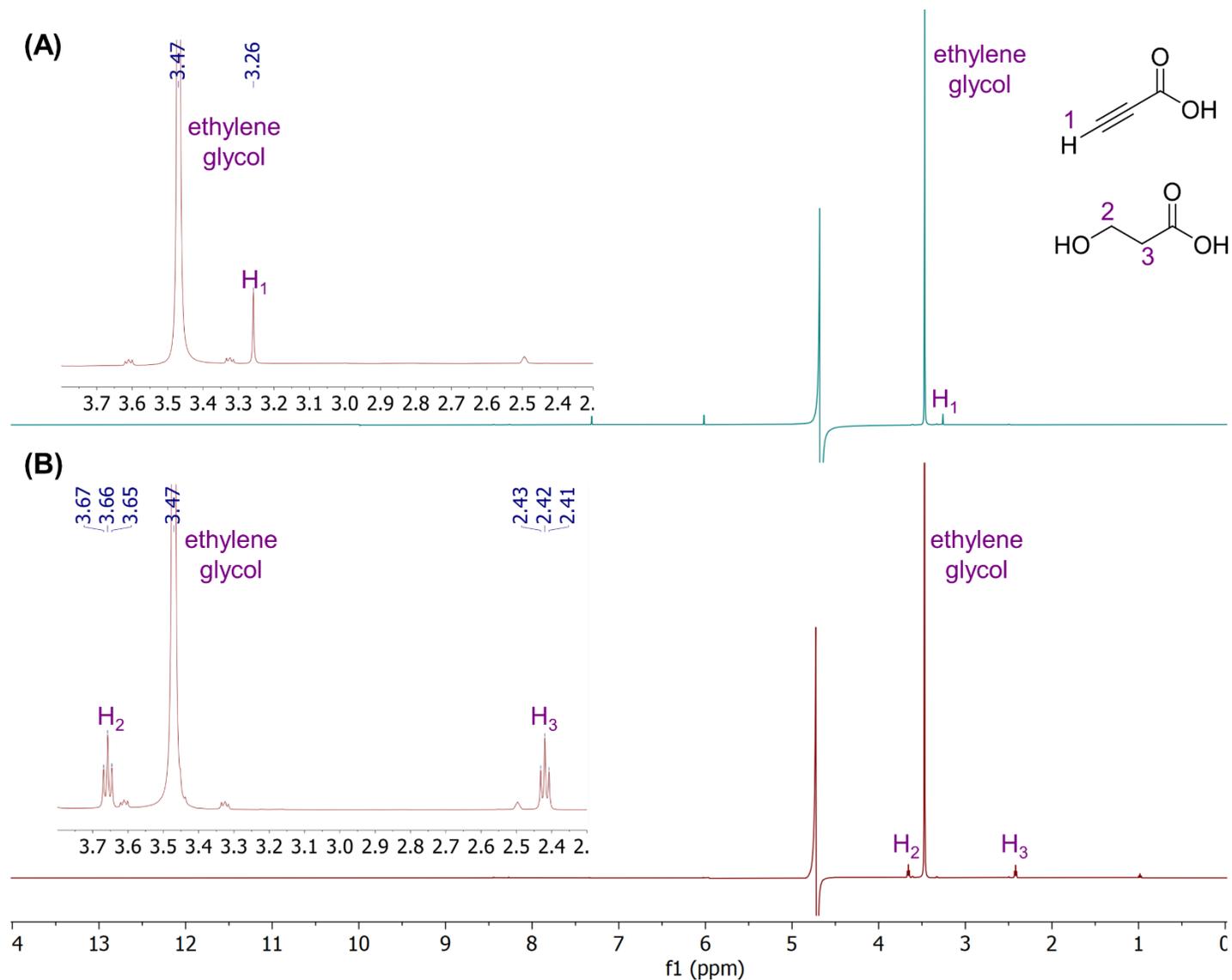


Figure 2.33. ^1H NMR of 3-HP synthesis with 100 mM ACA and 10 mM NADP(H). ^1H NMR spectra of the reaction at (A) 0 h and (B) 30 h. The panels on the left are zoomed in regions of each spectra indicating the ACA and 3-HP signals.

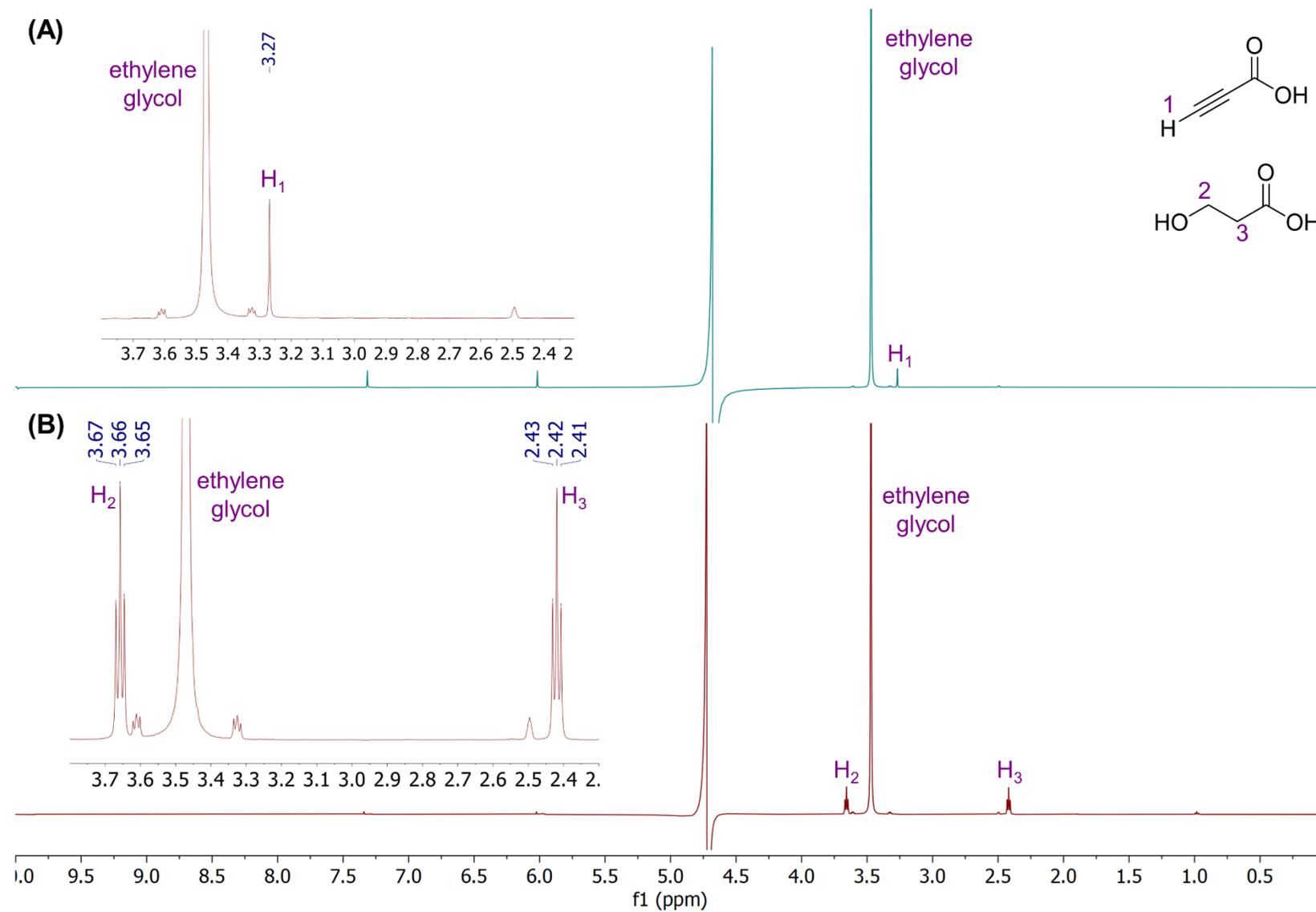


Figure 2.34. ^1H NMR of 3-HP synthesis with 100 mM ACA and 1 mM NADP(H). ^1H NMR spectra of the reaction at (A) 0 h and (B) 30 h. The panels on the left are zoomed in regions of each spectra indicating the ACA and 3-HP signals.

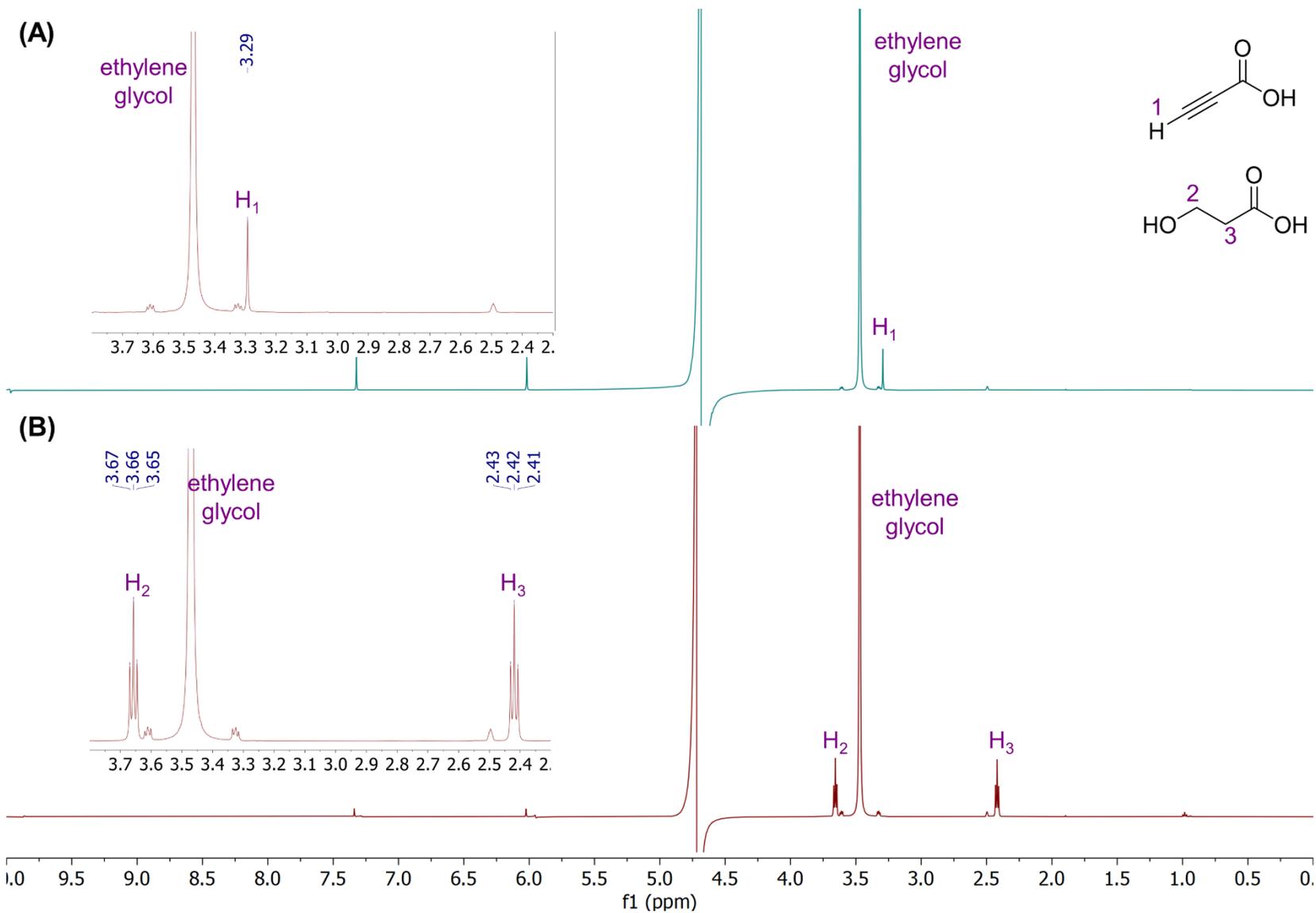


Figure 2.35. ^1H NMR of 3-HP synthesis with 100 mM ACA and 0.1 mM NADP(H). ^1H NMR spectra of the reaction at (A) 0 h and (B) 30 h. The panels on the left are zoomed in regions of each spectra indicating the ACA and 3-HP signals.

Since 3-HP synthesis from 100 mM ACA with cofactor regeneration was carried out successfully utilizing as low as 0.001 eq NADP(H) (0.1 mM), the same experiments were carried out using an initial concentration of 500 mM ACA (Figure 2.36). The scale-up reaction was carried out to test the feasibility of this biocatalytic route at higher substrate concentrations. It must be noted that the amount of enzyme used in the scale-up to 500 mM ACA was identical to that used in the set of reactions with 100 mM ACA (2 U: 15 U: 15 U of Cg10062(E114N): YdfG: PTDH). Therefore, incomplete consumption of ACA after 61 h observed via HPLC in all three reactions, was expected. The second set of ^1H NMR analyses for these reactions were carried out 72 h after reaction initiation (Figures 2.37–2.39). Utilizing 2 U of Cg10062(E114N) with specific activity of 5 U mg^{-1} , complete hydration of 500 mM ACA (500 μmol) would occur within approximately 4 h assuming that Cg10062(E114N) maintains 100% hydratase activity over the course of the reaction. However, a thorough examination of the results indicate that the rate of ACA hydration was inhibited at increasing cofactor concentrations, consistent with previous observations. As with 100 mM ACA, the reaction containing 0.1 eq (50 mM) NADP(H) was the least efficient. While unreacted ACA was observed by HPLC after 61 h in all three reactions (with 50 mM, 5 mM and 0.5 mM NADPH) (Figure 2.36), the NMR spectra indicate that given sufficient reaction time (72 h), quantitative conversion of ACA to 3-HP can be achieved (Figure 2.39). Collectively, results obtained from these experiments is clearly indicative of an inhibitory effect on Cg10062(E114N) caused by increasing concentrations of cofactor.

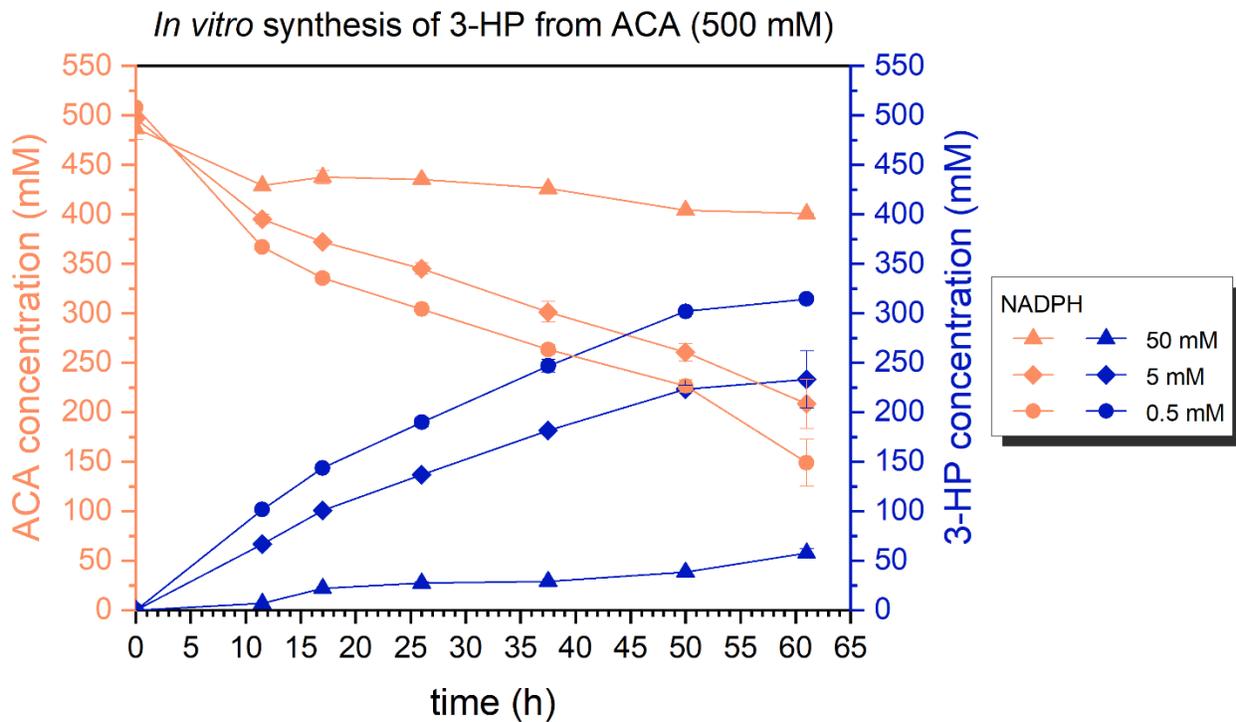


Figure 2.36. *In vitro* synthesis of 3-HP from 500 mM ACA. Conversion of ACA to 3-HP was carried out using three different molar equivalents of NADP(H) relative to ACA: 50 mM (triangles), 5 mM (diamonds), and 0.5 mM (circles). ACA (orange) and 3-HP (blue) were quantified by HPLC.

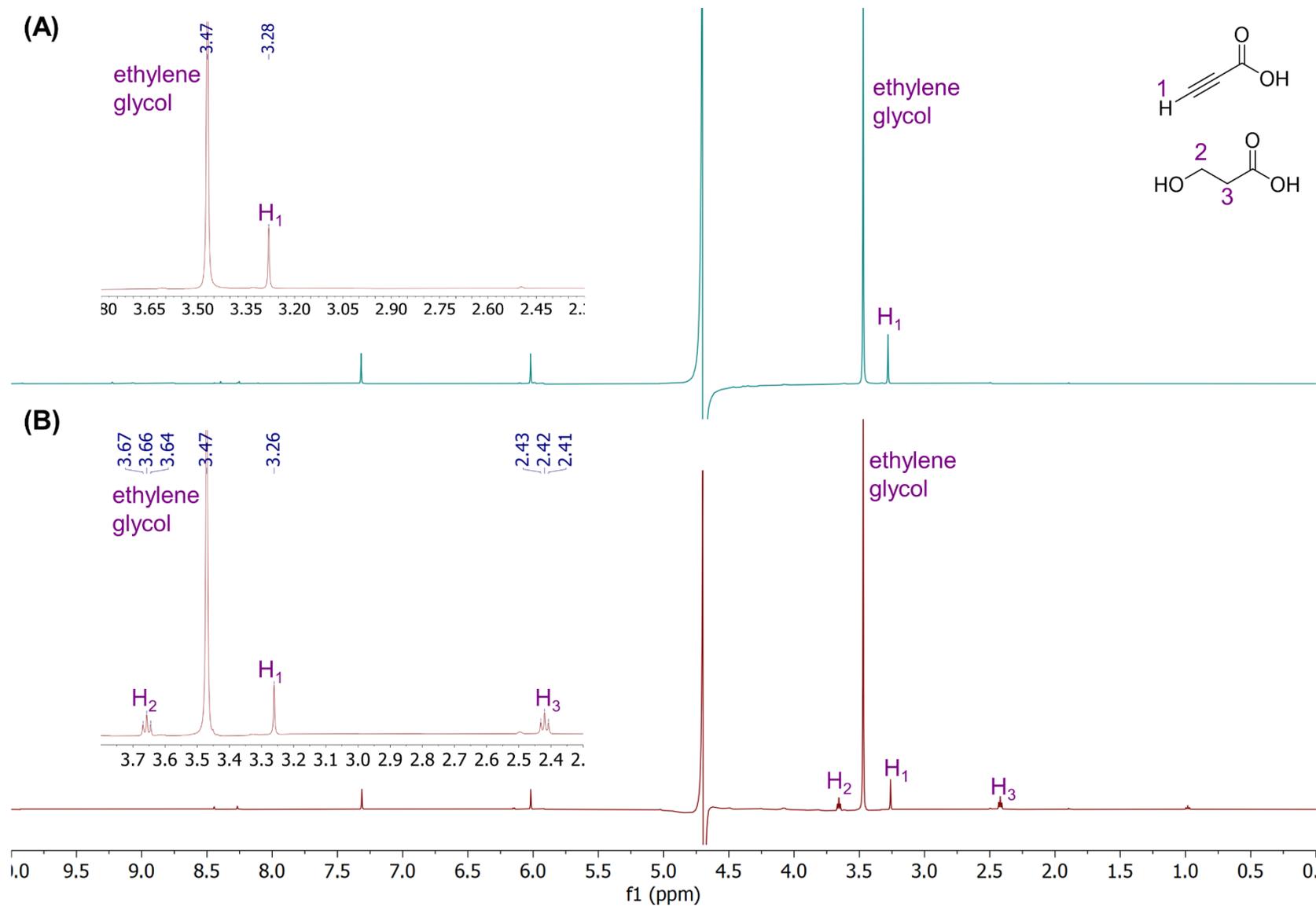


Figure 2.37. ^1H NMR of 3-HP synthesis with 500 mM ACA and 50 mM NADP(H). ^1H NMR spectra of the reaction at (A) 0 h and (B) 72 h. The panels on the left are zoomed in regions of each spectra indicating the ACA and 3-HP signals.

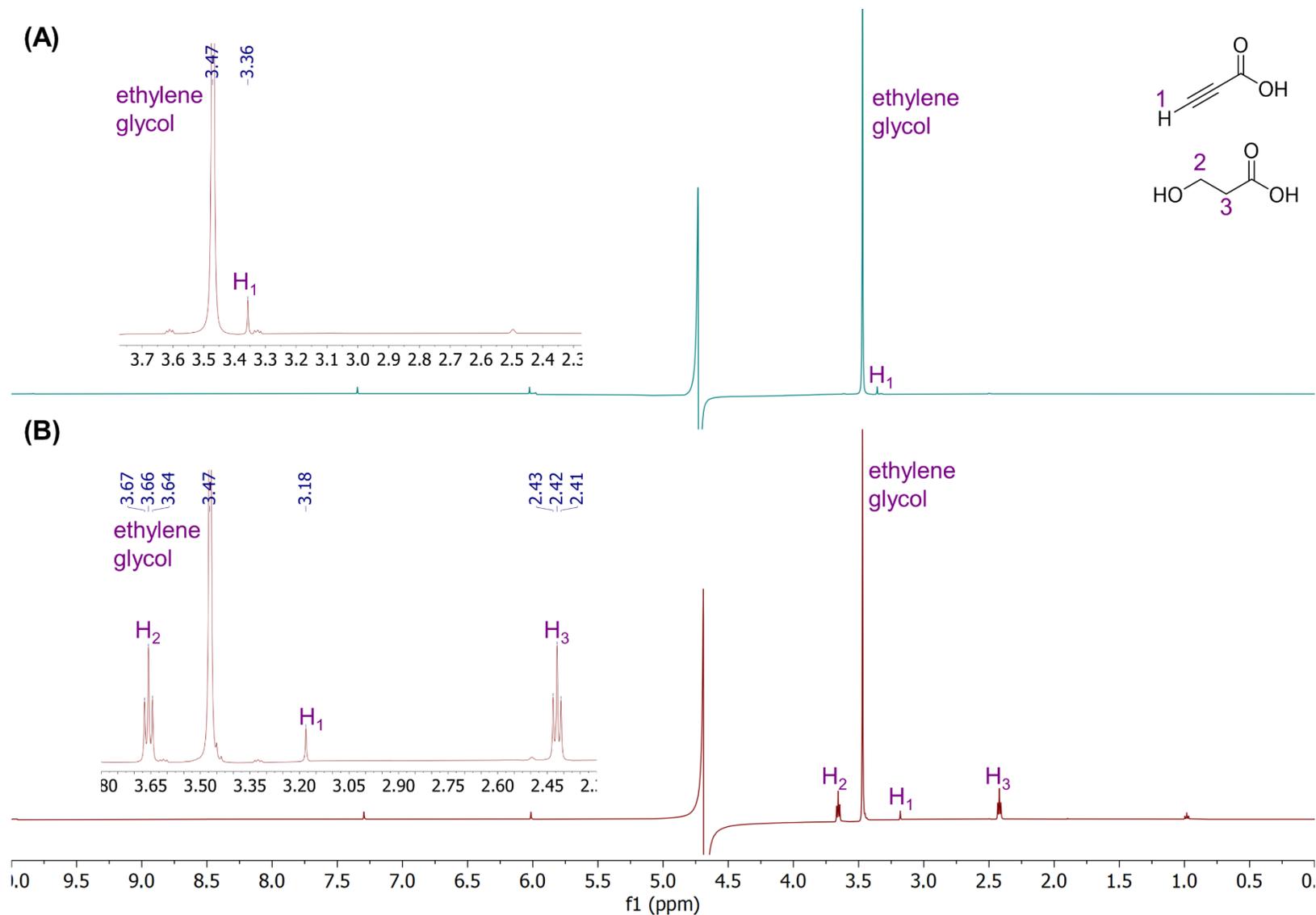


Figure 2.38. ^1H NMR of 3-HP synthesis with 500 mM ACA and 5 mM NADP(H). ^1H NMR spectra of the reaction at (A) 0 h and (B) 72 h. The panels on the left are zoomed in regions of each spectra indicating the ACA and 3-HP signals.

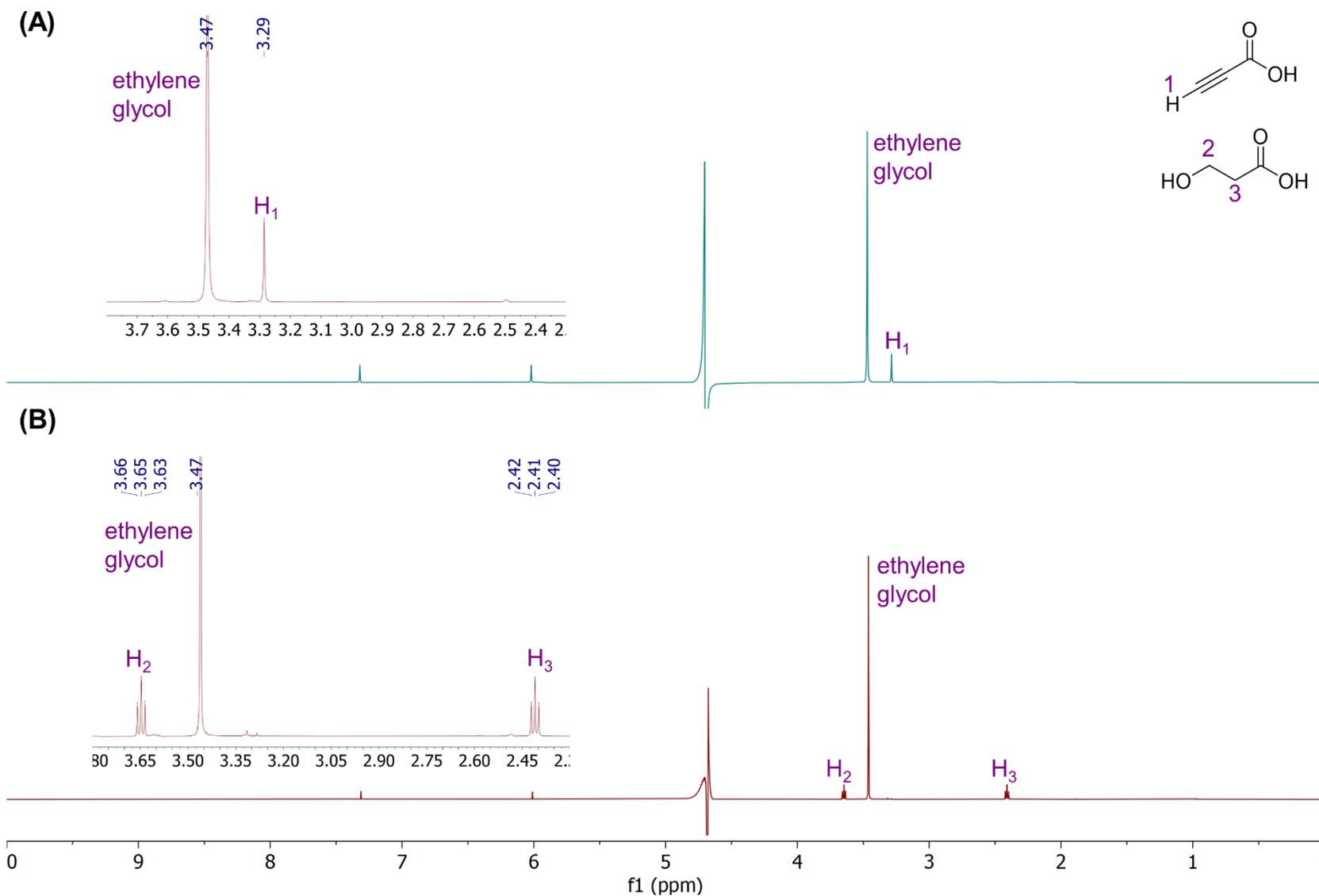


Figure 2.39. ^1H NMR of 3-HP synthesis with 500 mM ACA and 0.5 mM NADP(H). ^1H NMR spectra of the reaction at (A) 0 h and (B) 72 h. The panels on the left are zoomed in regions of each spectra indicating the ACA and 3-HP signals.

2.7. Cg10062(E114N) inhibition by NADP(H)

The results obtained from *in vitro* 3-HP synthesis from ACA indicated that the rates of ACA hydration declined with increasing NADP(H) concentrations, pointing towards inhibition of Cg10062(E114N) by the cofactor. Enzyme inhibition is an important regulatory process in biological systems and is typically classified into three main inhibition mechanisms: competitive, non-competitive, and uncompetitive (Figure 2.40). In a competitive inhibition process, the substrate (S) and the inhibitor (I) compete for binding at the catalytic site of the free enzyme (E).⁷⁷ Competitive binding is often observed when the inhibitor (I) bears close structural resemblance to the substrate (S) and may be reversed by significantly increasing the substrate concentration.^{77,78} Experimentally, competitive inhibition of enzymes lead to an increase in the K_m since a higher substrate concentration is required to out-compete the inhibitor binding at the active site.⁷⁸ However, this competitive binding does not affect the V_{max} of the given enzyme.⁷⁸ In contrast, non-competitive binding occurs when the inhibitor interacts with a different site on the enzyme that the substrate does not bind to. Non-competitive binding cannot be reversed using increased substrate concentrations since the inhibitor and substrate bind at separate sites on the enzyme.⁷⁸ Thus, the amount of inhibited enzyme (EI) remains unchanged over a range of substrate concentrations, leading to an overall decrease in V_{max} . Nevertheless, the K_m remains unaltered in the presence of the inhibitor since competitive binding does not occur at the catalytic site.^{77,78} The third mode of inhibition is uncompetitive and involves binding of the inhibitor only to the enzyme-substrate (ES) complex, leading to reduced V_{max} and K_m values.^{77,78} The formation of the enzyme-substrate-inhibitor (ESI) complex typically does not allow product formation, likely due to conformation changes of the enzyme, resulting in the reduced V_{max} . The reduced K_m value is a result of the increase in the inactive ESI complex, causing a shift in equilibrium to increase the concentration of the ES complex, thus resulting in less free enzyme.⁷⁸

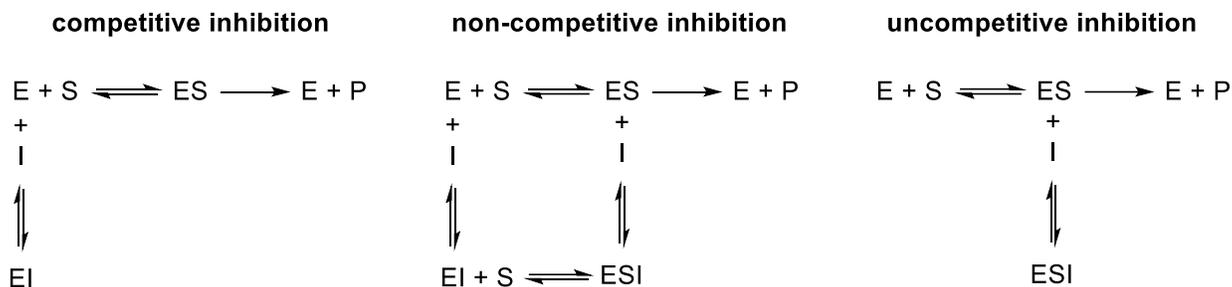


Figure 2.40. Modes of enzyme inhibition

The inhibition constant K_i is the equilibrium constant for the dissociation of the enzyme-inhibitor complex and is a measure of the potency of an inhibitor independent of the enzyme concentration.⁷⁹ To determine the K_i experimentally, the initial rates of the enzyme-catalyzed reaction are determined while varying the substrate concentration without the addition of any inhibitor. The kinetics experiments are then repeated for the same range of substrate concentrations, with the addition of a fixed concentration of inhibitor for each assay. This process is repeated with 4–6 inhibitor concentrations. A non-linear curve fit using a software such as Origin can then be used to determine the K_i . Since Cg10062(E114N) does not require a nicotinamide cofactor for catalysis, it was coupled to MSAD and NADH-dependent ADH to determine enzyme activity using the assay shown below (Figure 2.41). In order to determine the K_i , varied concentrations of inhibitor NADPH must be used in addition to the NADH required for the coupled enzyme assay. Since both cofactors absorb strongly at 340 nm ($\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$), the Cg10062 coupled enzyme assay could not be used for K_i determination. Instead, multiple attempts to determine the K_i of NADPH inhibition of Cg10062(E114N) were made using a commercial kit that quantifies NADH via fluorescence (see Section 4.15.11). The PicoProbe™ NADH Fluorometric Assay Kit (K338) utilizes a NADH cycling enzyme and a fluorometric probe that is highly specific for NADH and does not detect NAD^+ or NADPH (Figure 2.42).

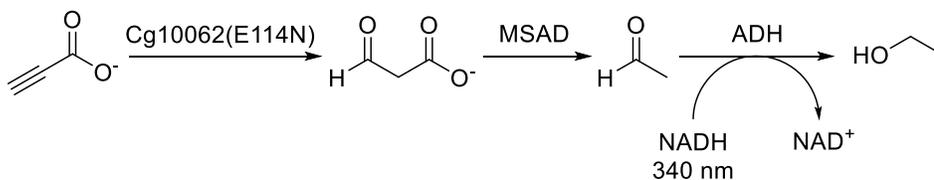
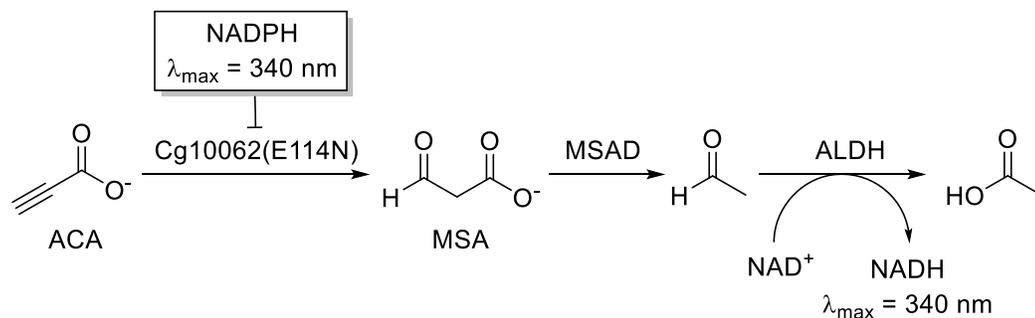


Figure 2.41. The coupled enzyme assay used to measure Cg10062(E114N) activity.

The Cg10062 coupled enzyme assay shown above was modified to replace the NADH-dependent alcohol dehydrogenase (ADH) with a NAD⁺-dependent aldehyde dehydrogenase (ALDH) (Figure 2.42). The kinetics were carried out using a discontinuous assay in which samples (200 μ L) containing all the assay components (except ACA) were combined in microfuge tubes. Once initiated with the addition of ACA, samples were quenched with 3 volumes of ACN (600 μ L) after 20 s. Quenching the enzyme assays using a strong acid was not suitable due to the instability of NADH at low pH.⁸⁰ Following manufacturer's instructions, samples were incubated at 60 °C to degrade unreacted NAD⁺. All samples were incubated with a master mix of PicoProbe™ and the NADH cycling enzyme mix in a 96-well microtiter plate (black). The NADH is oxidized by PicoProbe™, resulting in a fluorescent signal and the NADH cycling enzyme ensures amplification of the fluorescent signal by continuous reduction of NAD⁺ to NADH. Therefore, NAD⁺ degradation is critical in avoiding overestimation of NADH formed by the modified Cg10062 assay (Figure 2.42). The 96-well plate was sealed with an aluminum seal and incubated for 30 min at 25 °C, prior to measuring the fluorescence (ex/em = 535/587 nm). Further information regarding the structure, identity, or mechanism of action of PicoProbe™ and the NADH cycling enzyme mix is considered proprietary information and were not made available with purchase of the kit. Fluorescence readings obtained for each sample was used to calculate the NADH concentrations using a standard curve that was prepared using the NADH standard provided with the kit (Figure 2.43).



Step 1: degrade NAD⁺ at 60 C (30 min)

Step 2: Addition of PicoProbe and NADH cycling enzyme

Step 3: Quantify ex/em = 535/587 nm

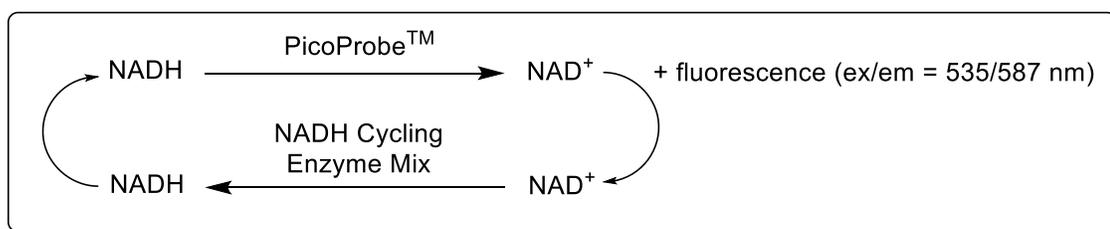


Figure 2.42. Quantification of NADH using PicoProbeTM. ADH was replaced with ALDH in the Cg10062 enzyme assay which was prepared in 100 mM sodium phosphate pH 8. Cg10602(E114N) assays were quenched with ACN. The kinetic assays were repeated with the addition of 3-4 NADPH concentrations. NADH quantification of all samples were carried out following instructions provided with the PicoProbeTM NADH Fluorometric Assay Kit.

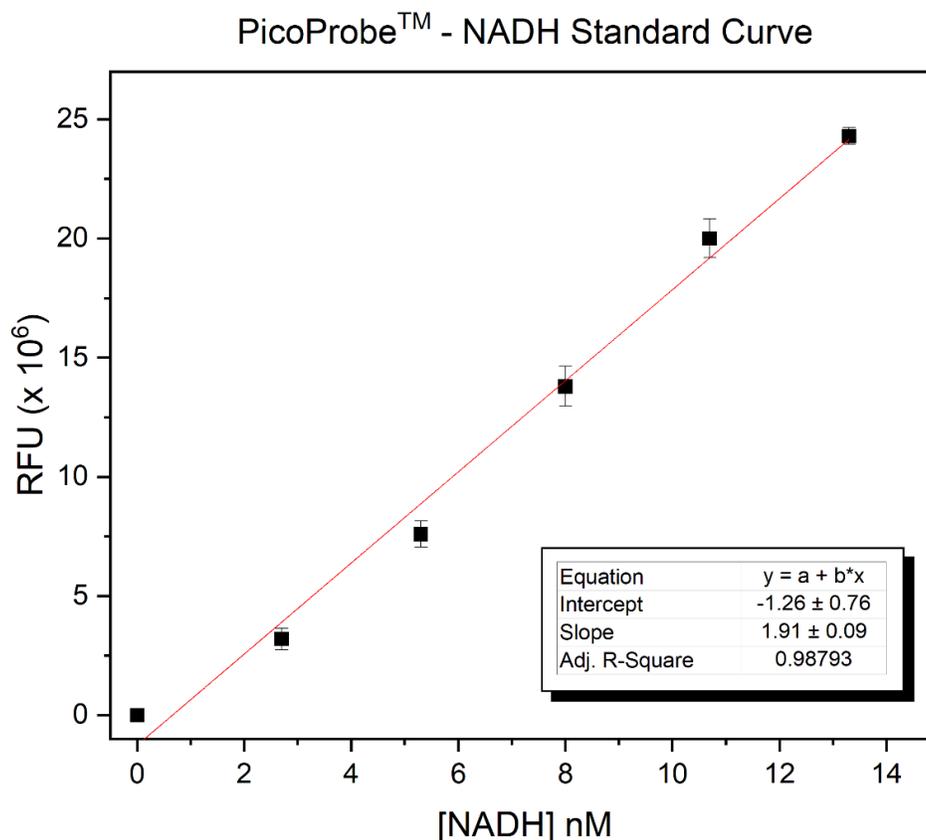


Figure 2.43. PicoProbe™- NADH standard curve

Prior to measuring the kinetics using a range of ACA concentrations, an initial screen was carried out using several concentrations of NADPH to determine a suitable range of NADPH concentrations to be used to observe inhibition. The modified coupled enzyme assay (Figure 2.42) was carried out using 45 μM ACA ($K_{m(\text{ACA})}$ of Cg10062(E114N) and four concentrations of NADPH (10, 20, 30 and 40 mM). The concentrations of NADH detected in the quenched samples correspond to the initial rate of ACA hydration (0–20 s) by Cg10062(E114N) since both coupling enzymes, MSAD and ALDH, were in excess. Thus, decreasing concentrations of NADH correlate to lower Cg10062(E114N) activity. The declining NADH concentrations at higher NADPH levels confirmed the inhibitory effect of NADPH on Cg10062(E114N) (Figure 2.44).

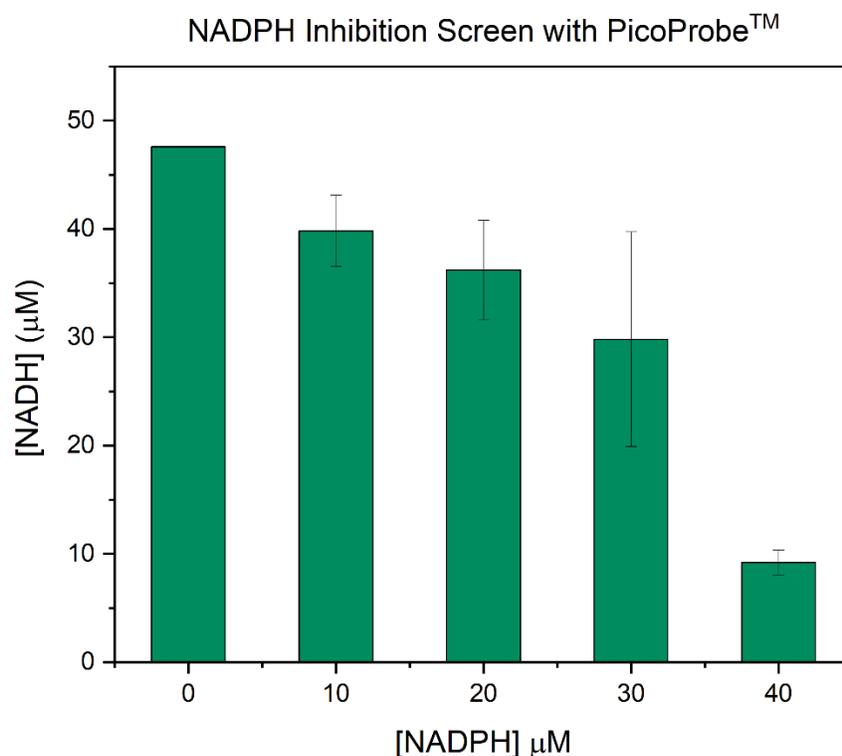


Figure 2.44. NADPH Inhibition Screen for Cg10062(E114N) using PicoProbe™. The concentration of ACA (45 µM) and all other assay components (except NADPH) were kept constant.

Based on the initial screen for inhibiting concentrations of NADPH, several attempts were made to determine K_i using the methods described above with varying NADPH concentrations (0, 10, 20 and 30 mM) were unsuccessful. The accurate measurement of Michaelis Menten kinetics using this discontinuous assay was challenging and the expected increase in initial velocity with increasing substrate concentrations was not observed (i.e., datapoints did not fit the Michaelis Menten model even in the absence of inhibitor; data not shown).

Instead, the IC_{50} of NADPH was determined as a measure of its inhibitory effect on Cg10062(E114N). The IC_{50} is defined as the inhibitor concentration required to halve the V_{max} of an enzyme under specified assay conditions and is a relative parameter that is dependent on the enzyme and substrate concentrations used.⁷⁹ As with the K_i , smaller IC_{50} values indicate stronger inhibition of the enzyme. However, in comparison with the K_i , the IC_{50} is obtained by measuring the initial rates of the enzyme-catalyzed reaction at a fixed concentration of substrate across a

range of inhibition concentrations. The inhibition of Cg10062(E114N) was monitored in the presence of ACA (500 μ M) and a range of NADPH concentrations (0–65 μ M) (Figure 2.45). While 85% hydratase activity is maintained in the presence of 20 μ M NADPH, 30 μ M NADPH caused approximately 55% decrease in Cg10062(E114N) activity. Under the specified assay conditions for this experiment, the IC_{50} of NADPH leading to a 50% reduction in hydratase activity of Cg10062(E114N) was determined to be 27.6 μ M. Furthermore, since Cg10062(E114N) was saturated with ACA, the initial rates measured correspond to the V_{max} of the enzyme. The decrease in V_{max} with increasing NADPH rules out competitive inhibition. However, these assays alone do not indicate whether NADPH inhibits Cg10062(E114N) through uncompetitive or non-competitive inhibition.

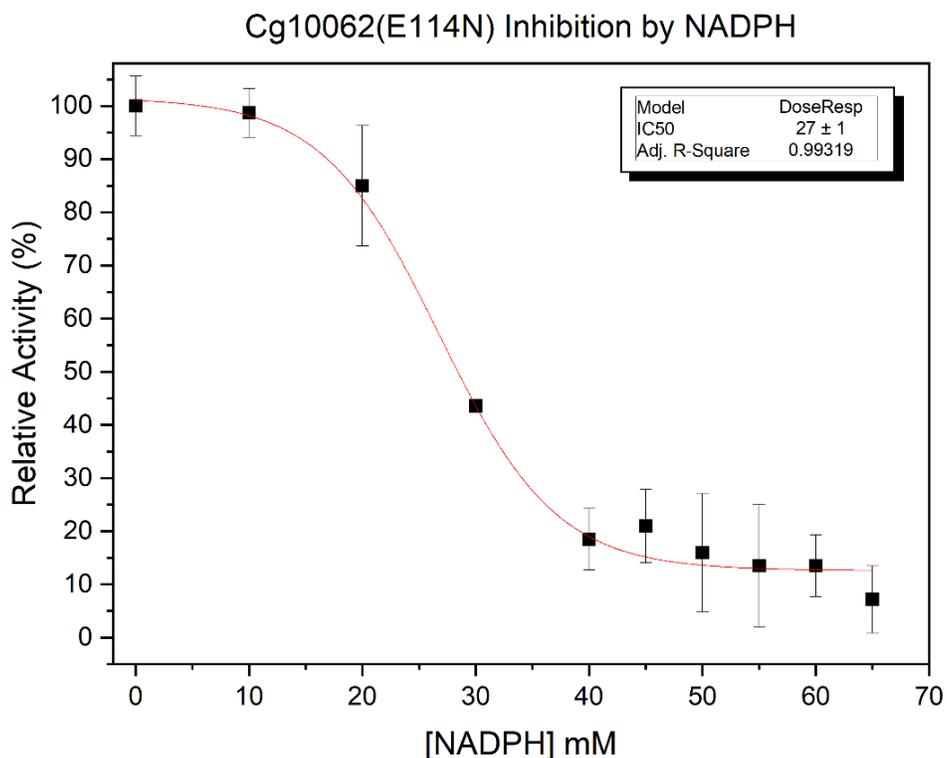


Figure 2.45. IC_{50} of NADPH for Cg10062(E114N) inhibition

2.8. *In vivo* synthesis of 3-HP

Following successful *in vitro* synthesis of 3-HP, preliminary studies were carried out to investigate *in vivo* production of 3-HP by *E. coli*. Plasmid pAS3.112 (Figure 2.46) harboring the genes expressing Cg10062(E114N) and YdfG downstream of an IPTG-inducible *trc* promoter was used to evaluate 3-HP formation in rich (LB) and minimal (M9) media. PTDH was not included on the plasmid for cofactor recycling and the biocatalytic synthesis of 3-HP from ACA was solely reliant on the intracellular availability of NADPH. A *trc* promoter was used in place of a T7 promoter to reduce metabolic burden on *E. coli* during protein expression.⁸¹ This consideration was especially important for the cells cultured in M9 media, which contain a defined and limited amount of resources for cell growth and biomass production. With the use of a T7 promoter, nutrients and precursors that would otherwise be available for energy production and cell growth are channeled towards the highly energy-demanding expression of the recombinant proteins encoded by the plasmid. BL21/pAS3.112 was cultured in 25 mL M9 or LB media (in baffled shake flasks). The media was supplemented with 100 mM ACA allowing the cells to utilize ACA for 3-HP synthesis while relying on glucose for cell growth. The LB cultures were grown for 9 h after IPTG induction and the M9 cultures were allowed 72 h after IPTG induction since cell growth appeared significantly slower relative to the LB cultures.

In both rich and minimal media, trace quantities of 3-HP were observed in the media using ¹H NMR. The presence of 3-HP was detected 9 h after induction in LB media (Figure 2.47). Trace quantities of 3-HP were also observed without induction, most likely due to leaky expression associated with IPTG-inducible promoters.⁸² Similarly, 3-HP was detected in trace amounts in both the uninduced and IPTG-induced M9 cultures at 72 h (Figure 2.48).

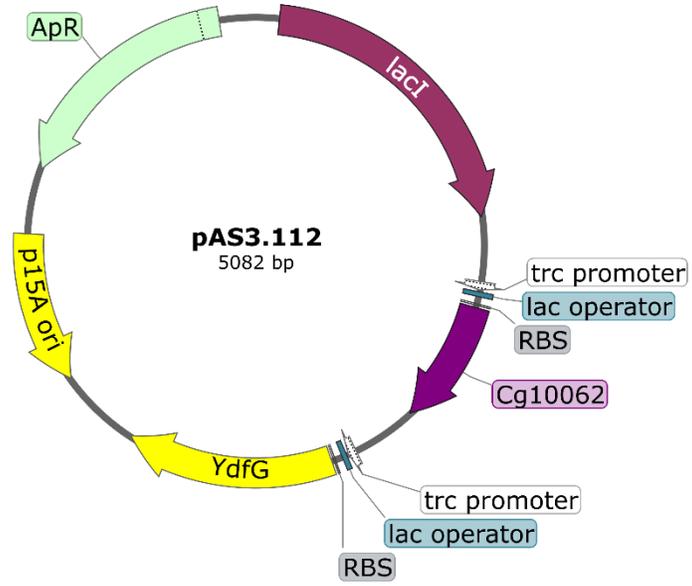


Figure 2.46. Plasmid map of pAS3.112 was used for *in vivo* synthesis of 3-HP.

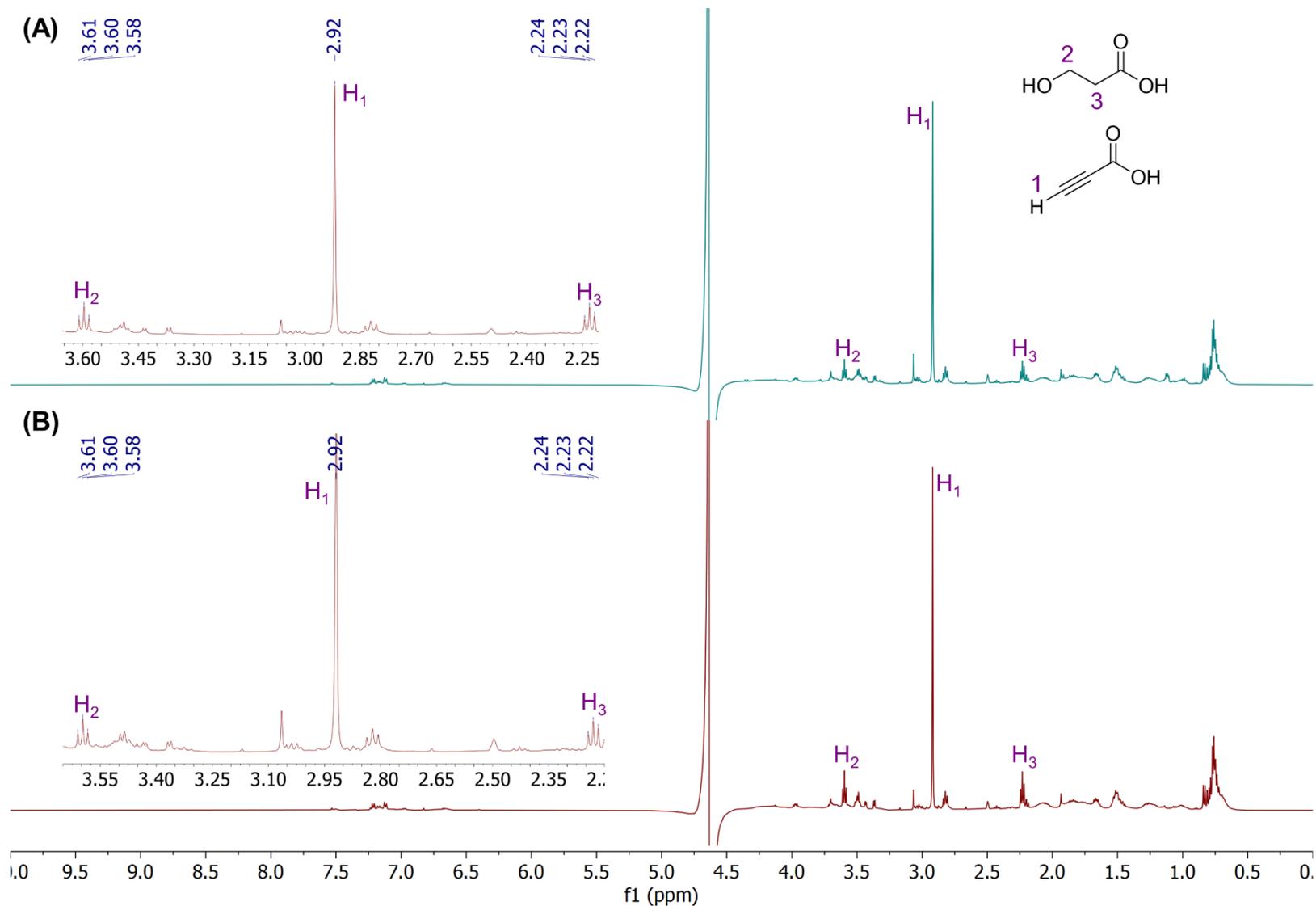


Figure 2.47. *In vivo* 3-HP synthesis by BL21/pAS3.112 in LB media supplemented with ACA. ^1H NMR spectra of the media from the (A) uninduced culture and (B) IPTG-induced culture after 9 h. The panels on the left are zoomed in regions of indicating the ACA and 3-HP signals.

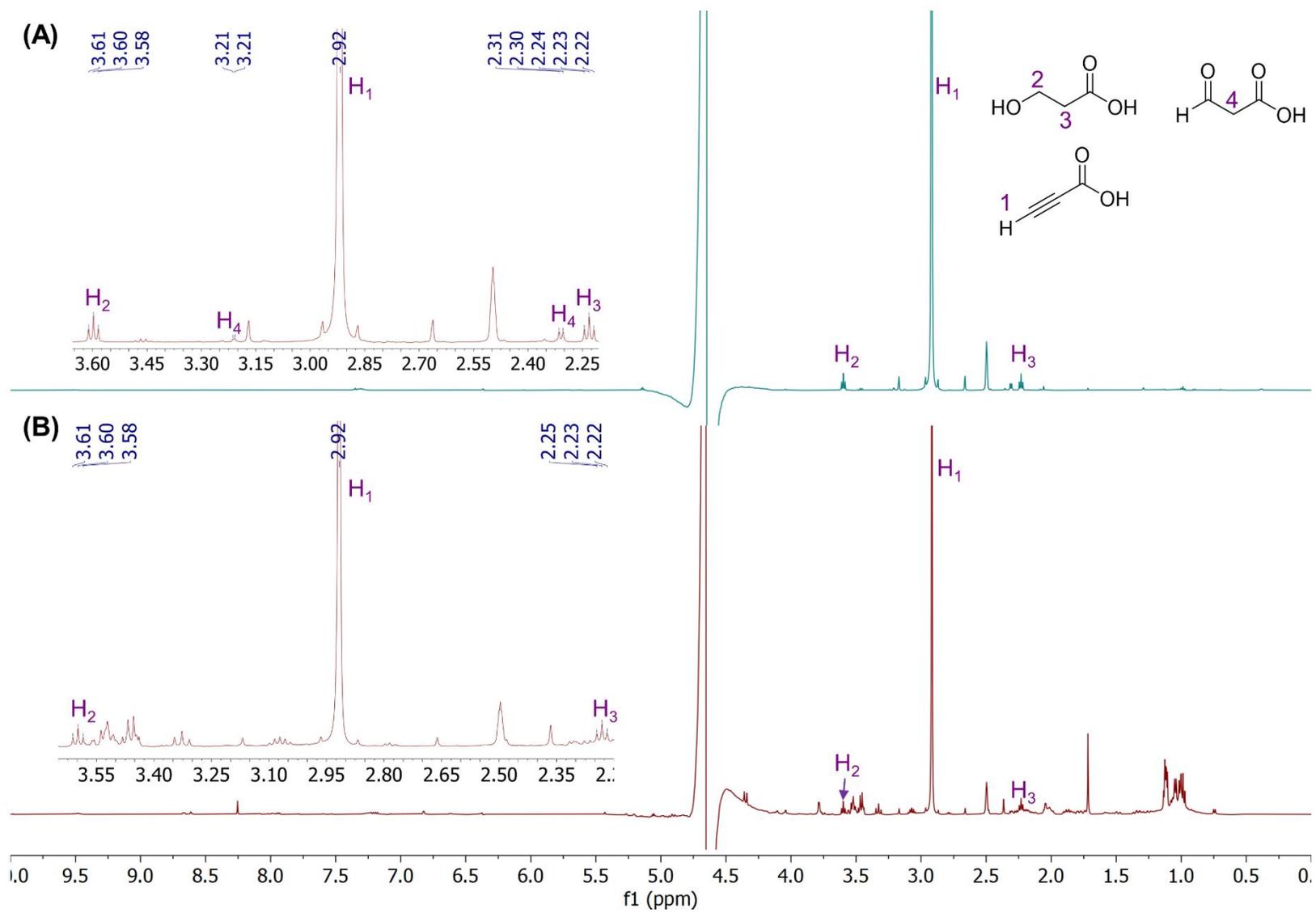


Figure 2.48. *In vivo* 3-HP synthesis by BL21/pAS3.112 in M9 media supplemented with ACA. ^1H NMR spectra of the media from the (A) uninduced culture and (B) IPTG-induced culture after 72 h. The panels on the left are zoomed in regions of indicating the ACA and 3-HP signals.

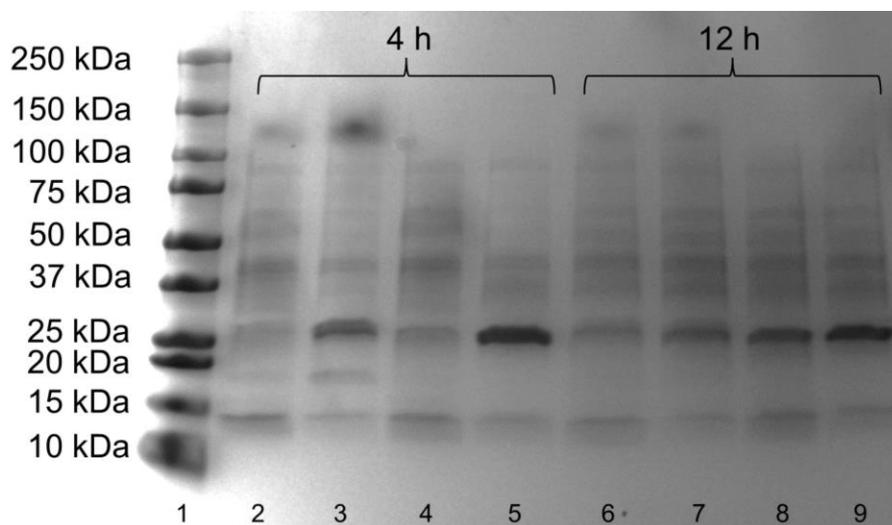


Figure 2.49. Expression of Cg10062(E114N and YdfG by BL21/pAS3.112 in M9 media. Samples were analyzed at 4 h and 12 h after IPTG induction. BL21/pBbA1a-RFP expressing RFP was used as a control. Lanes 1: 10 kDa ladder; 2: lysate of uninduced BL21/pBbA1a-RFP (4 h); 3: lysate of IPTG-induced BL21/pBbA1a-RFP (4 h); 4: lysate of uninduced BL21/pAS3.112 (4 h); 5: lysate of induced BL21/pAS3.112 (4 h); 6: lysate of uninduced BL21/pBbA1a-RFP (12 h); 7: lysate of IPTG-induced BL21/pBbA1a-RFP (12 h); 8: lysate of uninduced BL21/pAS3.112 (12 h); 9: lysate of IPTG-induced BL21/pAS3.112 (12 h). Bands at approximately 25 kDa and 28 kDa correspond to RFP and YdfG, respectively. Prominent band indicating Cg10062(E114N) (19 kDa) expression was not observed.

Although 3-HP was only detected in trace amounts by NMR, these experiments are the first indication of 3-HP synthesis by *E. coli* supplemented with ACA. To examine if the inefficiency of 3-HP formation was caused by poor expression of Cg10062(E114N) and YdfG, crude cell lysates of BL21/pAS3.112 cultured in M9 media were visualized by SDS-PAGE (Figure 2.49). A control strain, BL21/pBbA1a-RFP, harboring the *rfp* gene downstream of P_{trc} was included. Expression of each enzyme was monitored in the absence and presence of 1 mM IPTG at two time intervals (4 and 12 h). Expression of the RFP protein (25 kDa) was observed within 4 h of induction of BL21/pBbA1a-RFP and after 12 h of growth, a faint band corresponding to RFP was observed in the uninduced culture. Expression of YdfG (28 kDa) by BL21/pAS3.112 was seen within 4 h of IPTG induction, although a band corresponding to Cg10062(E114N) monomer (19 kDa) was not detected. Similar to RFP expression under uninduced conditions, YdfG expression was seen at 12 h of growth in uninduced BL21/pAS3.112. The absence of a prominent band for

Cg10062(E114N) suggests inefficient expression of the protein under the growth conditions used for the *in vivo* experiments and the presence of unreacted ACA in the cell cultures may also be attributed to poor efficiency of ACA transport into the cell or low hydratase activity by Cg10062(E114N). Further optimization of the plasmid using different promoters and transcriptional elements will have to be considered to improve protein expression to facilitate efficient 3-HP synthesis from ACA. It must also be noted that the results of these experiments do not provide any indication of the efficiency of ACA uptake by *E. coli*.

2.9. Conclusion

Since 3-HP was first identified as a top platform chemical to be produced from renewable sugars in 2004, the number of studies reporting the development of biocatalytic syntheses of 3-HP has risen exponentially. Although extensive work has been carried out using glycerol and glucose substrates, reaching high titers in a variety of recombinant microbes, to date, none of these processes have been implemented as an industrially relevant technology. Considerable efforts are required in terms of optimizing titers, productivity, and yield. Herein, ACA was presented as a highly attractive substrate for 3-HP synthesis in a relatively simple, two-step, three-enzyme pathway. ACA can be derived from CH₄ and CO₂ via methane dehydrodimerization and subsequent carboxylation of the resulting acetylene. ACA obtained via this process represents three “fixed” carbon atoms; two from CH₄ and one from CO₂. Relative to glucose, ACA does not have any established uses in the food industry and has previously not been used for the production of commodity chemicals or building block chemicals such as 3-HP. Furthermore, while glucose and glycerol are examples of plant-derived carbon fixation, ACA does not require the same resources that would otherwise be needed for industrial farming to generate carbon feedstock. Additionally, both CH₄ and CO₂ are highly potent greenhouse gases that are considered waste gases where CO₂ is usually released into the atmosphere and excess CH₄ from hydraulic fracturing is often flared, thereby producing more CO₂. Instead, their use in ACA production is a highly preferable alternative both economically and ecologically.

The use of the novel variant Cg10062(E114N) to produce MSA from ACA, adds to the economic value of the synthetic process since the tautomerase does not require any metal cofactors or coenzymes. Additionally, all three enzymes used in this study maintained catalytic activity for over a year when stored at -20 °C with a cryoprotectant such as ethylene glycol, eliminating the need for routine enzyme purification for an *in vitro* process. Quantitative conversion of ACA to 3-HP was observed starting with both 100 mM and 500 mM starting material. To address the requirement of the expensive NADPH cofactor, PTDH was used to ensure continuous regeneration of NADP(H), resulting in a highly efficient 3-HP production process. Since the work described in this chapter includes batch production of 3-HP *in vitro*, members of the Draths group have begun developing an enzyme-immobilized continuous flow synthesis of 3-HP using the pathway described here. As an alternative, 3-hydroxyisobutyrate dehydrogenase (MmsB) native to *Pseudomonas putida* KT2440 is also being explored for reduction of MSA to 3-HP, as it utilizes NADH for catalysis as opposed to NADPH-dependent YdfG.⁸³ In addition to being able to use a relatively cheaper cofactor, MmsB displayed higher specific activity for MSA relative to YdfG (results not shown). A highly active enzyme for MSA reduction is preferred in order to maintain excellent mass balance and minimize spontaneous decarboxylation of MSA, prior to its transformation to 3-HP. Since Cg10062(E114N) was inhibited by high concentrations of NADPH, the use of NADH-dependent MmsB is preferred over YdfG. However, it remains to be investigated whether NADH itself inhibits Cg10062(E114N).

Preliminary studies have also been carried out to demonstrate the production of 3-HP *in vivo* in trace quantities, provided glucose is added to the media for cell growth. Analysis of protein expression and detection of low levels of 3-HP confirmed that further optimization is required before this pathway can be implemented in a platform organism such as *E. coli*. The most important requirement for successful synthesis of 3-HP using a recombinant microbe is the efficient uptake of ACA into the intracellular matrix. Preliminary work has indicated that ACA does not appear to be toxic to *E. coli* (see chapter one). However, *E. coli* is unable to grow efficiently

on exclusively ACA and transport of ACA into the cells must be addressed. Based on the current body of literature, other significant limitations associated with *in vivo* synthesis of 3-HP is the toxicity of 3-HP to cells and the recovery of 3-HP from fermentation broth. Most developed biocatalytic routes include the formation of either 3-HPA or MSA as the immediate precursor to 3-HP. The accumulation of either compound is toxic to cell growth and hinders 3-HP production overall.^{4,27,31,32,36} The toxicity of MSA and 3-HPA has been attributed to covalent modification of enzymes due to their reactivity with the side-chains of amino acid residues such as lysine (amino group), cysteine (SH) and histidine (imidazole) residues.^{84–86} However, this limitation can be addressed by ensuring that the downstream enzyme responsible for MSA reduction or 3-HPA oxidation is active and expressed in sufficient quantities by the cell.^{4,31,32,36} Moreover, like most other organic acids, high concentrations of 3-HP can cause disruption of the pH-balance of the cells leading to growth inhibition, where concentrations of over 200 mM have been observed to significantly impact growth of recombinant *K. pneumoniae* and strains of *E. coli*.^{32,87} In part, these drawbacks have been addressed by using acid-tolerant strains of *E. coli* or conducting fermentation processes at a pH lower than the pK_a of 3-HP (4.51).^{88–92} The hydrophilicity of 3-HP makes 3-HP extraction challenging, Recently, Cargill patented a novel 3-HP recovery method from fermentation broth, which comprises of concentrating the fermentation media by evaporation and acidification of the resulting media to facilitate extraction of 3-HP.⁹³ In conclusion, despite the growing body of research on developing sustainable syntheses of 3-HP, thorough examination of the existing pathways must be carried out to address the limitations with regard to 3-HP toxicity (which in turn affect titers) and successful extraction of 3-HP from fermentation media.

REFERENCES

REFERENCES

- (1) de Fouchécour, F.; Sánchez-Castañeda, A.-K.; Saulou-Bérion, C.; Spinnler, H. É. Process Engineering for Microbial Production of 3-Hydroxypropionic Acid. *Biotechnol. Adv.* 2018, 36 (4), 1207–1222.
- (2) Werpy, T.; Petersen, G. *Top Value Added Chemicals from Biomass: Volume I -- Results of Screening for Potential Candidates from Sugars and Synthesis Gas*; DOE/GO-102004-1992, 15008859; 2004; p DOE/GO-102004-1992, 15008859. <https://doi.org/10.2172/15008859>.
- (3) Zhang, D.; Hillmyer, M. A.; Tolman, W. B. A New Synthetic Route to Poly[3-Hydroxypropionic Acid] (P[3-HP]): Ring-Opening Polymerization of 3-HP Macrocyclic Esters. *Macromolecules* 2004, 37 (22), 8198–8200.
- (4) Kumar, V.; Ashok, S.; Park, S. Recent Advances in Biological Production of 3-Hydroxypropionic Acid. *Biotechnol. Adv.* 2013, 31 (6), 945–961.
- (5) Bhagwat, S. S.; Li, Y.; Cortés-Peña, Y. R.; Brace, E. C.; Martin, T. A.; Zhao, H.; Guest, J. S. Sustainable Production of Acrylic Acid via 3-Hydroxypropionic Acid from Lignocellulosic Biomass. *ACS Sustain. Chem. Eng.* 2021, 9 (49), 16659–16669.
- (6) Ohara, T.; Sato, T.; Shimizu, N.; Prescher, G.; Schwind, H.; Weiberg, O.; Marten, K.; Greim, H.; Shaffer, T. D.; Nandi, P. Acrylic Acid and Derivatives. In *Ullmann's Encyclopedia of Industrial Chemistry*; John Wiley & Sons, Ltd, 2020; pp 1–21. https://doi.org/10.1002/14356007.a01_161.pub4.
- (7) Dishisha, T.; Pyo, S.-H.; Hatti-Kaul, R. Bio-Based 3-Hydroxypropionic- and Acrylic Acid Production from Biodiesel Glycerol via Integrated Microbial and Chemical Catalysis. *Microb. Cell Factories* 2015, 14 (1), 200.
- (8) *Cargill cooperation*. <http://www.novozymes.com/en/news/news-archive/2008/01/44469> (accessed 2022-06-14).
- (9) Bidy, M. J.; Scarlata, C.; Kinchin, C. *Chemicals from Biomass: A Market Assessment of Bioproducts with Near-Term Potential*; NREL/TP--5100-65509, 1244312; 2016; p NREL/TP--5100-65509, 1244312. <https://doi.org/10.2172/1244312>.
- (10) *BASF, Cargill and Novozymes achieve milestone in bio-based acrylic acid process | Cargill*. <https://www.cargill.com/news/releases/2013/NA3075513.jsp> (accessed 2021-06-05).
- (11) *BASF exits bio-acrylic acid partnership with Novozymes and Cargill - bioplastics MAGAZINE*. <https://www.bioplasticsmagazine.com/en/news/meldungen/20150202Cargill-Novozymes.php> (accessed 2022-06-14).

- (12) *Novozymes, Cargill continue bio-acrylic acid partnership as BASF exits.* <http://www.novozymes.com/en/news/news-archive/2015/01/novozymes-cargill-continue-bio-acrylic-acid-partnership-basf-exits> (accessed 2022-06-15).
- (13) *Procter & Gamble and Cargill collaborate to bring nature-powered innovation, fueling the future for more sustainable products | Cargill.* <https://www.cargill.com/2020/procter-gamble-and-cargill-collaborate-to-bring-nature-powered> (accessed 2021-06-05).
- (14) *NREL Shifts Carbon Fiber Research into Second Gear.* <https://www.nrel.gov/news/features/2018/nrel-shifts-carbon-fiber-research-into-second-gear.html> (accessed 2022-06-14).
- (15) Karp, E. M.; Eaton, T. R.; Sánchez i Nogué, V.; Vorotnikov, V.; Biddu, M. J.; Tan, E. C. D.; Brandner, D. G.; Cywar, R. M.; Liu, R.; Manker, L. P.; Michener, W. E.; Gilhespy, M.; Skoufa, Z.; Watson, M. J.; Fruchey, O. S.; Vardon, D. R.; Gill, R. T.; Bratis, A. D.; Beckham, G. T. Renewable Acrylonitrile Production. *Science* 2017, 358 (6368), 1307–1310.
- (16) Tang, S.; Hu, C. Design, Preparation and Properties of Carbon Fiber Reinforced Ultra-High Temperature Ceramic Composites for Aerospace Applications: A Review. *J. Mater. Sci. Technol.* 2017, 33 (2), 117–130.
- (17) Meng, F.; McKechnie, J.; Turner, T.; Wong, K. H.; Pickering, S. J. Environmental Aspects of Use of Recycled Carbon Fiber Composites in Automotive Applications. *Environ. Sci. Technol.* 2017, 51 (21), 12727–12736.
- (18) *Energy Department Announces \$11 Million to Advance Renewable Carbon Fiber Production from Biomass.* [Energy.gov. https://www.energy.gov/eere/articles/energy-department-announces-11-million-advance-renewable-carbon-fiber-production](https://www.energy.gov/eere/articles/energy-department-announces-11-million-advance-renewable-carbon-fiber-production) (accessed 2021-06-06).
- (19) Barile, C.; Casavola, C. 19 - Mechanical Characterization of Carbon Fiber-Reinforced Plastic Specimens for Aerospace Applications. In *Mechanical and Physical Testing of Biocomposites, Fibre-Reinforced Composites and Hybrid Composites*; Jawaid, M., Thariq, M., Saba, N., Eds.; Woodhead Publishing Series in Composites Science and Engineering; Woodhead Publishing, 2019; pp 387–407.
- (20) Grasselli, R. K.; Trifirò, F. Acrylonitrile from Biomass: Still Far from Being a Sustainable Process. *Top. Catal.* 2016, 59 (17), 1651–1658.
- (21) Della Pina, C.; Falletta, E.; Rossi, M. Oxidation of Allyl Alcohol in the Presence of a Gold Catalyst: A Route to 3-Hydroxypropionic Acid. *ChemSusChem* 2009, 2 (1), 57–58.
- (22) Matsumoto, M.; Shibasaki, M.; Yuasa, R.; Kondo, K. Salting-out Extraction of 3-Hydroxypropionic Acid with Reactive Extraction and Aqueous Two-Phase Systems. *Solvent Extr. Res. Dev. Jpn.* 2017, 24 (2), 141–147.
- (23) Burgé, G.; Chemarin, F.; Moussa, M.; Saulou-Bérion, C.; Allais, F.; Spinnler, H.-É.; Athès, V. Reactive Extraction of Bio-Based 3-Hydroxypropionic Acid Assisted by Hollow-Fiber Membrane Contactor Using TOA and Aliquat 336 in n-Decanol. *J. Chem. Technol. Biotechnol.* 2016, 91 (10), 2705–2712.

- (24) Moussa, M.; Burgé, G.; Chemarin, F.; Bounader, R.; Saulou-Bérion, C.; Allais, F.; Spinnler, H.-E.; Athès, V. Reactive Extraction of 3-Hydroxypropionic Acid from Model Aqueous Solutions and Real Bioconversion Media. Comparison with Its Isomer 2-Hydroxypropionic (Lactic) Acid. *J. Chem. Technol. Biotechnol.* 2016, 91 (8), 2276–2285.
- (25) Miltenberger, K. Hydroxycarboxylic Acids, Aliphatic. In *Ullmann's Encyclopedia of Industrial Chemistry*; John Wiley & Sons, Ltd, 2000. https://doi.org/10.1002/14356007.a13_507.
- (26) Behr, A.; Botulinski, A.; Carduck, F.-J.; Schneider, M. Process for Preparing Alkali Metal Salts of 3-Hydroxypropionic Acid. US5321156A, June 14, 1994.
- (27) Leonidas, M.; Evangelos, T.; Paul, C. New Trends in Microbial Production of 3-Hydroxypropionic Acid. *Curr. Biochem. Eng.* 2014, 1 (2), 141–154.
- (28) Henry, C. S.; Broadbelt, L. J.; Hatzimanikatis, V. Discovery and Analysis of Novel Metabolic Pathways for the Biosynthesis of Industrial Chemicals: 3-Hydroxypropanoate. *Biotechnol. Bioeng.* 2010, 106 (3), 462–473.
- (29) Jiang, X.; Meng, X.; Xian, M. Biosynthetic Pathways for 3-Hydroxypropionic Acid Production. *Appl. Microbiol. Biotechnol.* 2009, 82 (6), 995–1003.
- (30) Pina, C. D.; Falletta, E.; Rossi, M. A Green Approach to Chemical Building Blocks. The Case of 3-Hydroxypropanoic Acid. *Green Chem.* 2011, 13 (7), 1624–1632.
- (31) Zhao, P.; Tian, P. Biosynthesis Pathways and Strategies for Improving 3-Hydroxypropionic Acid Production in Bacteria. *World J. Microbiol. Biotechnol.* 2021, 37 (7), 117. <https://doi.org/10.1007/s11274-021-03091-6>.
- (32) Vidra, A.; Németh, Á. Bio-Based 3-Hydroxypropionic Acid: A Review. *Period. Polytech. Chem. Eng.* 2018, 62 (2), 156–166.
- (33) Jers, C.; Kalantari, A.; Garg, A.; Mijakovic, I. Production of 3-Hydroxypropanoic Acid from Glycerol by Metabolically Engineered Bacteria. *Front. Bioeng. Biotechnol.* 2019, 7, 124.
- (34) Wang, Q.; Yang, P.; Xian, M.; Feng, L.; Wang, J.; Zhao, G. Metabolic Engineering of *Escherichia coli* for Poly(3-Hydroxypropionate) Production from Glycerol and Glucose. *Biotechnol. Lett.* 2014, 36 (11), 2257–2262.
- (35) Chen, Z.; Huang, J.; Wu, Y.; Wu, W.; Zhang, Y.; Liu, D. Metabolic Engineering of *Corynebacterium glutamicum* for the Production of 3-Hydroxypropionic Acid from Glucose and Xylose. *Metab. Eng.* 2017, 39, 151–158.
- (36) Song, C. W.; Kim, J. W.; Cho, I. J.; Lee, S. Y. Metabolic Engineering of *Escherichia coli* for the Production of 3-Hydroxypropionic Acid and Malonic Acid through β -Alanine Route. *ACS Synth. Biol.* 2016, 5 (11), 1256–1263.
- (37) Yang, F.; Hanna, M. A.; Sun, R. Value-Added Uses for Crude Glycerol--a Byproduct of Biodiesel Production. *Biotechnol. Biofuels* 2012, 5 (1), 13.
- (38) Wang, Z.; Zhuge, J.; Fang, H.; Prior, B. A. Glycerol Production by Microbial Fermentation: A Review. *Biotechnol. Adv.* 2001, 19 (3), 201–223.

- (39) Bobik, T. A.; Havemann, G. D.; Busch, R. J.; Williams, D. S.; Aldrich, H. C. The Propanediol Utilization (Pdu) Operon of *Salmonella enterica* serovar typhimurium LT2 Includes Genes Necessary for Formation of Polyhedral Organelles Involved in Coenzyme B₁₂-Dependent 1,2-Propanediol Degradation. *J. Bacteriol.* 1999, 181 (19), 5967–5975.
- (40) Luo, L. H.; Seo, J.-W.; Baek, J.-O.; Oh, B.-R.; Heo, S.-Y.; Hong, W.-K.; Kim, D.-H.; Kim, C. H. Identification and Characterization of the Propanediol Utilization Protein PduP of *Lactobacillus reuteri* for 3-Hydroxypropionic Acid Production from Glycerol. *Appl. Microbiol. Biotechnol.* 2011, 89 (3), 697–703.
- (41) Sriramulu, D. D.; Liang, M.; Hernandez-Romero, D.; Raux-Deery, E.; Lünsdorf, H.; Parsons, J. B.; Warren, M. J.; Prentice, M. B. *Lactobacillus reuteri* DSM 20016 Produces Cobalamin-Dependent Diol Dehydratase in Metabolosomes and Metabolizes 1,2-Propanediol by Disproportionation. *J. Bacteriol.* 2008, 190 (13), 4559–4567.
- (42) Dishisha, T.; Pereyra, L. P.; Pyo, S.-H.; Britton, R. A.; Hatti-Kaul, R. Flux Analysis of the *Lactobacillus reuteri* Propanediol-Utilization Pathway for Production of 3-Hydroxypropionaldehyde, 3-Hydroxypropionic Acid and 1,3-Propanediol from Glycerol. *Microb. Cell Factories* 2014, 13 (1), 76.
- (43) Zhu, J.-G.; Ji, X.-J.; Huang, H.; Du, J.; Li, S.; Ding, Y.-Y. Production of 3-Hydroxypropionic Acid by Recombinant *Klebsiella pneumoniae* Based on Aeration and ORP Controlled Strategy. *Korean J. Chem. Eng.* 2009, 26 (6), 1679–1685.
- (44) Li, Y.; Wang, X.; Ge, X.; Tian, P. High Production of 3-Hydroxypropionic Acid in *Klebsiella pneumoniae* by Systematic Optimization of Glycerol Metabolism. *Sci. Rep.* 2016, 6, 26932.
- (45) Sabet-Azad, R.; Sardari, R. R. R.; Linares-Pastén, J. A.; Hatti-Kaul, R. Production of 3-Hydroxypropionic Acid from 3-Hydroxypropionaldehyde by Recombinant *Escherichia coli* Co-Expressing *Lactobacillus reuteri* Propanediol Utilization Enzymes. *Bioresour. Technol.* 2015, 180, 214–221.
- (46) Zhang, Y.; Zayed, H. M.; Yun, J.; Zhang, G.; Wang, Y.; Qi, X. Notable Improvement of 3-Hydroxypropionic Acid and 1,3-Propanediol Coproduction Using Modular Coculture Engineering and Pathway Rebalancing. *ACS Sustain. Chem. Eng.* 2021, 9 (12), 4625–4637.
- (47) Johnson, E. A.; Lin, E. C. *Klebsiella pneumoniae* 1,3-Propanediol:NAD⁺ Oxidoreductase. *J. Bacteriol.* 1987, 169 (5), 2050–2054.
- (48) Thi Nguyen, T.; Lama, S.; Kumar Ainala, S.; Sankaranarayanan, M.; Singh Chauhan, A.; Rae Kim, J.; Park, S. Development of *Pseudomonas asiatica* as a Host for the Production of 3-Hydroxypropionic Acid from Glycerol. *Bioresour. Technol.* 2021, 329, 124867.
- (49) Zhao, P.; Ma, C.; Xu, L.; Tian, P. Exploiting Tandem Repetitive Promoters for High-Level Production of 3-Hydroxypropionic Acid. *Appl. Microbiol. Biotechnol.* 2019, 103 (10), 4017–4031.
- (50) Xu, X.; Zhang, G.; Wang, L.; Ma, B.; Li, C. Quantitative Analysis on Inactivation and Reactivation of Recombinant Glycerol Dehydratase from *Klebsiella pneumoniae* XJPD-Li. *J. Mol. Catal. B Enzym.* 2009, 56 (2), 108–114.

- (51) Toraya, T. Radical Catalysis in Coenzyme B₁₂-Dependent Isomerization (Eliminating) Reactions. *Chem. Rev.* 2003, 103 (6), 2095–2128.
- (52) Daniel, R.; Bobik, T. A.; Gottschalk, G. Biochemistry of Coenzyme B₁₂-Dependent Glycerol and Diol Dehydratases and Organization of the Encoding Genes. *FEMS Microbiol. Rev.* 1998, 22 (5), 553–566.
- (53) Nasir, A.; Ashok, S.; Shim, J. Y.; Park, S.; Yoo, T. H. Recent Progress in the Understanding and Engineering of Coenzyme B₁₂-Dependent Glycerol Dehydratase. *Front. Bioeng. Biotechnol.* 2020, 8.
- (54) Mori, K.; Tobimatsu, T.; Hara, T.; Toraya, T. Characterization, Sequencing, and Expression of the Genes Encoding a Reactivating Factor for Glycerol-Inactivated Adenosylcobalamin-Dependent Diol Dehydratase *. *J. Biol. Chem.* 1997, 272 (51), 32034–32041.
- (55) Honda, S.; Toraya, T.; Fukui, S. In Situ Reactivation of Glycerol-Inactivated Coenzyme B₁₂-Dependent Enzymes, Glycerol Dehydratase and Diol Dehydratase. *J. Bacteriol.* 1980, 143 (3), 1458–1465.
- (56) Liu, C.; Ding, Y.; Xian, M.; Liu, M.; Liu, H.; Ma, Q.; Zhao, G. Malonyl-CoA Pathway: A Promising Route for 3-Hydroxypropionate Biosynthesis. *Crit. Rev. Biotechnol.* 2017, 37 (7), 933–941.
- (57) Liu, C.; Wang, Q.; Xian, M.; Ding, Y.; Zhao, G. Dissection of Malonyl-Coenzyme A Reductase of *Chloroflexus aurantiacus* Results in Enzyme Activity Improvement. *PLOS ONE* 2013, 8 (9), e75554.
- (58) Lama, S.; Kim, Y.; Nguyen, D. T.; Im, C. H.; Sankaranarayanan, M.; Park, S. Production of 3-Hydroxypropionic Acid from Acetate Using Metabolically-Engineered and Glucose-Grown *Escherichia coli*. *Bioresour. Technol.* 2021, 320, 124362.
- (59) Zarzycki, J.; Brecht, V.; Müller, M.; Fuchs, G. Identifying the Missing Steps of the Autotrophic 3-Hydroxypropionate CO₂ Fixation Cycle in *Chloroflexus aurantiacus*. *Proc. Natl. Acad. Sci.* 2009, 106 (50), 21317–21322.
- (60) Wang, Y.; Sun, T.; Gao, X.; Shi, M.; Wu, L.; Chen, L.; Zhang, W. Biosynthesis of Platform Chemical 3-Hydroxypropionic Acid (3-HP) Directly from CO₂ in *Cyanobacterium synechocystis* sp. PCC 6803. *Metab. Eng.* 2016, 34, 60–70.
- (61) Jiang, X.-R.; Yan, X.; Yu, L.-P.; Liu, X.-Y.; Chen, G.-Q. Hyperproduction of 3-Hydroxypropionate by *Halomonas bluephagenesis*. *Nat. Commun.* 2021, 12 (1), 1513.
- (62) Yu, S.; Yao, P.; Li, J.; Ren, J.; Yuan, J.; Feng, J.; Wang, M.; Wu, Q.; Zhu, D. Enzymatic Synthesis of 3-Hydroxypropionic Acid at High Productivity by Using Free or Immobilized Cells of Recombinant *Escherichia coli*. *J. Mol. Catal. B Enzym.* 2016, 129, 37–42.
- (63) Fujisawa, H.; Nagata, S.; Misono, H. Characterization of Short-Chain Dehydrogenase/Reductase Homologues of *Escherichia coli* (YdfG) and *Saccharomyces cerevisiae* (YMR226C). *Biochim. Biophys. Acta BBA - Proteins Proteomics* 2003, 1645 (1), 89–94.

- (64) Loh, K. D.; Gyaneshwar, P.; Markenscoff Papadimitriou, E.; Fong, R.; Kim, K.-S.; Parales, R.; Zhou, Z.; Inwood, W.; Kustu, S. A Previously Undescribed Pathway for Pyrimidine Catabolism. *Proc. Natl. Acad. Sci.* 2006, *103* (13), 5114–5119.
- (65) Kim, K.-S.; Pelton, J. G.; Inwood, W. B.; Andersen, U.; Kustu, S.; Wemmer, D. E. The Rut Pathway for Pyrimidine Degradation: Novel Chemistry and Toxicity Problems. *J. Bacteriol.* 2010, *192* (16), 4089–4102.
- (66) Kapust, R. B.; Waugh, D. S. Controlled Intracellular Processing of Fusion Proteins by TEV Protease. *Protein Expr. Purif.* 2000, *19* (2), 312–318.
- (67) DNASU Plasmid | TEVgp1 (*T. etch virus*)
 In pMHTDelta238 (vector expressing the TEV protease). DNASU Plasmids. <http://dnasu.org/DNASU/GetCloneDetail.do?cloneid=84286> (accessed 2022-05-31).
- (68) Mordhorst, S.; N. Andexer, J. Round, Round We Go – Strategies for Enzymatic Cofactor Regeneration. *Nat. Prod. Rep.* 2020, *37* (10), 1316–1333.
- (69) Costas, A. M. G.; White, A. K.; Metcalf, W. W. Purification and Characterization of a Novel Phosphorus-Oxidizing Enzyme from *Pseudomonas stutzeri* WM88*210. *J. Biol. Chem.* 2001, *276* (20), 17429–17436.
- (70) Woodyer, R.; van der Donk, W. A.; Zhao, H. Relaxing the Nicotinamide Cofactor Specificity of Phosphite Dehydrogenase by Rational Design. *Biochemistry* 2003, *42* (40), 11604–11614.
- (71) Woodyer, R.; van der Donk, W. A.; Zhao, H. Optimizing a Biocatalyst for Improved NAD(P)H Regeneration: Directed Evolution of Phosphite Dehydrogenase. *Comb. Chem. High Throughput Screen.* 2006, *9* (4), 237–245.
- (72) Woodyer, R.; Zhao, H.; Donk, W. A. van der. Mechanistic Investigation of a Highly Active Phosphite Dehydrogenase Mutant and Its Application for NADPH Regeneration. *FEBS J.* 2005, *272* (15), 3816–3827.
- (73) Johannes, T. W.; Woodyer, R. D.; Zhao, H. Efficient Regeneration of NADPH Using an Engineered Phosphite Dehydrogenase. *Biotechnol. Bioeng.* 2007, *96* (1), 18–26.
- (74) Seelbach, K.; Riebel, B.; Hummel, W.; Kula, M.-R.; Tishkov, V. I.; Egorov, A. M.; Wandrey, C.; Kragl, U. A Novel, Efficient Regenerating Method of NADPH Using a New Formate Dehydrogenase. *Tetrahedron Lett.* 1996, *37* (9), 1377–1380.
- (75) Davis, B. G.; Celik, A.; Davies, G. J.; Ruane, K. M. Novel Enzyme. WO2010041055A2, April 15, 2010.
- (76) Mertens, N.; Remaut, E.; Fiers, W. Tight Transcriptional Control Mechanism Ensures Stable High-Level Expression from T7 Promoter-Based Expression Plasmids. *Bio/Technology* 1995, *13* (2), 175–179.
- (77) Ochs, R. S. Understanding Enzyme Inhibition. *J. Chem. Educ.* 2000, *77* (11), 1453.
- (78) Bachan Upadhyay, L. S.; Verma, N. Enzyme Inhibition Based Biosensors: A Review. *Anal. Lett.* 2013, *46* (2), 225–241.

- (79) Burlingham, B. T.; Widlanski, T. S. An Intuitive Look at the Relationship of K_i and IC_{50} : A More General Use for the Dixon Plot. *J. Chem. Educ.* 2003, 80 (2), 214.
- (80) Wu, J. T.; Wu, L. H.; Knight, J. A. Stability of NADPH: Effect of Various Factors on the Kinetics of Degradation. *Clin. Chem.* 1986, 32 (2), 314–319.
- (81) Li, Z.; Rinas, U. Recombinant Protein Production-Associated Metabolic Burden Reflects Anabolic Constraints and Reveals Similarities to a Carbon Overfeeding Response. *Biotechnol. Bioeng.* 2021, 118 (1), 94–105.
- (82) Studier, F. W.; Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *J. Mol. Biol.* 1986, 189 (1), 113–130.
- (83) Hanko, E. K. R.; Minton, N. P.; Malys, N. Characterisation of a 3-Hydroxypropionic Acid-Inducible System from *Pseudomonas putida* for Orthogonal Gene Expression Control in *Escherichia coli* and *Cupriavidus necator*. *Sci. Rep.* 2017, 7 (1), 1724.
- (84) Di Lorenzo, M.; Hidalgo, A.; Molina, R.; Hermoso, J. A.; Pirozzi, D.; Bornscheuer, U. T. Enhancement of the Stability of a Prolipase from *Rhizopus oryzae* toward Aldehydes by Saturation Mutagenesis. *Appl. Environ. Microbiol.* 2007, 73 (22), 7291–7299.
- (85) Franken, B.; Eggert, T.; Jaeger, K. E.; Pohl, M. Mechanism of Acetaldehyde-Induced Deactivation of Microbial Lipases. *BMC Biochem.* 2011, 12 (1), 10.
- (86) Weber, H. K.; Stecher, H.; Faber, K. Sensitivity of Microbial Lipases to Acetaldehyde Formed by Acyl-Transfer Reactions from Vinyl Esters. *Biotechnol. Lett.* 1995, 17 (8), 803–808.
- (87) Tingirikari, J. M. R.; Ahmed, S.; Yata, V. K. Chapter 6 - 3-Hydroxy-Propionic Acid. In *Platform Chemical Biorefinery*; Kaur Brar, S., Jyoti Sarma, S., Pakshirajan, K., Eds.; Elsevier: Amsterdam, 2016; pp 101–117.
- (88) Kildegaard, K. R.; Wang, Z.; Chen, Y.; Nielsen, J.; Borodina, I. Production of 3-Hydroxypropionic Acid from Glucose and Xylose by Metabolically Engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* 2015, 2, 132–136.
- (89) Warnecke, T. e.; Lynch, M. d.; Lipscomb, M. I.; Gill, R. t. Identification of a 21 Amino Acid Peptide Conferring 3-Hydroxypropionic Acid Stress-Tolerance to *Escherichia coli*. *Biotechnol. Bioeng.* 2012, 109 (5), 1347–1352.
- (90) Kildegaard, K. R.; Hallström, B. M.; Blicher, T. H.; Sonnenschein, N.; Jensen, N. B.; Sherstyk, S.; Harrison, S. J.; Maury, J.; Herrgård, M. J.; Juncker, A. S.; Forster, J.; Nielsen, J.; Borodina, I. Evolution Reveals a Glutathione-Dependent Mechanism of 3-Hydroxypropionic Acid Tolerance. *Metab. Eng.* 2014, 26, 57–66.
- (91) Nguyen-Vo, T. P.; Liang, Y.; Sankaranarayanan, M.; Seol, E.; Chun, A. Y.; Ashok, S.; Chauhan, A. S.; Kim, J. R.; Park, S. Development of 3-Hydroxypropionic-Acid-Tolerant Strain of *Escherichia coli* W and Role of Minor Global Regulator YieP. *Metab. Eng.* 2019, 53, 48–58.

- (92) Sankaranarayanan, M.; Ashok, S.; Park, S. Production of 3-Hydroxypropionic Acid from Glycerol by Acid Tolerant Escherichia Coli. *J. Ind. Microbiol. Biotechnol.* 2014, 41 (7), 1039–1050.
- (93) Abraham, T. W.; Allen, E.; Bonhert, E. C.; Frank, C. L.; Hahn, J. J.; Tsobanakis, P. Recovery of 3-Hydroxypropionic Acid. WO2014144367A1, September 18, 2014.

CHAPTER THREE: 3-HP BIOSENSOR FOR EVOLUTION OF Cg10062

3.1. Introduction

A novel route to *in vitro* biosynthesis of 3-HP from ACA was discussed in detail in chapter two. In spite of the high efficiency and quantitative formation of the desired building block chemical from ACA, this biocatalytic process may be optimized further to improve the formation of 3-HP. In particular, enzyme evolution of Cg10062(E114N) to obtain non-decarboxylating variant with improved overall catalytic efficiency with ACA remains a top priority. The E114N variant was discovered by rational mutagenesis of the active site of Cg10062. Site-directed mutagenesis is an indispensable tool in the field of molecular biology and enzyme evolution, and remains a useful technique for elucidation of enzyme mechanisms as well as optimization of enzyme activity.¹⁻³ A rational design approach can be used to alter substrate specificity or product profiles,⁴ as was the case for Cg10062. Although site-directed mutagenesis is a useful tool used to generate individual variants, the process can be a laborious and time-consuming.⁵⁻⁷ Generating a single variant of Cg10062 requires the initial steps of site-directed mutagenesis on a parent plasmid, transformation, isolating plasmid DNA, diagnostic restriction digests and sanger sequencing. Subsequent characterization of the variants requires labor-intensive enzyme purification and characterization via various analytical techniques such as SDS-PAGE, UV-based assays, and NMR studies. Characterization of individual Cg10062 variants is particularly laborious since a coupled enzyme assay with MSAD and ADH (see chapter four) is required to distinguish between decarboxylating and non-decarboxylating variants. Furthermore, the sequence space that can be targeted via site-directed mutagenesis is limited, while generation of mutant libraries via random mutagenesis (which requires high-throughput screening methods) allows more complete exploration of the sequence space.⁸ While rational mutagenesis of key catalytic residues in Cg10062 proved useful in improving the product distribution (MSA versus acetaldehyde) from ACA hydration, the biocatalytic route would benefit from further optimization of Cg10062(E114N) which can be achieved more easily using high-throughput screening methodology.

Among high-throughput screening methods developed over the past few decades, the use of genetically-encoded biosensors driven by small-molecule inducers has proven to be a successful biomolecular tool.⁹ Biosensors are typically designed to provide an output signal, such as a fluorescent readout, that is proportional to the intracellular concentration of the targeted metabolite being monitored. However, while diverse regulatory pathways exist in nature, an inducible system pertaining to ACA or the preferred Cg10062 hydration product (MSA) is yet to be discovered. In such cases, it is not uncommon to rely on the use of an indirect biosensor, where the undetectable compound is converted to a metabolite that can be detected by an existing inducible system.^{10,11} For example, Rogers and Church describe the development of an acrylate based-biosensor for the detection of 3-HP in *E. coli*.¹¹ Inside the cells, 3-HP is converted to acrylate in a series of enzyme-catalyzed reactions. Thereafter, acrylate activates AcuR, a transcriptional regulator native to *Rhodobacter sphaeroides*, to induce a fluorescent output signal.^{11,12} This ability to monitor a target metabolite in real-time via a fluorescent signal allows the flexibility of screening large mutant libraries (1×10^9) that is not possible using traditional methods.¹¹ To achieve these rate of screening, additional tools such as fluorescence-activated cell sorting (FACS) must be used in combination with the biosensor.¹³

Inspired by the use of indirect biosensors to monitor production of undetectable compounds, the use of genetically-encoded biosensors was considered for enabling efficient screening of Cg10062 variants. IPTG-induced expression of Cg10062(E114N) and YdfG in *E. coli* has previously resulted in the formation of trace amounts of 3-HP (see chapter two). While an inducible system for ACA or MSA detection has not been described, co-expression Cg10062(E114N) with YdfG should enable conversion of ACA to 3-HP, allowing the use of 3-HP-responsive biosensor as an indirect measurement of Cg10062 hydratase activity. For example, a biosensor strain expressing a novel variant of Cg10062, with improved catalytic efficiency relative to Cg10062(E114N) would result in faster conversion of ACA to 3-HP. In screening for novel variants with improved kinetics, its crucial that they maintain hydratase only activity with ACA. A

hydratase-only variant with higher catalytic efficiency would lead to a stronger signal relative to a decarboxylating variant that results in lower 3-HP formation. One of the most recent studies highlighted by Park and coworkers, inspired by a native 3-HP degradation pathway in *Pseudomonas denitrificans* was studied in detail for utilization in Cg10062 evolution.^{14–18}

3.1.1. 3-HP degradation in *Pseudomonas denitrificans*

In 2013, Zhou et. al reported the first synthetic pathway to 3-HP in *Pseudomonas denitrificans* using glycerol.¹⁹ The authors report an unexpectedly low yield of 3-HP and thorough analysis indicated that 3-HP was being assimilated and degraded by *P. denitrificans*, thereby preventing the expected titers.^{14,15} Complete genome sequencing and metabolite analysis indicated that 3-HP metabolism occurred via two key enzymes in *P. denitrificans*, a 3-hydroxyisobutyrate dehydrogenase (HpdH4) and methylmalonate semialdehyde dehydrogenase (MmsA) (Figure 3.1).^{20,21} Genome analysis indicated that a gene a LysR-type transcriptional regulators, *mmsR*, was transcribed divergently upstream of *mmsA*.¹⁷

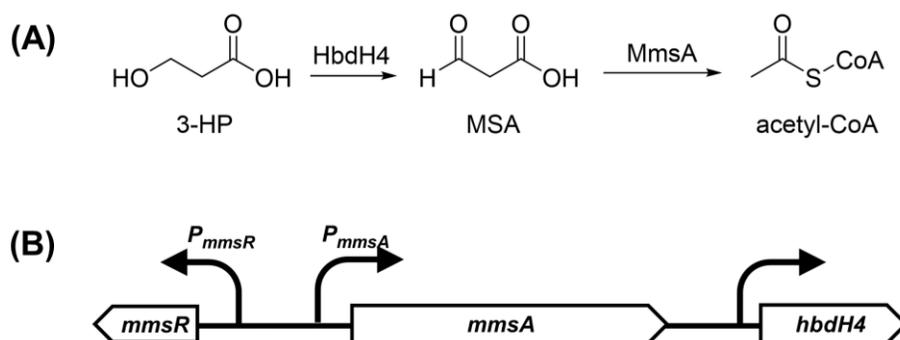


Figure 3.1. 3-HP degradation pathway in *P. denitrificans*. (A) Key enzymes involved in 3-HP degradation. (B) Transcriptional regulator *mmsR* is divergently transcribed upstream of *mmsA* in *P. denitrificans*. (HpdH4, 3-hydroxyisobutyrate dehydrogenase IV; MmsA, methylmalonate semialdehyde dehydrogenase).

The LysR-type transcriptional regulator (LTTR) family are DNA-binding proteins first discovered by Henikoff et. al in 1988 and is one of the most abundant protein families (>40,000) found in bacteria.^{22,23} The LTTR family are implicated in diverse physiological functions and the number of LTTRs found in different organisms can vary significantly.^{22,23} Sequence alignment of

LTTRs from various organisms have shown that they are comprised of a DNA-binding domain (helix-turn-helix motif) towards the N-terminus and a cofactor-binding domain near the C-terminus.^{16,22,24} Interaction of LTTRs with their target sequence results in conformational changes of the DNA and subsequent recruitment of RNA polymerase for transcription.^{16,24} LTTRs are divergently transcribed from its promoter in the vicinity of the gene which it is regulating, and most LTTRs are known repress their own transcription, as observed in the 3-HP catabolic pathway in *P. denitrificans*.^{17,18,25}

3.1.2. Genetically-encoded 3-HP biosensor in *Pseudomonas denitrificans*

Nguyen et. al describe the development of the 3-HP biosensor in *P. denitrificans* using various elements of the 3-HP catabolic pathway identified in the strain.¹⁷ The biosensor was developed in *P. denitrificans* using a single plasmid (Figure 3.2) which contained the *mmsR* gene downstream of P_{zwf} promoter (native to *P. denitrificans*). On the same plasmid, the *gfp* gene encoding the green fluorescent protein was cloned downstream of P_{mmsA} .¹⁷ The MmsR transcriptional regulator binds 3-HP and the resulting 3-HP–MmsR complex binds P_{mmsA} to enable transcription of the *rfp* reporter gene. The basal RFP signal from the biosensor strain was minimal relative to the fluorescent signal observed with 3-HP induction, indicating that in the absence of 3-HP, apo-MmsR represses transcription of *rfp*.^{16,17} Characterization of the biosensor with varying 3-HP concentrations displayed a dose-dependent increase in RFP signal intensity relative to the 3-HP concentration with a dynamic range of 0.1–10 mM.^{17,18} Additional studies in which random mutations were introduced in the *mmsA* promoter region, to increase the sensitivity and dynamic range of the biosensor were carried out. However, mutagenesis of the operator sites of P_{mmsA} led to significant reduction or lack of induction by 3-HP and increased background fluorescence in the absence of 3-HP.¹⁷ These findings suggested that the mutations reduced the overall binding affinity of MmsR towards P_{mmsA} , weakening its ability to repress the promoter effectively.¹⁷ Therefore, the authors recommend minimal to no alteration of the native P_{mmsA} sequence required for effective induction and application of the biosensor. The genetically-encoded MmsR-based

biosensor was introduced into *P. putida* and *E. coli* to demonstrate the versatility of the biosensor, although low levels of induction were observed relative to *P. denitrificans*.¹⁷ The biosensor was also applied as a screening tool for engineering an aldehyde dehydrogenase (ALDH) in *P. denitrificans*.¹⁸

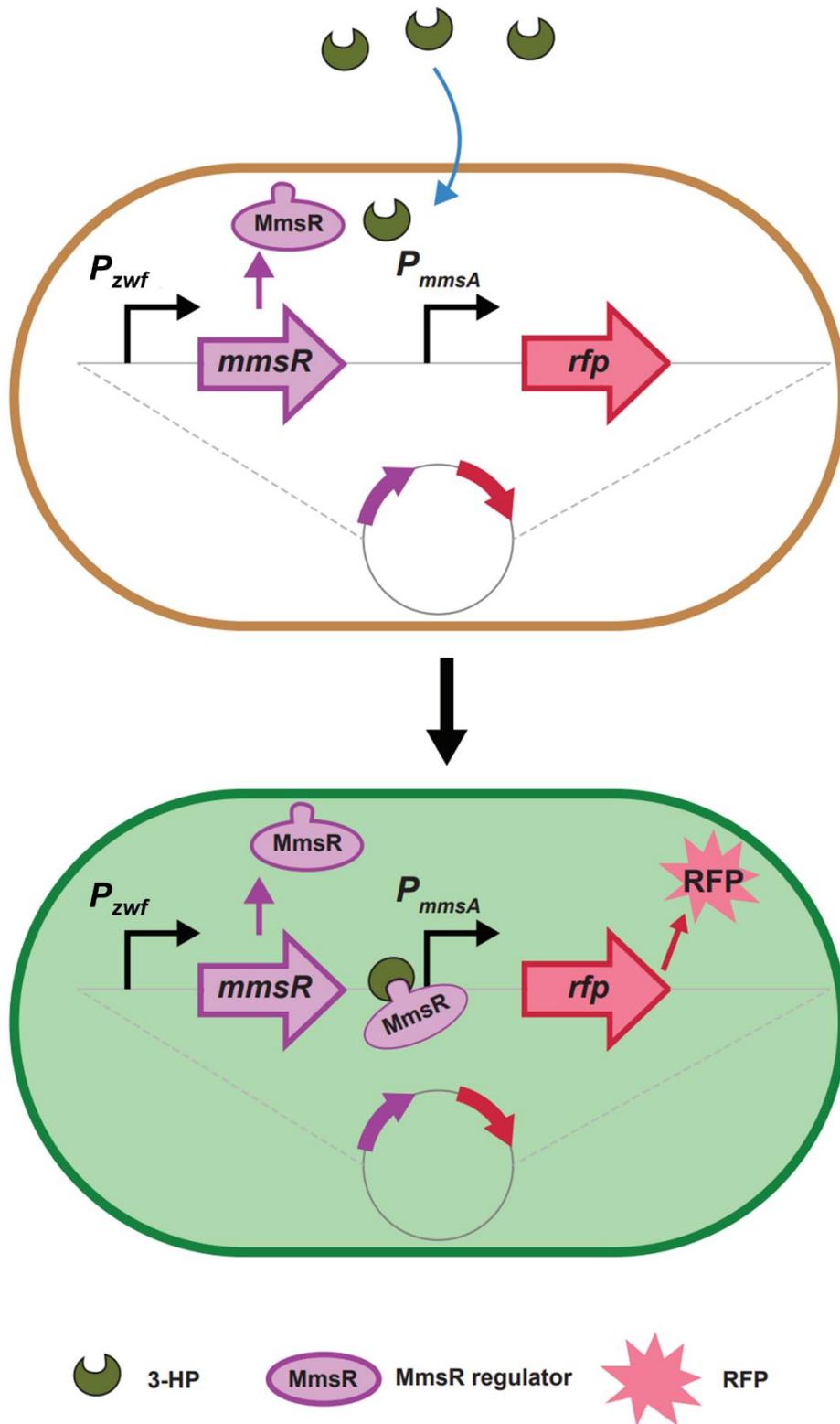


Figure 3.2. A genetically-encoded 3-HP biosensor in *P. denitrificans*. Genes *mmsR* and *rfp* are encoded on a plasmid downstream of *zwf* and *mmsA* promoters, respectively. Transcription of *rfp* is repressed in the absence of 3-HP. In the presence of 3-HP, the MmsR-3-HP complex binds the *mmsA* promoter to enable transcription of *rfp*.

The level of tunability, dynamic range and sensitivity desired from a biosensor is typically dependent on its application.⁹ For example, utilizing the 3-HP biosensor developed by Nguyen et. al^{17,18} in a 3-HP production strain with the target of improving 3-HP titers will require a biosensor that has low sensitivity, allowing detection of significantly high concentrations of 3-HP without reaching saturation. For the evolution of Cg10062(E114N), a relatively higher sensitivity where a higher fluorescence response is triggered at lower 3-HP concentrations is desired since the biosensor will only be used as a high-throughput screening tool of Cg10062 mutant libraries. Unlike in an application focusing on improving 3-HP titers, the evolution of Cg10062(E114N) can be carried out at lower 3-HP concentrations with exogenous addition of ACA to the media.

3.2. Development of a 3-HP biosensor for evolution of Cg10062 in *E. coli* MG1655

As discussed previously (chapter one), the E114N variant of Cg10062 was discovered through rational mutagenesis of the known active site residues of the tautomerase enzyme. However, engineering of Cg10062(E114N) to further improve its catalytic efficiency for hydration of ACA is required. In this study, the 3-HP biosensor described above will be applied for the evolution of Cg10062(E114N) in *E. coli* MG1655. Modifications made to the previously described biosensor to enable fluorescence output as a function of Cg10062-catalyzed hydration of ACA in *E. coli* MG1655 is described in this section.

The development of the biosensor in *E. coli* MG1655 was carried out using two plasmids. Since previous studies demonstrated that RFP proved to be better than GFP in terms of stability,¹⁸ the reporter plasmid pAS5.003 (Figure 3.3) was constructed using the ampicillin-resistant BglBrick vector pBbA1a-RFP that carries the *rfp* gene downstream of the IPTG-inducible *trc* promoter.^{26,27} The *trc* promoter was replaced by the *mmsA* promoter (*mmsR-mmsA* intergenic region in *P. denitrificans*) and the *mmsR* gene from *P. denitrificans* was cloned in to the same plasmid, upstream of *rfp*. The gene encoding MmsR was transcribed using a high-strength constitutive promoter, P_{J23119}, from the Anderson promoter library.²⁸ Strong expression of *mmsR* results in high concentrations of MmsR, enabling adequate repression of P_{*mmsA*} in the absence of 3-HP. In

the presence of 3-HP, the transcriptional regulator MmsR binds 3-HP, and this complex binds to P_{mmsA} to enable transcription of the *rfp* gene from pAS5.003.¹⁷

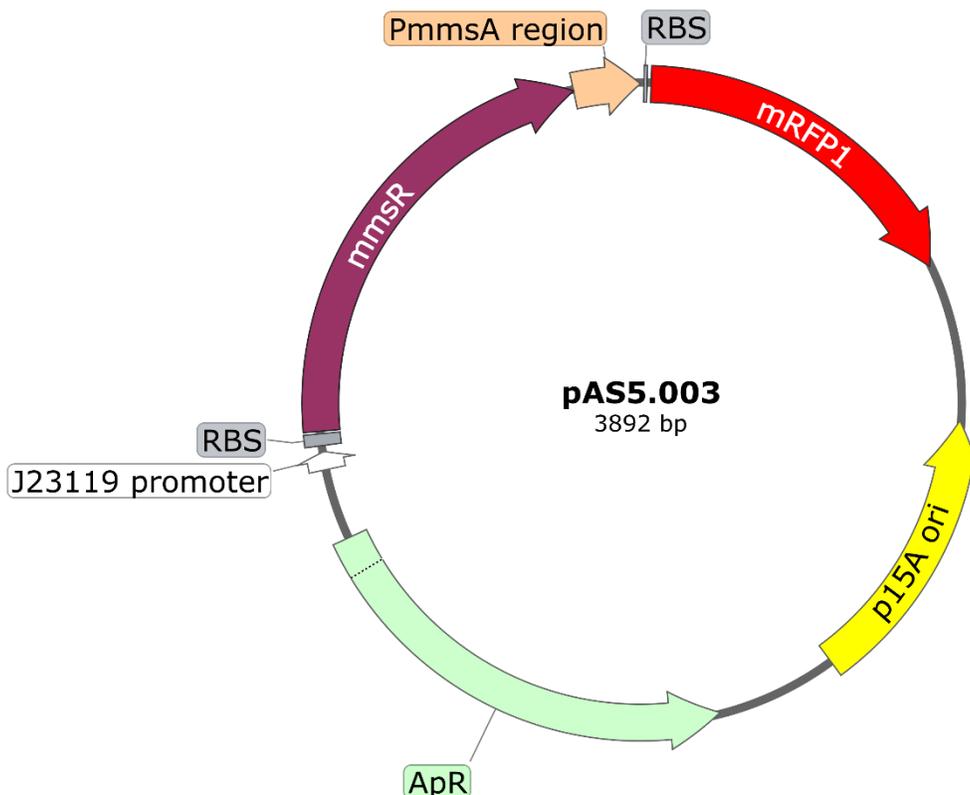


Figure 3.3. Plasmid map of pAS5.003 used for expression of MmsR and RFP.

The second component of the biosensor are the two enzymes required for conversion of ACA to 3-HP. pET-28a(+) carrying a kanamycin resistance gene was selected for the cloning of *cg10062(E114N)* and *ydfG*, where each gene was placed downstream of the medium-high strength P_{J23102} , to construct pAS5.008 (Figure 3.4).²⁸ Electrocompetent *E. coli* MG1655 was transformed with pAS5.003 resulting in strain designated *E. coli* 3HPC. Electrocompetent *E. coli* 3HPC was transformed with pAS5.008 resulting in strain *E. coli* 3HPC-E114N carrying plasmids pAS5.003 and pAS5.008 (MG1655/pAS5.003/pAS5.008). For comparison of the fluorescent output signals caused by varying Cg10062 activity, a plasmid encoding wild type *cg10062* and

ydfG was constructed using the methods described above for pAS5.008. This resulting plasmid pAS5.005 (Figure 3.5) was transformed into 3HPC to obtain strain 3HPC-WT (MG1655/pAS5.003/pAS5.005).

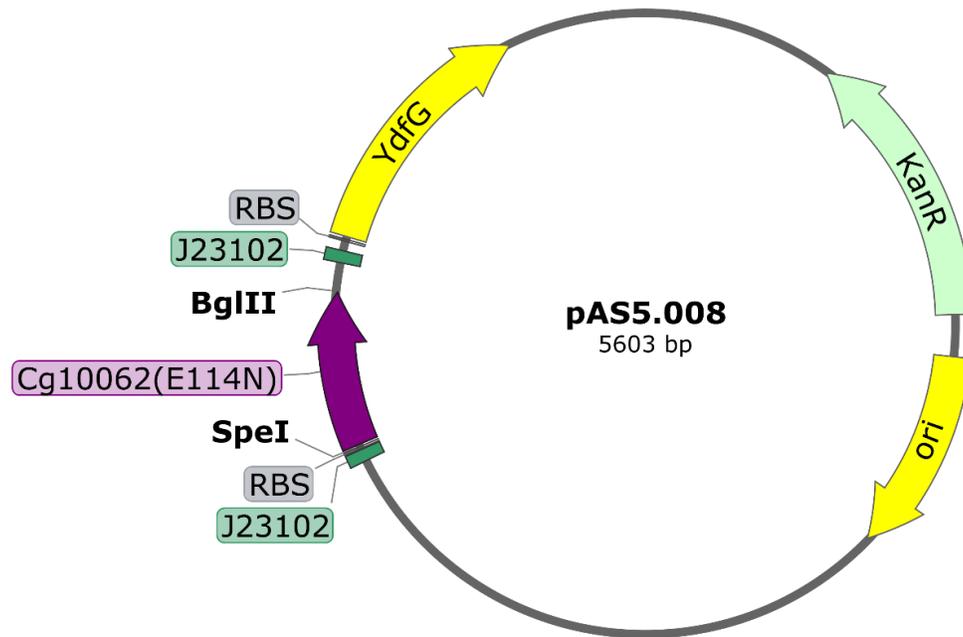


Figure 3.4. Plasmid map of pAS5.008 used for expression of *Cg10062(E114N)* and *YdfG*.

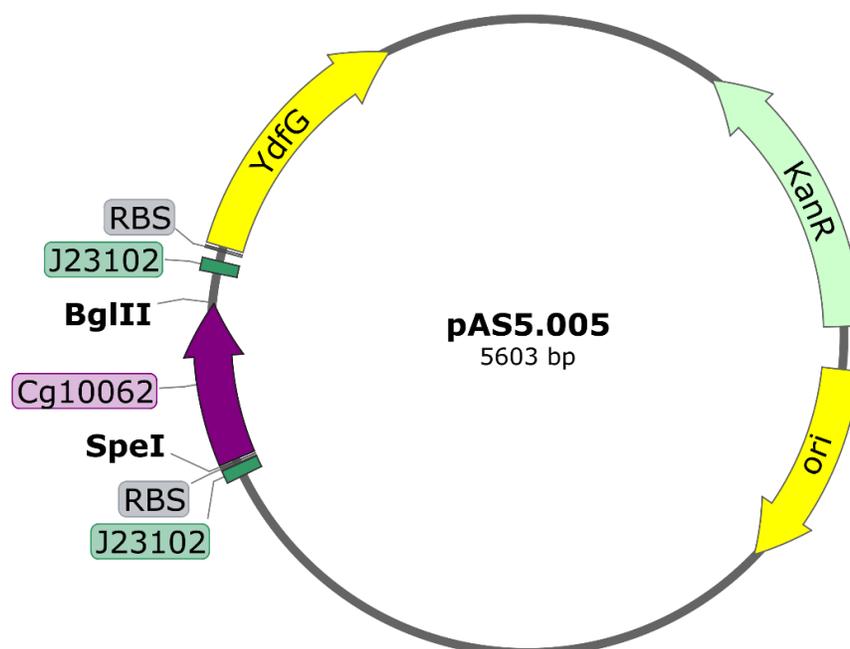


Figure 3.5. Plasmid map of pAS5.005 used for expression of Cg10062 and YdfG.

The complete 3-HP biosensor comprising of a two plasmid system is shown in Figure 3.6. ACA transported into the intracellular matrix of *E. coli* MG1655 was expected to be hydrated by Cg10062(E114N) or other novel ACA-hydrating variant produced via random mutagenesis, to produce MSA, which would subsequently be reduced to 3-HP by YdfG. The 3-HP binds the LysR-type transcriptional regulator MmsR, which in turn recruits RNA polymerase for transcription of *rfp*. Since *apo*-MmsR has been previously identified as a negative regulator of P_{mmsA} , in the absence of 3-HP, minimal expression of RFP was expected.¹⁷ Since the 3-HP biosensor will be employed as an indirect biosensor to screen for novel non-decarboxylating variants of Cg10062, efficient ACA uptake into cells and subsequent quantitative conversion of ACA to 3-HP is crucial in order to observe a dose- and time-dependent correlation between intracellular levels of 3-HP and RFP intensity. Assuming efficient ACA transport into the cell, a decarboxylating variant of Cg10062 (for example wild type Cg10062) favors acetaldehyde production from ACA. This translates to low levels of 3-HP formation and should ideally be reflected in low RFP intensity. On the other hand, a kinetically-improved non-decarboxylating hydratase would be expected to

display a stronger fluorescent signal assuming efficient ACA transport into the cells and rapid conversion of MSA to 3-HP by YdfG.

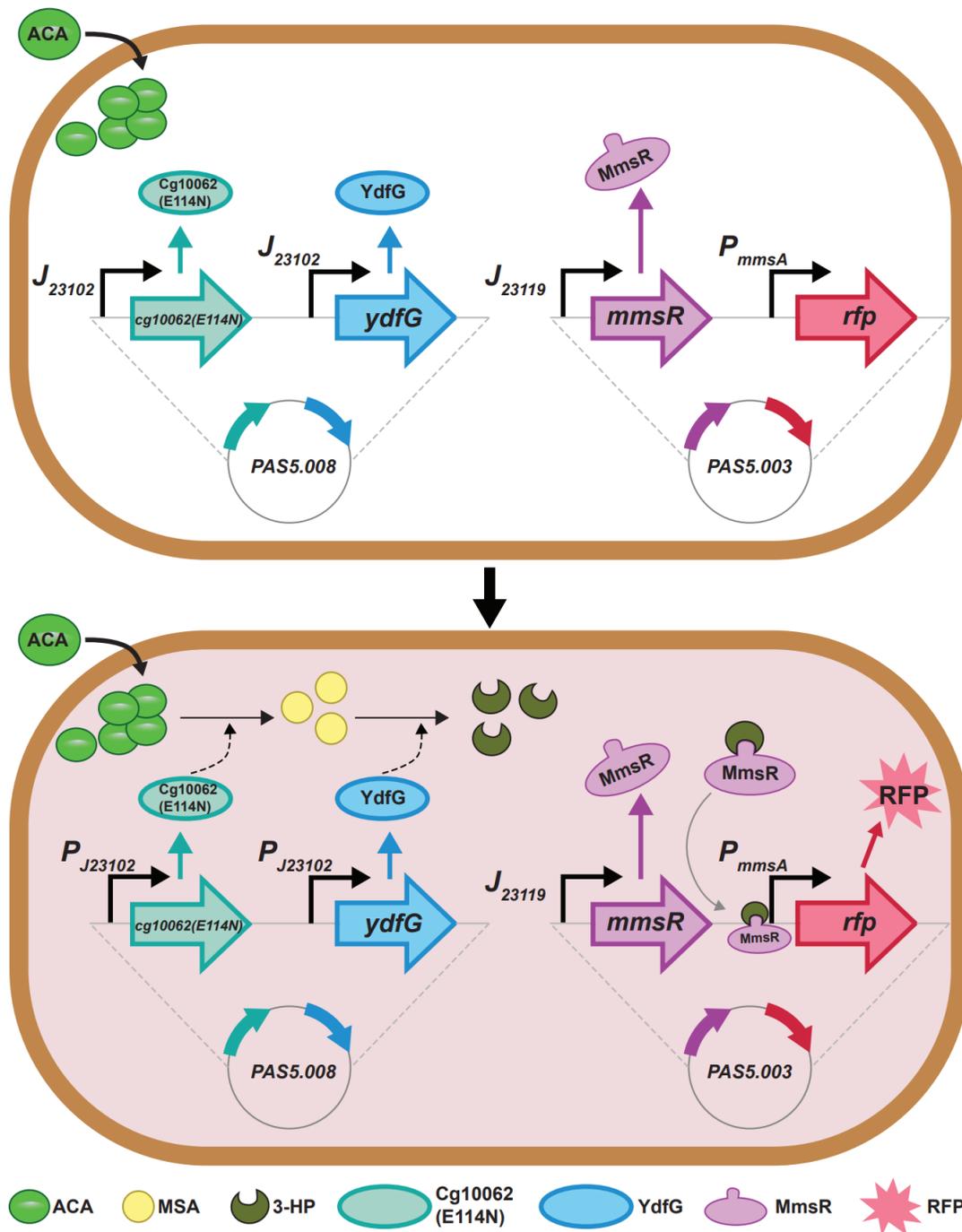


Figure 3.6. 3-HP biosensor for evolution of Cg10062(E114N) in *E. coli* MG1655. ACA entering the cell is converted to 3-HP via MSA by Cg10062(E114N) and YdfG. The 3-HP-MmsR complex binds the *mmsA* promoter to enable transcription of *rfp* in the presence of 3-HP.

3.3. Characterization of 3-HP biosensor in *E. coli* MG1655

Prior to utilizing the biosensor for evolution of Cg10062(E114N) by indirect measurement of 3-HP, the biosensor must be characterized and optimized for a time- and dose-dependent response of the fluorescent signal. For the initial characterization of the 3-HP biosensor in *E. coli* MG1655, the OD₆₀₀ and fluorescence intensity were measured over time upon exposure of the cells to a range of 3-HP concentrations. Assuming that 3-HP is efficiently transported into the cells, it was expected that an increase in 3-HP concentrations would result in a corresponding increase in the fluorescence output signal. Observed fluorescence was normalized to the OD₆₀₀ and is reported as the specific fluorescence (average RFU/OD₆₀₀). Overnight cultures of 3HPC, 3HPC-WT and 3HPC-E114N grown in M9 media containing glucose were used to prepare fresh cultures with an initial OD₆₀₀ of ~0.1. Aliquots of these cultures (180 µL) were transferred onto a black clear-bottomed 96-well plate that would allow simultaneous measurements of the OD₆₀₀ and fluorescence (ex/em = 535/585 nm).^{18,29} The plate was sealed with a Genetix BreatheSeal film and allowed to grow to an OD₆₀₀ of 0.5–0.7. However, after >10 h, a significant increase in OD₆₀₀ was not observed and the use of the porous film caused evaporation of the media at 37 °C.

In a different approach, the cultures of 3HPC, 3HPC-WT and 3HPC-E114N were allowed to grow to an OD₆₀₀ of 0.5 – 0.7 in culture tubes prior to transfer into a 96-well plate. This ensured that the initial OD₆₀₀ of the cultures were identical at the beginning of the induction experiment. The BreatheSeal film used prior was replaced with a MicroAmp™ Optical Adhesive film which was successful in allowing aeration of the cultures without any loss of sample volume. Triplicate samples (180 µL) of each strain were induced with a range of 3-HP concentrations: 0, 0.5, 1, 2.5, 5, 10, 25, 50, 75 and 100 mM and OD₆₀₀ and fluorescence measurements were obtained at 0, 3, 6, and 12 h. Since the cells were directly supplemented with 3-HP instead of ACA, assuming a dose-dependent response to 3-HP, a similar trend in inducibility was expected across the three strains. The lack of a plasmid carrying Cg10062 and YdfG in the 3HPC strain was not expected to impact the specific fluorescence significantly since the 3-HP was expected to bind the MmsR

regulator without first requiring conversion by the two enzymes. However, although the cells continued to grow, an increase in fluorescence corresponding to the increasing 3-HP concentrations was not observed (data not shown). All RFU measurements obtained at every time point was lower than the background fluorescence observed in wells without any addition of 3-HP.

One observation made during the development of the biosensor was the pink appearance of *E. coli* colonies harboring the pBbA1a-RFP plasmid where RFP is regulated by P_{trc} , when grown on solid media. Transformants plated on LB/Ap or M9/glu/Ap plates appeared pink without the need for IPTG induction, indicating leaky expression of RFP due to poor regulation. In contrast, *E. coli* carrying the reporter plasmid pAS5.003 which encodes RFP downstream of the high-strength P_{J23119} appeared white on both LB/Ap and M9/glu plates, an indication of tight repression of the P_{mmsA} -guided *rfp* gene. *E. coli* MG1655/pBbA1a-RFP cultured in minimal media as described above and induced with varying concentrations of IPTG (0, 0.01, 0.1 and 1 mM) was carried out to observe the response of the RFP signal to increasing concentrations of the inducer (Figure 3.7). The specific background fluorescence of RFP immediately after IPTG induction remained the same across all samples ($\sim 2 \times 10^6$). Over the course of experiment, a steady increase in the RFP signal proportional to the IPTG concentration was observed as a result of IPTG-induced overexpression of the reporter protein. The leaky expression of RFP commonly observed with IPTG inducible promoters was evident with a >2-fold increase in the signal after 12 h in uninduced cells.³⁰ Additionally, while a notable increase in RFP signal was observed between 0.01 and 0.1 mM, across all timepoints, the RFP signal remained more or less constant between the 0.1 mM and 1 mM IPTG samples indicating that 0.1 mM IPTG is adequate for full induction.

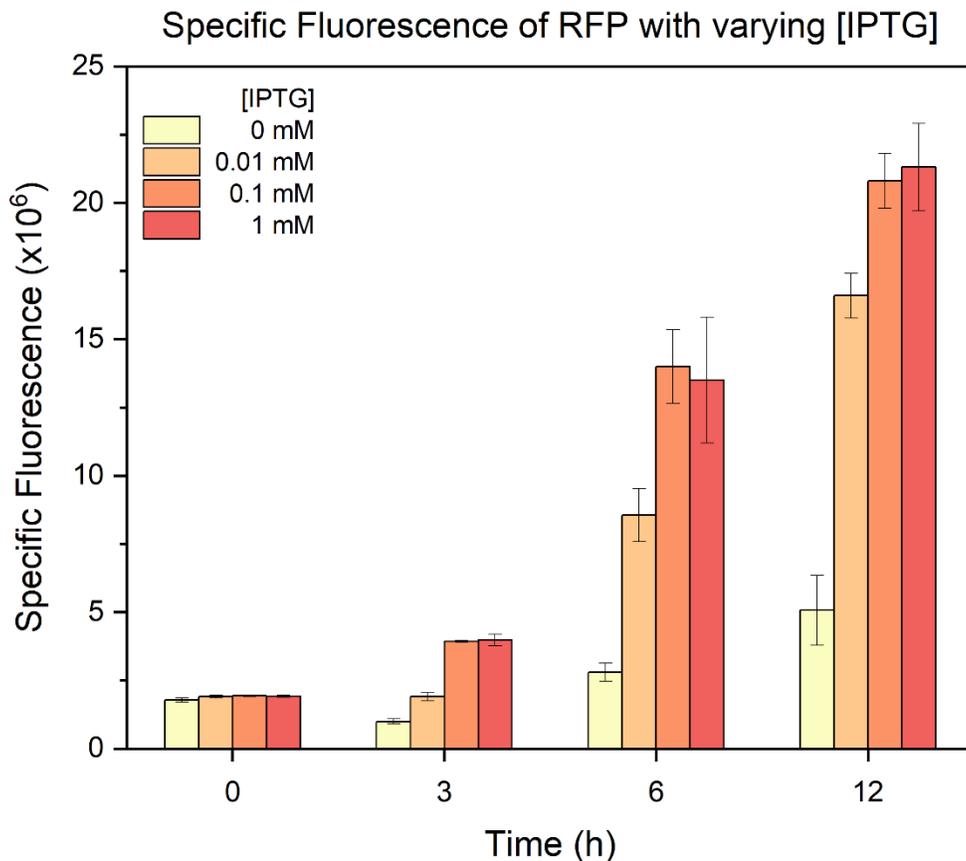


Figure 3.7. RFP fluorescent signal in *E. coli* MG1655/pBbA1a-RFP. The specific fluorescence of RFP was measured using a range of IPTG concentrations over a 12 h time course.

Based on these observations, it was hypothesized that the use of a high-strength promoter such as P_{J23119} enabled high levels of the MmsR inside the cell thus requiring higher concentrations of the inducer molecule in order to stimulate transcription of the *rfp* gene. The induction experiment was repeated with concentrations of 0, 75, 100, 125, 150, 175, 200, 250 and 300 mM 3-HP. However, the expected increase in RFP intensity (assuming transport of 3-HP into the cells) was not observed even with higher 3-HP concentrations. Similar observations were made when the cells were cultured in M9/glu media supplemented with the same concentrations of ACA (0–300 mM). On the other hand, the effects of 3-HP toxicity on cell growth were readily apparent. Starting with approximately the same OD_{600} , in all samples induced by 3-HP, a significant reduction in growth was observed across all samples containing 3-HP over a period of 12 h (Figure 3.8). As discussed in chapter two (section 2.8), previous studies have

reported inhibition of growth when 3-HP concentrations exceed 200 mM.^{31,32} Severe growth inhibition was also observed with the 3HPC-WT (Figure 3.9) and 3HPC-E114N strains, where the significant reduction of cell growth is more prominent at higher 3-HP concentrations. This pronounced growth inhibition can be attributed to a combination of 3-HP toxicity since growth inhibition was not observed in the absence of 3-HP. Interestingly, growth inhibition correlating to increasing ACA concentrations were not observed when the strains were cultured in the same concentrations of ACA suggesting that ACA itself may not be as toxic to cell growth relative to 3-HP (Figures 3.10–3.11).

The 3-HP-based growth inhibition of the biosensor strains leads to a significant hurdle in the determination of the dynamic range and sensitivity of the biosensor with respect to the inducer molecule. Testing the biosensor inducibility at even higher concentrations of 3-HP was not feasible, not only with respect to the observed growth inhibition. Therefore, optimization of the biosensor elements to observe a dose-dependent response at lower concentrations of 3-HP is necessary before it can be implemented for the evolution of Cg10062(E114N).

Growth Inhibition of 3HPC by 3-HP

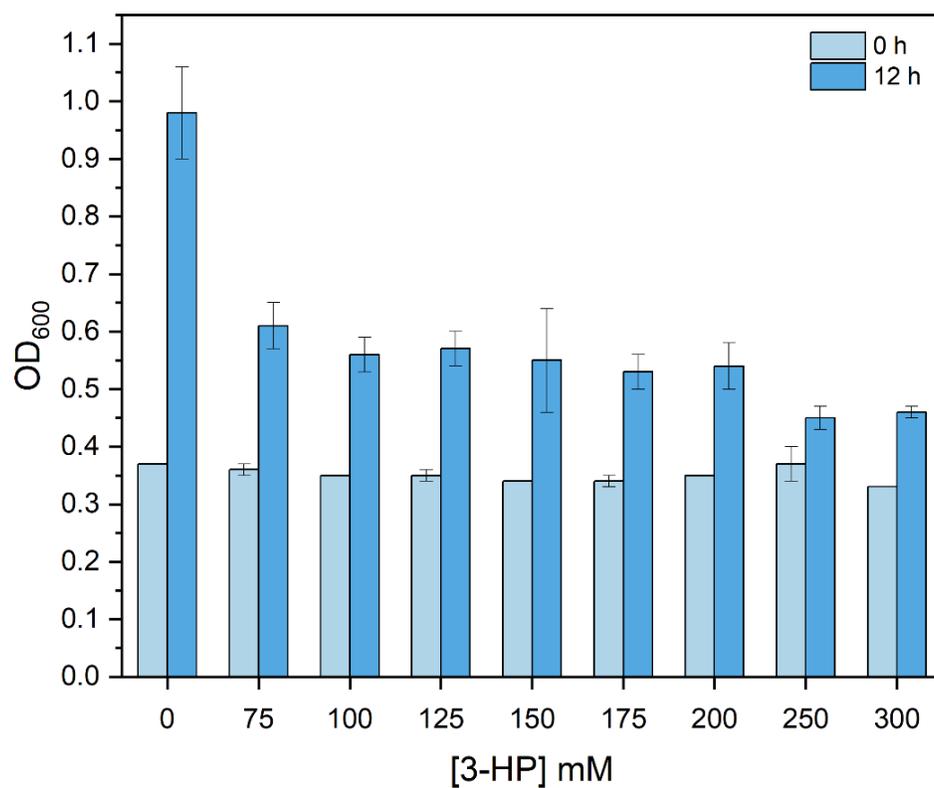


Figure 3.8. Growth inhibition of 3HPC with increasing 3-HP concentrations. The OD₆₀₀ of 3HPC strain induced with varying concentrations of 3-HP was measured at 0 h (light blue bars) and 12 h (dark blue bars).

Growth Inhibition of 3HPC-WT by 3-HP

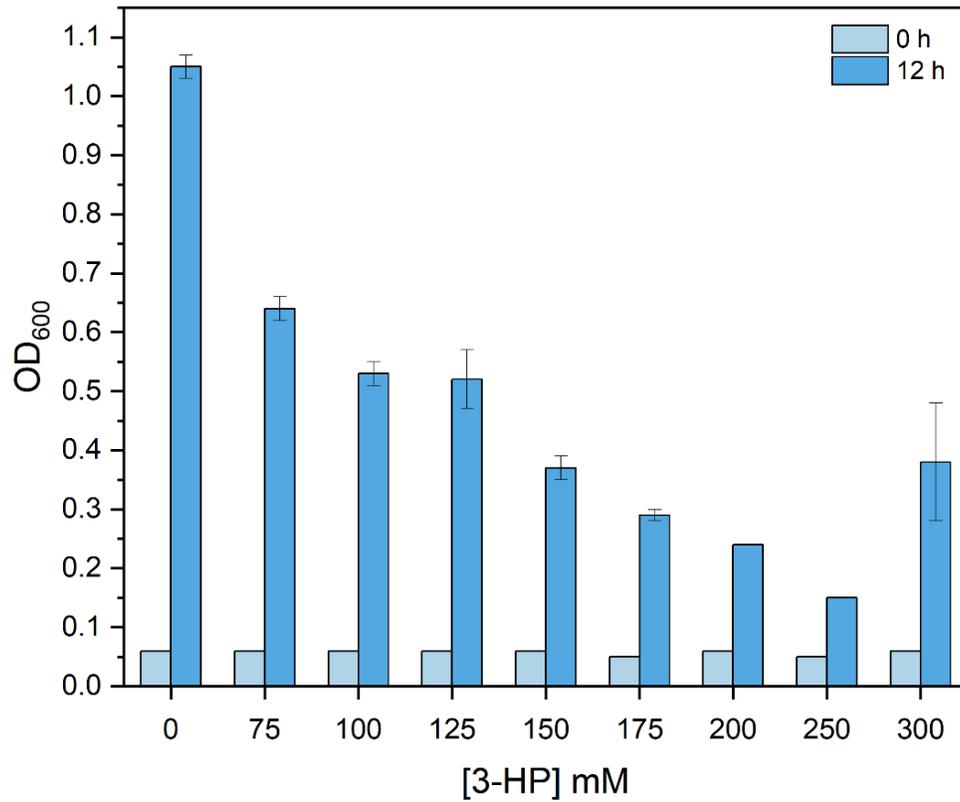


Figure 3.9. Growth inhibition of 3HPC-WT with increasing 3-HP concentrations. The OD₆₀₀ of 3HPC-WT strain induced with varying concentrations of 3-HP was measured at 0 h (light blue bars) and 12 h (dark blue bars).

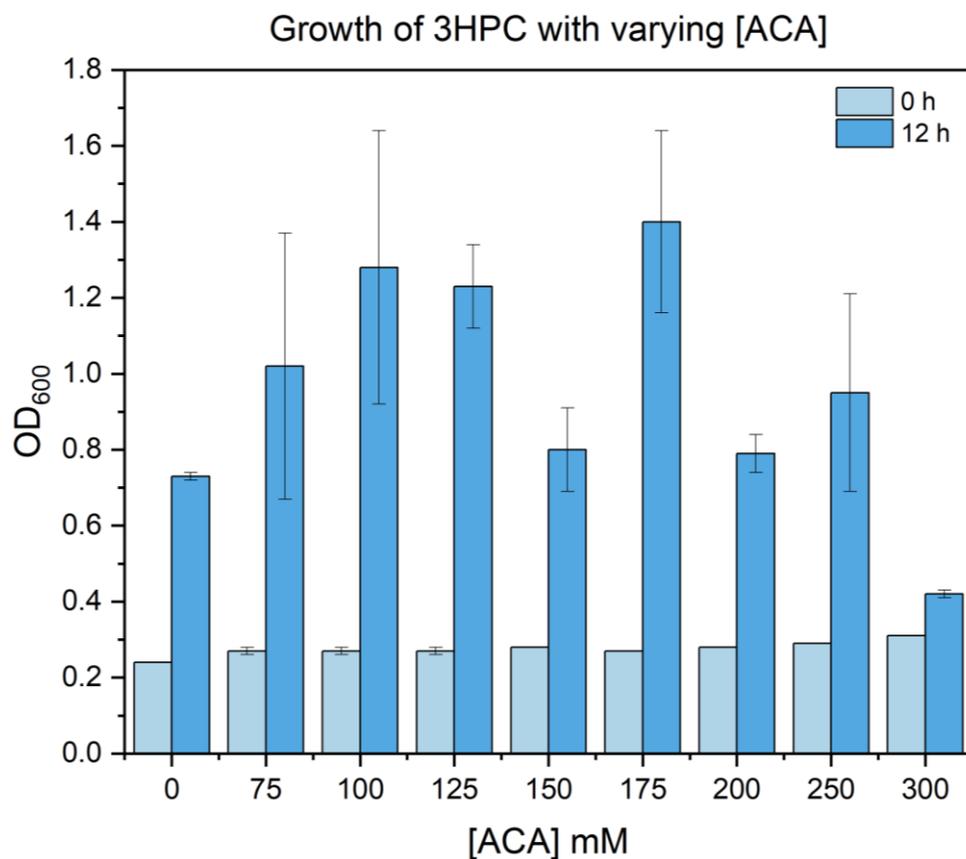


Figure 3.10. Growth of 3HPC with increasing ACA concentrations. The OD₆₀₀ of 3HPC strain induced with varying concentrations of ACA was measured at 0 h (light blue bars) and 12 h (dark blue bars).

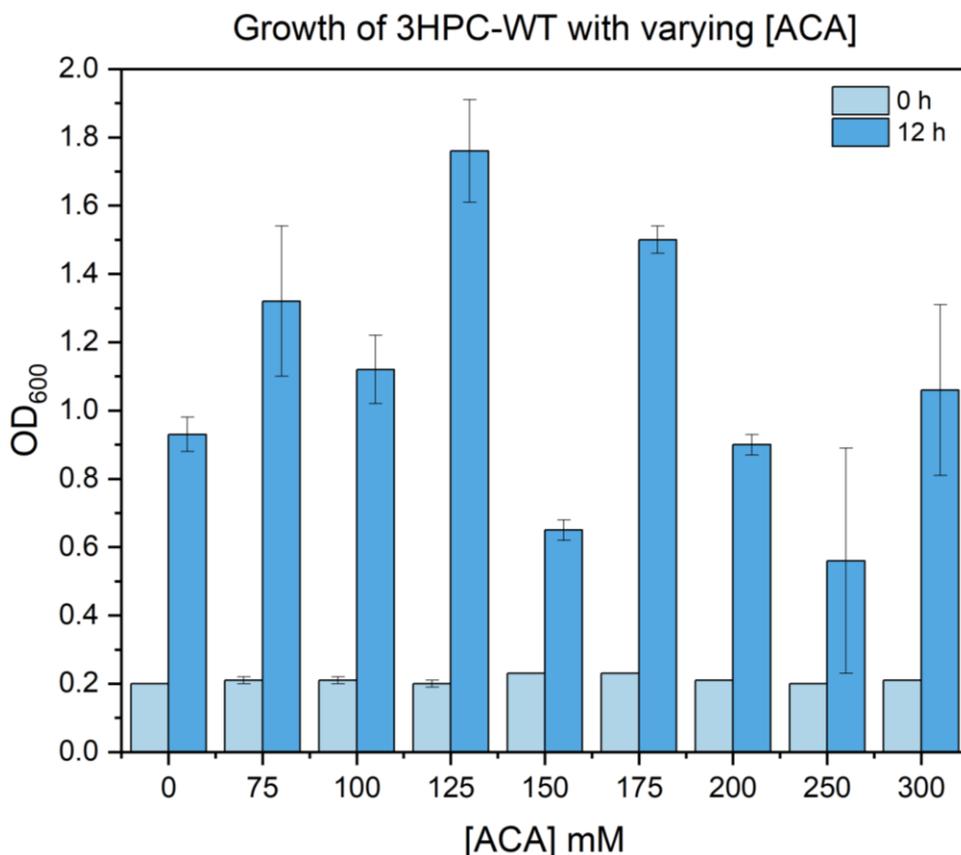


Figure 3.11. Growth of 3HPC-WT with increasing ACA concentrations. The OD₆₀₀ of 3HPC-WT strain induced with varying concentrations of ACA was measured at 0 h (light blue bars) and 12 h (dark blue bars).

3.4. ¹H NMR characterization of *in vivo* synthesis of 3-HP from ACA

The lack of inducibility of the 3HPC-E114N with a broad range of ACA concentrations prompted ¹H NMR analysis of the biosensor strain carrying the genes required for conversion of ACA to 3-HP. Prior results (see chapter two) indicated that the IPTG-induced overexpression of Cg10062(E114N) and YdfG allowed the formation of trace amounts of 3-HP *in vivo* when cells were provided with ACA. These results and the growth of *E. coli* expressing Cg10062 and variants on aldehyde indicator plates (chapter one) suggest that ACA is transported into the intracellular matrix of *E. coli*, albeit not very efficiently. The *in vivo* growth studies were carried out using plasmid pAS3.112 (Table 4.5), which enabled overexpression of Cg10062(E114N) and YdfG guided by a *trc* promoter. The formation of 3-HP detected by ¹H NMR in these was not quantifiable. In contrast, the biosensor plasmids carry the genes *cg10062(E114N)* and *ydfG*,

downstream of a high-strength constitutive promoter P_{J23102} .²⁶ The inclusion of the pAS5.003 reporter plasmid in the biosensor strain provides the added advantage of being able to detect an increase in fluorescence signal corresponding to 3-HP formation.

The three biosensor strains 3HPC, 3HPC-WT and 3HPC-E114N were cultured in minimal M9 media containing glucose for cell growth and ACA supplementation for 3-HP synthesis. An overnight culture grown in M9/glu was used to inoculate fresh M9/glu media (duplicate) to an initial OD_{600} of ~ 0.1 and one of the two duplicate cultures was supplemented with 50 mM ACA. An increase in fluorescence was not observed in any of the cultures containing ACA relative to the cultures without ACA. Culture supernatants were analyzed by 1H NMR 48 h after addition of ACA to identify any formation of 3-HP. Across all three strains, regardless of whether the media was supplemented with ACA or not, accumulation of acetate and ethanol were observed in the culture media (Figures 3.12-3.14). Peaks at δ 3.45 (q, $J = 7.1$ Hz, 2H) and δ 0.98 (t, $J = 7.1$ Hz, 3H), and δ 1.71 (s, 3H) correspond to ethanol and acetate, respectively. Peaks at δ 3.58 (t, $J = 6.6$ Hz, 2H) and δ 2.23 (t, $J = 6.6$ Hz, 2H), indicating *in vivo* production of 3-HP were observed in the supernatant from the 3HPC-E114N culture (Figure 3.10). However, unreacted ACA was also observed in the media caused either by inadequate transport of ACA into *E. coli* MG1655 or inefficient catalysis by the Cg10062(E114N)-YdfG pair or a combination of both factors. The high concentrations of ACA remaining after 48 h poses a significant limitation in attempts to utilize the 3-HP-based biosensor as an indirect method of screening *cg10062* mutant libraries. Analysis of 3HPC-WT (Figure 3.11) expectedly displayed only trace amounts of 3-HP present in the media. This would be expected since Cg10062(E114N) produces exclusively MSA while Cg10062 favors acetaldehyde production.^{33,34} ACA remaining in the media of both cultures indicates inefficient ACA transport into the cell and poses a significant limitation for application of the biosensor for Cg10062(E114N) evolution, which would only be effective if ACA transport is not a rate-limiting factor. The inefficient transformation of ACA to 3-HP (likely due to poor ACA transport) and the

lack of RFP inducibility in cells supplied with 3-HP prevent utilization of the biosensor for Cg10062(E114N) in its current form.

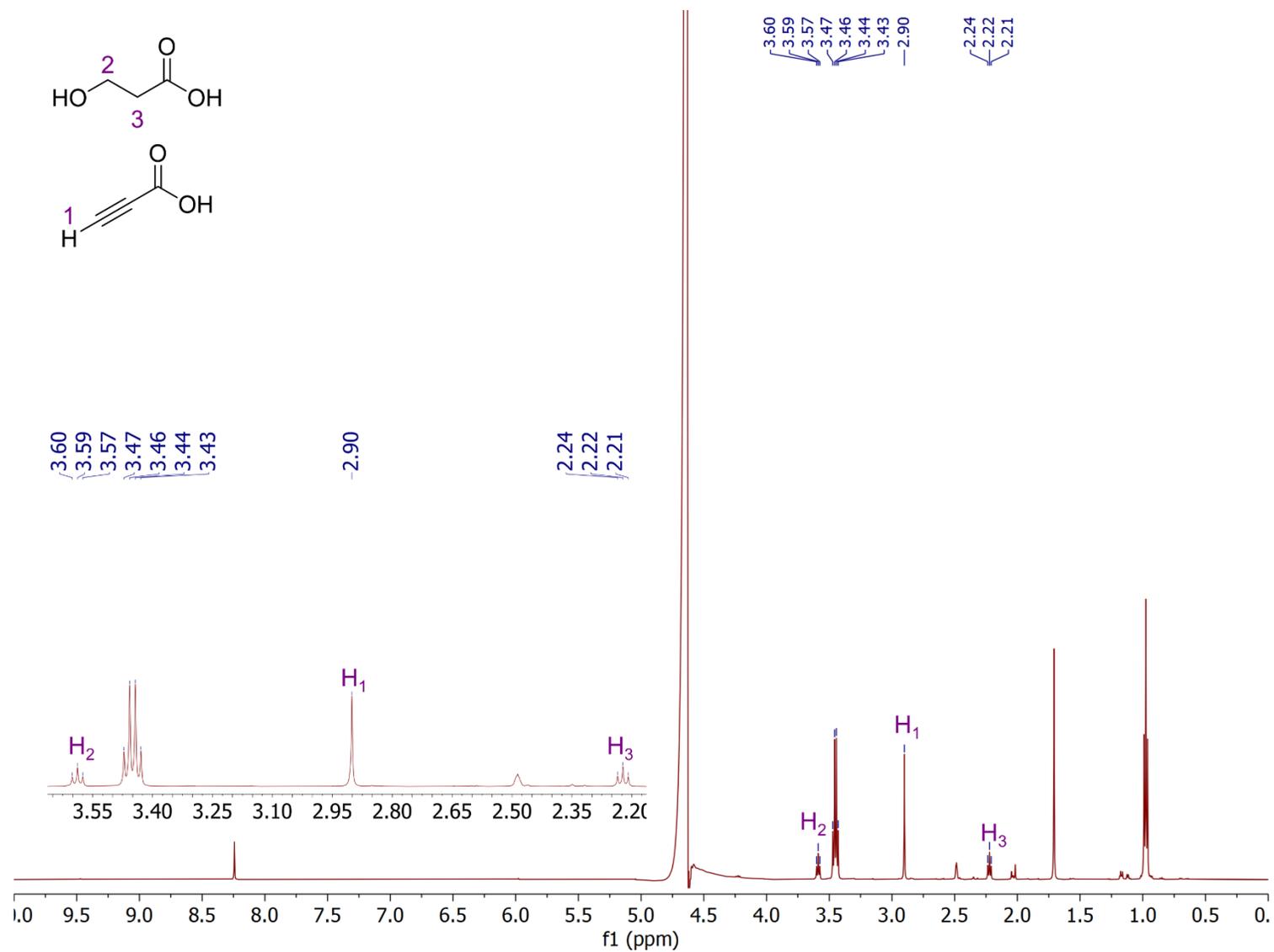


Figure 3.12. ^1H NMR of 3HPC-E114N in M9/glu/ACA media after 48 h. The panel on the top left is a zoomed in region of the spectra indicating the ACA and 3-HP signals.

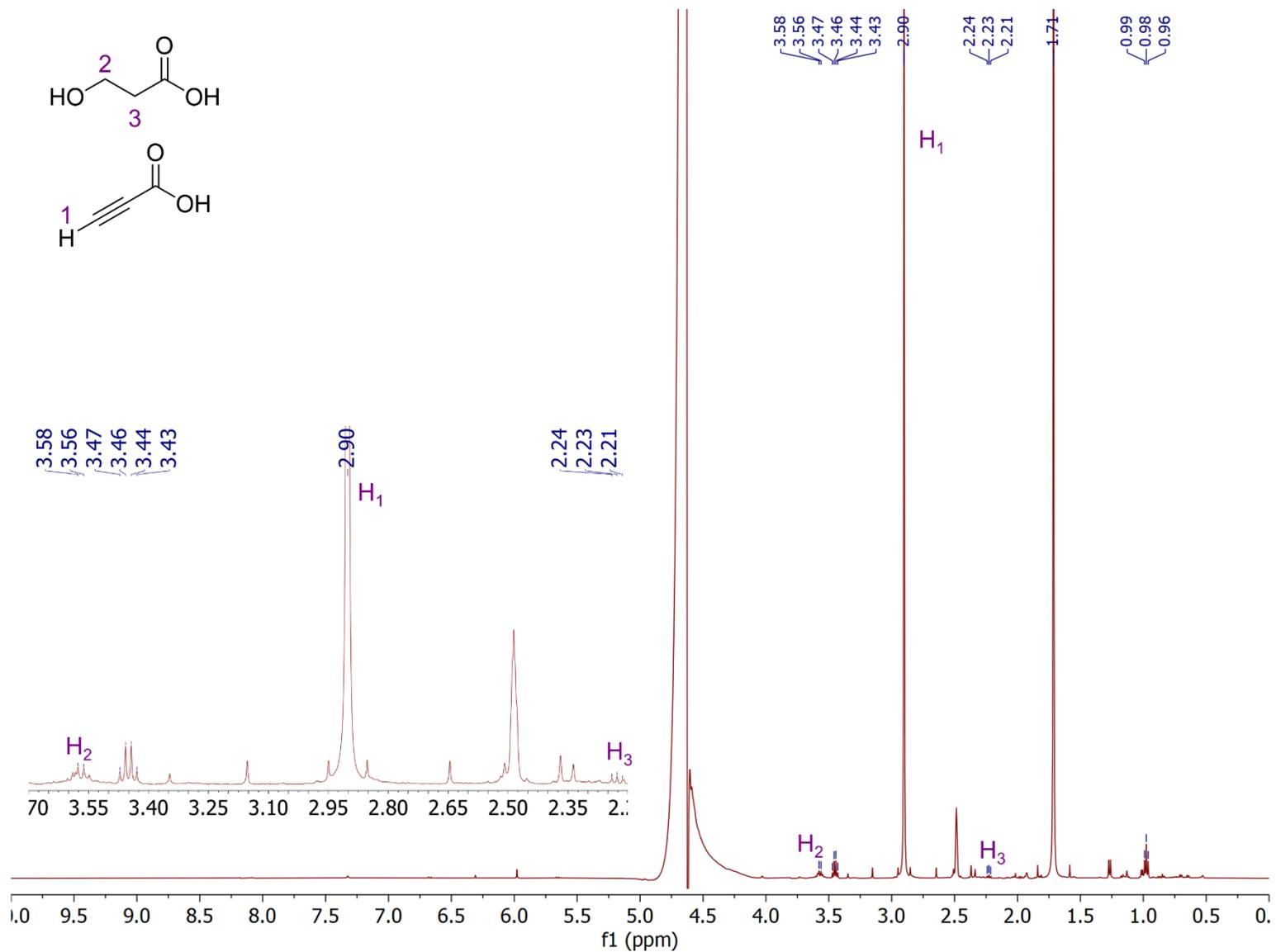


Figure 3.13. ^1H NMR of 3HPC-WT in M9/glu/ACA media after 48 h. The panel on the top left is a zoomed in region of the spectra indicating the ACA and 3-HP signals.

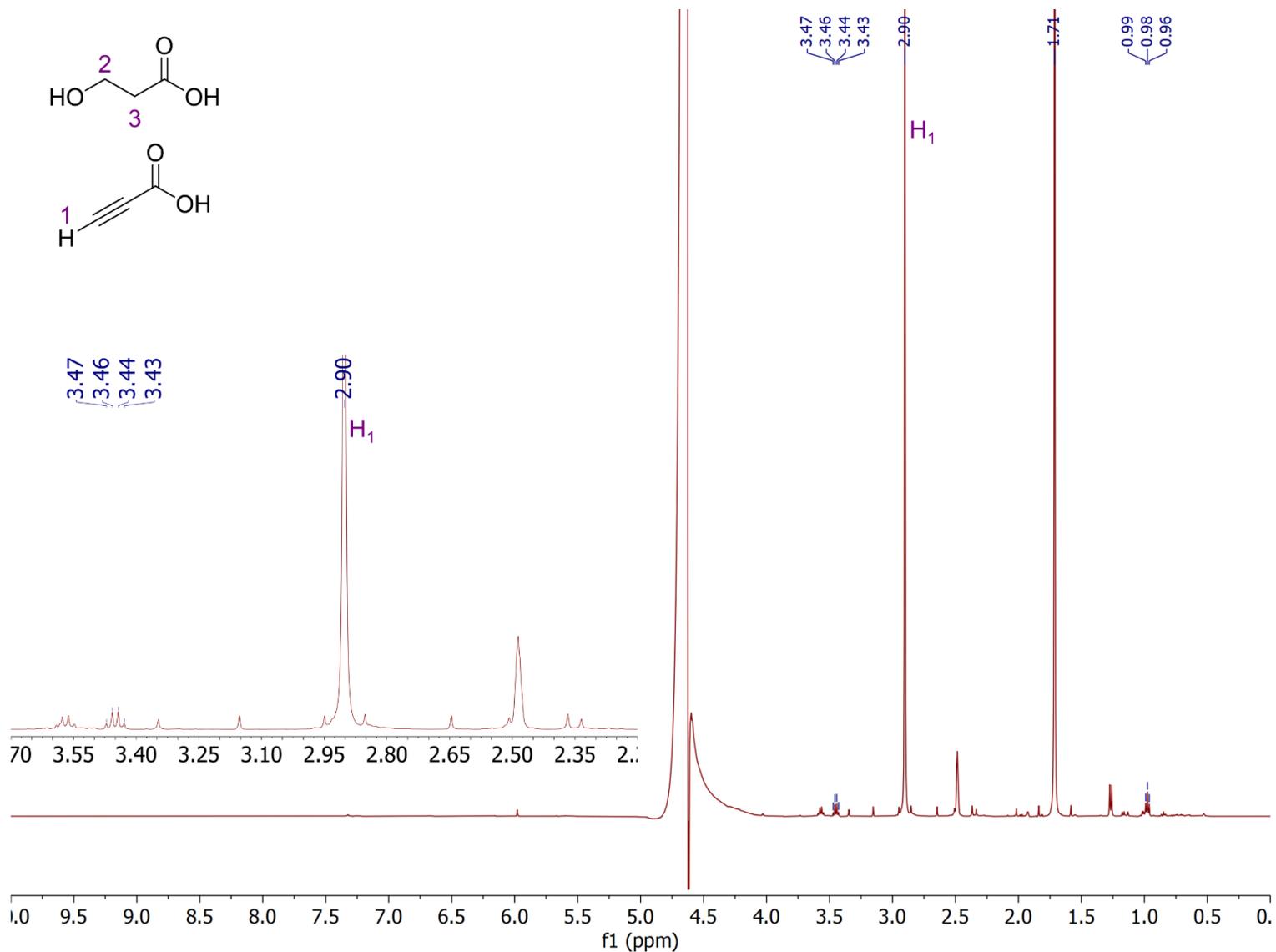


Figure 3.14. ^1H NMR of 3HPC in M9/glu/ACA media after 48 h. The panel on the top left is a zoomed in region of the spectra indicating the ACA.

3.5. Conclusion

Although genetically-encoded biosensors have become a valuable tool in molecular biology applications such as screening mutant libraries for enzyme evolution or determination of titers of a target metabolite *in vivo*, careful and thorough optimization of these systems are often required. Engineering biosensors for certain applications require fine-tuning of different biosensor elements that can vary significantly among different organisms. For instance, the MmsR-guided biosensor developed for 3-HP detection in *P. denitrificans* exhibits excellent time- and dose-dependent inducibility with 3-HP.¹⁷ However, introduction of the same biosensor system in *E. coli* and *P. putida* did not result in the same levels of inducibility.¹⁷

The 3-HP biosensor developed by Nguyen et. al,¹⁷ with some modifications, was selected as a screening tool for random mutagenesis of Cg10062. The promoter used for *mmsR* was replaced by P_{J23119} , a high-strength promoter designed for *E. coli*. As constructed and tested, the use of the 3-HP biosensor in *E. coli* MG1655 for Cg10062 evolution faces two significant challenges. Firstly, the lack of induction or fluorescent output signal even with a broad range of 3-HP concentrations (0–300 mM) is perhaps the biggest limitation. If higher concentrations of 3-HP (>300 mM) are required to achieve inducibility, this prevents the biosensor from being used as an effective screening tool due to the observed growth inhibition at high concentrations of 3-HP. Therefore, the biosensor requires further modification to achieve inducibility using lower concentrations of 3-HP. Assuming that 3-HP is transported into the intracellular matrix, it was hypothesized that the expression of large amounts of MmsR by the high-strength promoter leads to extremely tight repression of the P_{mmsA} -guided *rfp* gene. As discussed earlier in this work, the phenotypic characteristics of the biosensor strain which appears white, rather than pink, supports this hypothesis. The promoter strength can be altered via site-directed mutagenesis to replace the high-strength promoter with other medium and low strength promoters from the Anderson library.²⁸ It's expected that although the background fluorescence may increase, reducing the promoter strength will allow a dose-dependent response from the biosensor. Further optimization

of other transcriptional and translational elements such as RBS or plasmid copy numbers may be necessary to alter the dynamic range or sensitivity of the biosensor.

Secondly, while 3-HP was detected in 3HPC-E114N supplemented with ACA, quantitative conversion of ACA to 3-HP did not occur *in vivo*. This dilemma presents a roadblock for the screening strategy proposed in this study since successful screening for non-decarboxylating Cg10062 variants is possible only if ACA is efficiently transported into *E. coli* and subsequently converted to 3-HP. Although 3-HP was detected in the 3HPC-E114N media, the presence of ACA in the media indicates that ACA transport into *E. coli* may be limited under the growth conditions tested in this study. Although the exact mechanism of ACA uptake in *E. coli* is not known, it was hypothesized that monocarboxylate transporters in *E. coli* may be responsible for ACA transport.³⁵ Radio-labelled or ¹³C-labelled ACA may prove useful in identifying the degree and efficiency of ACA transport and discovery of method to increase transport into the cells.^{36,37} The diauxic growth of *E. coli*, where it preferentially transports and utilizes glucose for growth is an important consideration in improving ACA uptake. Comparison of transcriptome analysis of *E. coli* cultured in ACA-supplemented M9/glu media versus media containing exclusively ACA, may point towards up-regulation of certain genes that aid in ACA transport.^{38,39} Identifying its transport mechanism is a crucial factor in optimizing the *in vivo* conversion of ACA to 3-HP since the unreacted ACA in the media hints towards inefficient ACA uptake by the cell. It must be noted that altering the promoter strength of the *mmsR* gene and investigation of the ACA uptake mechanism are problems that may need to be addressed simultaneously for successful implementation of the biosensor. Tuning the biosensor elements to obtain a dose-dependent response to 3-HP induction will serve no purpose for Cg10062 evolution if the ACA substrate is not efficiently transported into the cell and converted to the desired inducer molecule. Genetically-encoded biosensors using small molecule-responsive transcription factors typically consist of a sensing module and a reporter module. The tuning and desired characteristics of each of these modules can vary significantly depending on the intended application or host strain and in order to successfully

employ the 3-HP biosensor as a quantitative biomolecular tool, careful optimization of each module must be addressed.

REFERENCES

REFERENCES

- (1) Li, D.; Wu, Q.; Reetz, M. T. Chapter Ten - Focused Rational Iterative Site-Specific Mutagenesis (FRISM). In *Methods in Enzymology*; Tawfik, D. S., Ed.; Enzyme Engineering and Evolution: General Methods; Academic Press, 2020; Vol. 643, pp 225–242.
- (2) Lingen, B.; Grötzinger, J.; Kolter, D.; Kula, M.-R.; Pohl, M. Improving the Carboligase Activity of Benzoylformate Decarboxylase from *Pseudomonas putida* by a Combination of Directed Evolution and Site-Directed Mutagenesis. *Protein Eng. Des. Sel.* 2002, 15 (7), 585–593.
- (3) Zhang, Z.-G.; Yi, Z.-L.; Pei, X.-Q.; Wu, Z.-L. Improving the Thermostability of *Geobacillus stearothermophilus* Xylanase XT6 by Directed Evolution and Site-Directed Mutagenesis. *Bioresour. Technol.* 2010, 101 (23), 9272–9278.
- (4) Labrou, N. E. Random Mutagenesis Methods for In Vitro Directed Enzyme Evolution. *Curr. Protein Pept. Sci.* 2010, 11 (1), 91–100.
- (5) Bansal, S.; Kundu, B. Chapter 39 - Protein Engineering: Methods and Applications. In *Advances in Protein Molecular and Structural Biology Methods*; Tripathi, T., Dubey, V. K., Eds.; Academic Press, 2022; pp 641–668.
- (6) Calzini, M. A.; Malico, A. A.; Mitchler, M. M.; Williams, G. J. Protein Engineering for Natural Product Biosynthesis and Synthetic Biology Applications. *Protein Eng. Des. Sel.* 2021, 34, 15.
- (7) Chen, H.; Ma, L.; Dai, H.; Fu, Y.; Wang, H.; Zhang, Y. Advances in Rational Protein Engineering toward Functional Architectures and Their Applications in Food Science. *J. Agric. Food Chem.* 2022, 70 (15), 4522–4533.
- (8) Yang, H.; Li, J.; Shin, H.; Du, G.; Liu, L.; Chen, J. Molecular Engineering of Industrial Enzymes: Recent Advances and Future Prospects. *Appl. Microbiol. Biotechnol.* 2014, 98 (1), 23–29.
- (9) Mannan, A. A.; Liu, D.; Zhang, F.; Oyarzún, D. A. Fundamental Design Principles for Transcription-Factor-Based Metabolite Biosensors. *ACS Synth. Biol.* 2017, 6 (10), 1851–1859.
- (10) Lee, S. K.; Keasling, J. D. A Propionate-Inducible Expression System for Enteric Bacteria. *Appl. Environ. Microbiol.* 2005, 71 (11), 6856–6862.
- (11) Rogers, J. K.; Church, G. M. Genetically Encoded Sensors Enable Real-Time Observation of Metabolite Production. *Proc. Natl. Acad. Sci.* 2016, 113 (9), 2388–2393.
- (12) Sullivan, M. J.; Curson, A. R. J.; Shearer, N.; Todd, J. D.; Green, R. T.; Johnston, A. W. B. Unusual Regulation of a Leaderless Operon Involved in the Catabolism of

- Dimethylsulfoniopropionate in *Rhodobacter sphaeroides*. *PLOS ONE* 2011, 6 (1), e15972.
- (13) Schallmeyer, M.; Frunzke, J.; Eggeling, L.; Marienhagen, J. Looking for the Pick of the Bunch: High-Throughput Screening of Producing Microorganisms with Biosensors. *Curr. Opin. Biotechnol.* 2014, 26, 148–154.
 - (14) Zhou, S.; Catherine, C.; Rathnasingh, C.; Somasundar, A.; Park, S. Production of 3-Hydroxypropionic Acid from Glycerol by Recombinant *Pseudomonas denitrificans*. *Biotechnol. Bioeng.* 2013, 110 (12), 3177–3187.
 - (15) Zhou, S.; Ashok, S.; Ko, Y.; Kim, D.-M.; Park, S. Development of a Deletion Mutant of *Pseudomonas denitrificans* That Does Not Degrade 3-Hydroxypropionic Acid. *Appl. Microbiol. Biotechnol.* 2014, 98 (10), 4389–4398.
 - (16) Zhou, S.; Ainala, S. K.; Seol, E.; Nguyen, T. T.; Park, S. Inducible Gene Expression System by 3-Hydroxypropionic Acid. *Biotechnol. Biofuels* 2015, 8 (1), 169.
 - (17) Nguyen, N. H.; Kim, J.-R.; Park, S. Development of Biosensor for 3-Hydroxypropionic Acid. *Biotechnol. Bioprocess Eng.* 2019, 24 (1), 109–118.
 - (18) Nguyen, N. H.; Kim, J.-R.; Park, S. Application of Transcription Factor-Based 3-Hydroxypropionic Acid Biosensor. *Biotechnol. Bioprocess Eng.* 2018, 23 (5), 564–572.
 - (19) Arasu, M. V.; Sarkar, R.; Sekar, B. S.; Kumar, V.; Rathnasingh, C.; Choi, J.; Song, H.; Seung, D.; Park, S. Isolation of a Novel *Pseudomonas* Species SP2 Producing Vitamin B₁₂ under Aerobic Condition. *Biotechnol. Bioprocess Eng.* 2013, 18 (1), 43–51.
 - (20) Lee, P.; Raj, S. M.; Zhou, S.; Ashok, S.; Edwardraja, S.; Park, S. 3-Hydroxyisobutyrate Dehydrogenase-I from *Pseudomonas denitrificans* ATCC 13867 Degrades 3-Hydroxypropionic Acid. *Biotechnol. Bioprocess Eng.* 2014, 19 (1), 1–7.
 - (21) Ainala, S. K.; Somasundar, A.; Park, S. Complete Genome Sequence of *Pseudomonas denitrificans* ATCC 13867. *Genome Announc.* 2013, 1 (3), e00257-13.
 - (22) Henikoff, S.; Haughn, G. W.; Calvo, J. M.; Wallace, J. C. A Large Family of Bacterial Activator Proteins. *Proc. Natl. Acad. Sci.* 1988, 85 (18), 6602–6606.
 - (23) Knapp, G. S.; Hu, J. C. Specificity of the *E. coli* LysR-Type Transcriptional Regulators. *PLOS ONE* 2010, 5 (12), e15189.
 - (24) Maddocks, S. E.; Oyston, P. C. F. Y. 2008. Structure and Function of the LysR-Type Transcriptional Regulator (LTTR) Family Proteins. *Microbiology* 154 (12), 3609–3623.
 - (25) Schell, M. A. Molecular Biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* 1993, 47 (1), 597–626.
 - (26) Anderson, Jc.; Dueber, J. E.; Leguia, M.; Wu, G. C.; Goler, J. A.; Arkin, A. P.; Keasling, J. D. BglBricks: A Flexible Standard for Biological Part Assembly. *J. Biol. Eng.* 2010, 4 (1), 1.

- (27) Lee, T. S.; Krupa, R. A.; Zhang, F.; Hajimorad, M.; Holtz, W. J.; Prasad, N.; Lee, S. K.; Keasling, J. D. BglBrick Vectors and Datasheets: A Synthetic Biology Platform for Gene Expression. *J. Biol. Eng.* 2011, 5 (1), 12.
- (28) *Promoters/Catalog/Anderson - parts.igem.org.*
<http://parts.igem.org/Promoters/Catalog/Anderson> (accessed 2022-06-30).
- (29) Campbell, R. E.; Tour, O.; Palmer, A. E.; Steinbach, P. A.; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. A Monomeric Red Fluorescent Protein. *Proc. Natl. Acad. Sci.* 2002, 99 (12), 7877–7882.
- (30) Mertens, N.; Remaut, E.; Fiers, W. Tight Transcriptional Control Mechanism Ensures Stable High-Level Expression from T7 Promoter-Based Expression Plasmids. *Bio/Technology* 1995, 13 (2), 175–179.
- (31) Vidra, A.; Németh, Á. Bio-Based 3-Hydroxypropionic Acid: A Review. *Period. Polytech. Chem. Eng.* 2018, 62 (2), 156–166.
- (32) Tingirikari, J. M. R.; Ahmed, S.; Yata, V. K. Chapter 6 - 3-Hydroxy-Propionic Acid. In *Platform Chemical Biorefinery*; Kaur Brar, S., Jyoti Sarma, S., Pakshirajan, K., Eds.; Elsevier: Amsterdam, 2016; pp 101–117.
- (33) Mathes Hewage, A.; Nayebi Gavvani, H.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochemistry* 2021, 60 (51), 3879–3886.
- (34) Huddleston, J. P.; Johnson, W. H.; Schroeder, G. K.; Whitman, C. P. Reactions of Cg10062, a *cis*-3-Chloroacrylic Acid Dehalogenase Homologue, with Acetylene and Allene Substrates: Evidence for a Hydration-Dependent Decarboxylation. *Biochemistry* 2015, 54 (19), 3009–3023.
- (35) Halestrap, A. P.; Wilson, M. C. The Monocarboxylate Transporter Family—Role and Regulation. *IUBMB Life* 2012, 64 (2), 109–119.
- (36) Albrich, J. M.; Gilbaugh, J. H.; Callahan, K. B.; Hurst, J. K. Effects of the Putative Neutrophil-Generated Toxin, Hypochlorous Acid, on Membrane Permeability and Transport Systems of *Escherichia coli*. *J. Clin. Invest.* 1986, 78 (1), 177–184.
- (37) Tucker, A. M.; Winkler, H. H.; Driskell, L. O.; Wood, D. O. S-Adenosylmethionine Transport in *Rickettsia prowazekii*. *J. Bacteriol.* 2003, 185 (10), 3031–3035.
- (38) Karapetyan, L.; Mikoyan, G.; Vassilian, A.; Valle, A.; Bolivar, J.; Trchounian, A.; Trchounian, K. *Escherichia coli* Dcu C4-Dicarboxylate Transporters Dependent Proton and Potassium Fluxes and FOF1-ATPase Activity during Glucose Fermentation at PH 7.5. *Bioelectrochemistry* 2021, 141, 107867.
- (39) Zhu, L.-W.; Zhang, L.; Wei, L.-N.; Li, H.-M.; Yuan, Z.-P.; Chen, T.; Tang, Y.-L.; Liang, X.-H.; Tang, Y.-J. Collaborative Regulation of CO₂ Transport and Fixation during Succinate Production in *Escherichia coli*. *Sci. Rep.* 2015, 5 (1), 17321.

CHAPTER FOUR: EXPERIMENTAL

4.1. Materials and equipment

Chemicals, biochemicals, Luria-Bertani (LB), Nutrient broth media components and buffer salts were purchased from MilliporeSigma, Becton, Dickinson and Company and Thermo Fisher Scientific. Acetylenecarboxylic acid (propionic acid), *cis*-3-chloroacrylic acid and 3-hydroxypropionic acid were purchased from MilliporeSigma. IPTG (isopropyl- β -D-1-thiogalactopyranoside) was purchased from Gold Biotechnology. Bio-Rad Protein Assay, Precision Plus Protein Electrophoresis Standards and Mini-PROTEAN[®] TGX[™] Precast 4-20% polyacrylamide gels were purchased from Bio-Rad. Restriction enzymes, T4 DNA ligase, Gibson Assembly[®] Cloning Kit, NEBuilder[®] HiFi DNA Assembly Cloning Kit, Q5[®] Site-Directed Mutagenesis Kit, Monarch[®] PCR and DNA Cleanup Kit, Q5[®] High-Fidelity DNA Polymerase and Taq DNA Polymerase were purchased from New England Biolabs. QIAprep Spin Miniprep and Maxiprep kits were purchased from Qiagen. Wizard[®] Genomic DNA Purification Kit was purchased from Promega. HisTrap[™] FF 1 mL and 5 mL pre-packed columns were purchased from Cytiva. Amicon Ultra-15 10K centrifugal filter units, 0.22 μ m and 0.45 μ m syringe filters were purchased from MilliporeSigma. UltraPure[™] agarose, low EEO agarose and MicroAmp[™] Optical Adhesive films were purchased from Thermo Fisher. dNTPs were purchased from Promega. Oligonucleotides were purchased from Integrated DNA Technologies. Commercially synthesized plasmids were obtained from Genscript. Alcohol dehydrogenase (catalog no: A7011) and aldehyde dehydrogenase (catalog no: 10171832001) from *Saccharomyces cerevisiae* were purchased from MilliporeSigma. PicoProbe[™] NADH Fluorometric Assay Kit (K338) was purchased from BioVision.

The OD₆₀₀ of cell cultures were measured using a NanoDrop[™] One Microvolume UV-Vis Spectrophotometer. Depending on the growth conditions, M9 or LB (1 mL) was used as a blank and sample (1 mL) were measured at 600 nm using a cuvette. In samples where the OD₆₀₀ > 1, a 10-fold dilution was carried out using M9 or LB to ensure an absorbance < 1. UV assays were

carried out on Molecular Devices SpectraMax[®] iD3 multi-mode microplate reader and Shimadzu UV2600 spectrophotometer. Protein purification was carried out on ÄKTA Start FPLC (Cytiva). NMR spectra were obtained using a 500 MHz Varian NMR spectrophotometer. HPLC analysis was carried out using RI detection on an Agilent 1100 HPLC system equipped with an Aminex HPX-87H column. PCR amplifications utilized a Bio-Rad DNA Engine[®] Peltier Thermal Cycler. Standard UV enzyme assays and 3-HP biosensor experiments were carried out in Grenier UV-Star[®] clear 96-well plates (catalog no: M3812) and Corning[®] costar black clear-bottomed 96-well plates (catalog no: CLS3603), respectively.

4.2. Culture media and stock solutions

Culture media were prepared according to standard protocols¹ using double deionized water and sterilized by autoclaving (liquid cycle, 121 °C, 25 min). LB (1 L) contained Bacto[™] tryptone (10 g), Difco[™] yeast extract (5 g) and NaCl (10 g) and was sterilized by autoclaving.² Nutrient Broth (NB) (1 L) contained beef extract (3 g) and peptone (5 g) and was sterilized by autoclaving. 2xYeast-tryptone (2xYT) (1 L) contained Bacto[™] tryptone (16 g), Difco[™] yeast extract (10 g) and NaCl (5 g) and was sterilized by autoclaving. SOB (1 L) contained Bacto[™] tryptone (20 g), Difco[™] yeast extract (5 g), NaCl (0.5 g), 250 mM KCl (10 mL) and the pH was adjusted to pH 7.0 using 10 N NaOH prior to autoclaving. Immediately prior to use, sterile 2 M MgCl₂ (5 mL) was added to SOB (1 L). SOC (1 L) was prepared by addition of 1 M D-glucose (20 mL) to sterile SOB medium (1 L). M9 salts (1 L) contained Na₂HPO₄ (6 g), KH₂PO₄ (3 g), NH₄Cl (1 g) and NaCl (0.5 g) and was autoclaved.¹ M9 media (1 L) was prepared by addition of sterile 20% (w/v) glucose (20 mL), 1 M MgSO₄ (2 mL), and 1 mg mL⁻¹ thiamine hydrochloride (1 mL). Where applicable, M9 media was supplemented with L-serine to a final concentration of 40 mg L⁻¹. CGXII media (1 L) contained (NH₄)₂SO₄ (20 g), urea (5 g), KH₂PO₄ (1 g), K₂HPO₄ (1 g), MgSO₄·7H₂O (0.25 g), MOPS (42 g), CaCl₂ (10 mg), FeSO₄·7H₂O (10 mg), CuSO₄ (0.2 mg) and NiCl₂·6H₂O (0.02 mg).³

Stock solutions of 2 M MgCl₂, 1 M MgSO₄, 250 mM KCl, 1 M glucose, and 20% (w/v) glucose were prepared separately using deionized water and sterilized by autoclaving. Stock solutions of thiamine hydrochloride (1 mg mL⁻¹), L-serine (40 mg mL⁻¹), ampicillin (50 mg mL⁻¹), kanamycin (50 mg mL⁻¹) and isopropyl β-D-1-thiogalactopyranoside (IPTG) (1 M) stock solutions were prepared in deionized water and sterilized via passage through a 0.22 μm filter. Antibiotics were added to all media where appropriate to the following final concentrations: ampicillin (Ap, 50 μg mL⁻¹) and kanamycin (Kan, 50 μg mL⁻¹). Solid media was prepared by the addition of Bacto™ Agar to a final concentration of 1.5% (w/v). Low EEO agarose was used for the preparation of M9 plates containing ACA (section 4.14.8).

4.3. Preparation and transformation of electrocompetent *E. coli*

Electrocompetent cells were prepared using standard laboratory operating procedures which were adapted from Sambrook and Russell.¹ A single bacterial colony from a streaked plate (< 5 days old) of the desired strain was used to inoculate an LB culture (5 mL), and the sample was incubated overnight with shaking (37 °C, 200 rpm). An aliquot of the overnight culture (2 mL) was used to inoculate 100 mL of sterile 2xYT media in a 500 mL baffled flask, and the sample was incubated at 37 °C until an OD₆₀₀ of 0.5-0.7 was reached. The cells were collected by centrifugation using a Fiberlite™ F12-6x500LEX fixed angle rotor (4,500 x g, 5 min, 4 °C). The supernatant was discarded and the cell pellet was gently resuspended in 100 mL cold, sterile deionized water to remove residual salts. The cells were pelleted by centrifugation (4,500 x g, 5 min, 4 °C), and the sample was carefully decanted to remove the supernatant. The wash step was repeated with an additional 100 mL cold, sterile deionized water. The cell pellet was resuspended in 100 mL cold, sterile glycerol (10% aqueous (v/v)) and centrifuged at (4,500 x g, 5 min, 4 °C). The supernatant was discarded and the cell pellet was resuspended in 0.5 mL glycerol (10% aqueous (v/v)). Aliquots (50 μL) were prepared in pre-chilled, sterile microcentrifuge tubes on ice and flash-frozen in liquid nitrogen prior to storage at -80 °C.

For the transformation of electrocompetent *E. coli* cells, plasmid DNA (2 μL of 1-5 $\text{ng } \mu\text{L}^{-1}$ in sterile deionized water) or purified, de-salted PCR or ligation products were combined with 50 μL of electrocompetent cells that were previously thawed on ice. Gentle pipetting of the DNA ensured sufficient mixing, and the sample was transferred to a cold, sterile Gene Pulser® electroporation cuvette (0.2 cm electrode gap). Electroporation was carried out using a Bio-Rad Gene Pulser II electroporation system (2.5 kV, 25 mF and 2000 Ω), which resulted in a time constant in the range of 5.16–5.25 ms. The cuvette was immediately placed on ice and 1 mL room temperature SOC media was added. The resulting recovery culture was incubated at 37 °C with shaking (1 h, 200 rpm). The cells were harvested by centrifugation at 13,000 rpm in a microcentrifuge, and 800 μL of the supernatant was discarded. The cell pellet was resuspended in the remaining media and 20 μL and 80 μL aliquots were plated onto LB solid media containing the appropriate antibiotic.

4.4. Isolation of plasmid DNA

For small scale purification of plasmid DNA, a single colony from a freshly transformed plate of *E. coli* was inoculated into LB (5 mL) containing the appropriate antibiotic and incubated overnight with shaking (37 °C, 200 rpm). Cells were harvested using a microcentrifuge (13,000 rpm, 1 min, rt). For large scale purification, a single colony from a freshly transformed plate of *E. coli* was inoculated into LB (100 mL) containing the appropriate antibiotic and incubated overnight with shaking (37 °C, 200 rpm). Cells were harvested by centrifugation (4,500 $\times g$, 10 min, rt). Isolation and purification of the plasmid DNA was carried out using, respectively, Qiagen Plasmid Miniprep and Midiprep Kits following the manufacturer's instructions. For the small scale purification, DNA was eluted from the spin column with 50 μL sterile, deionized water. For the large scale purification of DNA, the air-dried plasmid pellet was dissolved in 100–200 μL sterile, deionized water. All purified plasmids were stored at 4 °C.

4.5. Isolation of genomic DNA

To isolate genomic DNA from *E. coli* W3110 and *Pseudomonas denitrificans* (ATCC 13867), a single colony of each bacterium was inoculated into LB (5 mL) and incubated overnight with shaking (37 °C, 200 rpm). The cells were harvested using a microcentrifuge (17,000 x *g*, 1 min). Genomic DNA was isolated using Wizard® Genomic DNA Purification Kit following the protocol described by the manufacturer.

4.6. Restriction digestion of DNA

Digestion of plasmid DNA with restriction endonucleases followed by analysis of the products using gel electrophoresis was used to confirm the identity of plasmid DNA. A typical restriction digest (20 µL) contained the components listed in Table 4.1. To prepare larger quantities of DNA for cloning purposes, the digest was increased to 50 µL and all components were scaled appropriately. Restriction digests were incubated at 37 °C (1 h) and quenched by addition of 4 µL of 6x loading dye containing SDS (supplied by NEB) prior to visualization on a 0.7% agarose gel containing ethidium bromide (0.5 µg mL⁻¹) in 1X TAE (Tris-Acetate EDTA) buffer. 1X TAE was used as the running buffer with a standard voltage of 98 V (1 h).

Table 4.1. Components of a routine restriction digest

Component	Volume (µL)	Final Concentration
Plasmid DNA	varies	500-600 ng
10X Buffer	2	1X
Restriction enzyme 1	1	10 U
Restriction enzyme 2 (optional)	1	10 U
Deionized water	up to 20 µL	-

4.7. DNA sequencing

The DNA sequence of all plasmids prepared in this study were confirmed by Sanger sequencing prior to their use. Isolated DNA was sequenced at the Michigan State University

Research Technology Support Facility (MSU RTSF) Genomics Core using custom primers. For each plasmid sequenced, forward and reverse primers were designed 450 bp apart to obtain complete sequences of both DNA strands. Sequencing data received from MSU RTSF were analyzed using the “Align to Reference Sequence” tool on SnapGene 6.0.

4.8. PCR amplification

Unless otherwise specified, PCR amplifications utilized a Bio-Rad DNA Engine® Peltier Thermal Cycler using Q5® High-Fidelity DNA Polymerase. Reaction components for a typical amplification (50 μ L) are listed in Table 4.2. Typical cycling parameters are provided in Table 4.3. PCR samples were quenched with 10 μ L of 6x loading dye containing SDS (supplied by NEB) prior to visualization on a 0.7% agarose gel.

Table 4.2. Components of a routine PCR reaction

Component	Volume (μ L)	Final Concentration
5X Q5 reaction buffer	10	1X
10 mM dNTPs	1	200 μ M
10 μ M forward primer	2.5	0.5 μ M
10 μ M reverse primer	2.5	0.5 μ M
Template DNA	varies	1-5 ng
Q5 High-Fidelity DNA polymerase	0.5	0.02 U μ L ⁻¹
5X Q5 High GC Enhancer (optional)	10	1X
Deionized water	up to 50 μ L	-

Table 4.3. Standard temperature cycling parameters for PCR

Step	Temperature (°C)	Time
Initial Denaturation	98	30 sec
30 cycles	98	10 sec
	*50–72	30 sec
	72	30 sec kb ⁻¹
Final Extension	72	2 min
Hold	4	-

*Annealing temperature for each primer pair was determined using NEBasechanger.

4.9. Protein expression

Unless otherwise specified, purification of all proteins characterized in this study relied on inclusion of a His₆-tag on the C-terminus of the protein and subsequent immobilized metal affinity chromatography (IMAC). Protein expression plasmids were prepared from pET-21a(+) or pET-15b. Protein expression was carried out in *E. coli* BL21(DE3) using the T7 promoter expression system.⁴ *E. coli* BL21(DE3) was transformed with a plasmid encoding a His₆-tagged protein and a single colony was inoculated into LB/Ap (50 mL) and incubated overnight with shaking (37 °C, 200 rpm). For small scale expression, the overnight culture was used to inoculate LB/Ap (100 mL) to an initial OD₆₀₀ of 0.05. For larger scale expression, the overnight culture was used to inoculate LB/Ap (1 L). Cultures were incubated at 37 °C with shaking (200 rpm) until the OD₆₀₀ reached 0.5–0.7 (1.5–2 h). IPTG was added to a final concentration of 1 mM and cells were incubated with shaking at 30 °C (200 rpm, 8–10 h). Cells were collected by centrifugation using a Fiberlite™ F12-6x500LEX fixed angle rotor (4,500 x *g*, 10 min, 4 °C) and frozen cell paste was routinely stored at -20 °C for (0–30 days) prior to protein purification.

4.10. Protein purification

Proteins were purified on an ÄKTA Start FPLC system (Cytiva) using either HisTrap FF Nickel affinity columns (1 mL or 5 mL), depending on the original volume of cell culture. Cell pellets

were removed from -20 °C storage and thawed at room temperature. Cell pellets were resuspended in lysis buffer (20 mM sodium phosphate pH 7.4, 20 mM imidazole) to a concentration of 2 mL lysis buffer per gram of cell paste. The cell suspension was then lysed by at least two passages through a SLM Aminco French® Pressure cell (Thermo Scientific, Waltham, MA) at 18,000 psi. Cellular debris was separated from the cell lysate by centrifugation using a Fiberlite™ F21S-8x50y fixed angle rotor (47,500 x *g*, 4 °C, 30 min). Clarified lysate was decanted from the insoluble cell debris and filtered through a 0.45 µm sterile syringe filter. The protein purification process included five steps: equilibration, sample application, wash out unbound, elution and fractionation, and equilibration. Two buffers were used for protein purification on the FPLC: binding buffer contained 20 mM sodium phosphate, pH 7.4 and 500 mM NaCl, and elution buffer contained 20 mM sodium phosphate pH 7.4, 500 mM sodium chloride and 500 mM imidazole. A HisTrap FF column was equilibrated by passage of 5 column volumes of buffer (96:4 binding buffer: elution buffer) at a flow rate of 4.0 mL min⁻¹. Using the same ratio of buffers, the sample was loaded (flow rate 1.0 mL min⁻¹) and the column then eluted with 15 additional column volumes of buffer (4.0 mL min⁻¹). An imidazole gradient (20–500 mM) over 20 column volumes (96:4 binding buffer: elution buffer) at a flow rate of 4.0 mL min⁻¹ was used to elute bound protein. Fractions containing the desired protein, identified by an increase in absorbance at 280 nm were pooled, concentrated, and desalted using Amicon Ultra-15 10K filters. Proteins were stored in 100 mM sodium phosphate, pH 8 containing 20% glycerol, unless otherwise specified (see section 4.14.7.) Purification from 1 L of cell culture typically yielded 50–150 mg of purified protein. After a routine purification, cell lysate, FPLC fractions, and purified enzymes were visualized using SDS-PAGE.

4.11. Protein quantification

Protein concentrations of cell lysates were quantified using Bradford Assay Reagent purchased from Bio-Rad. The standard Bradford assay (5 mL) was used to quantify protein in crude cell lysates. Cell lysate (10 µL) was diluted with deionized water (90 µL) and Bradford

reagent (5 mL, room temperature) was added. Samples were incubated at room temperature for 5 min prior to measurement of the absorbance at 595 nm. Protein concentration was determined based on a standard curve prepared using bovine serum albumin standard (Figure 4.1).

Quantification of purified enzymes was carried out using guanidinium-hydrochloride.⁵ Purified enzyme (10 μL) was diluted with 990 μL of 6.0 M guanidinium-hydrochloride. A control sample containing 10 μL of the enzyme storage buffer, and 990 μL was also prepared. The samples were incubated at room temperature for 5 min prior to measurement of A_{280} . Enzyme concentrations were calculated using the molar extinction coefficients (280 nm) and molecular weights (including His-tag) using Protein Calculator v3.4⁶ (Table 4.4).

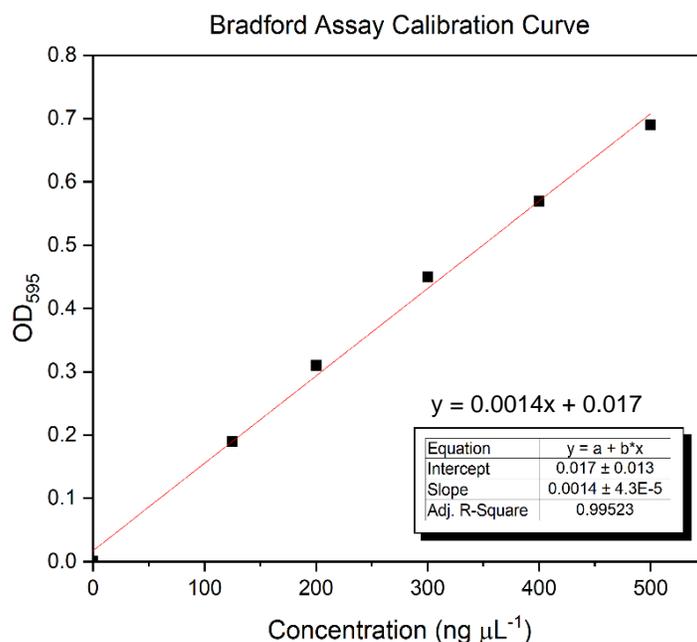


Figure 4.1. Standard calibration curve for the Bradford protein assay

Table 4.4. Molecular weights and molar extinction coefficients of enzymes

Enzyme	MW (Da)	$\epsilon_{280 \text{ nm}}$ ($\text{M}^{-1} \text{ cm}^{-1}$)
Cg10062 (WT)	19013	30440
Cg10062(E114Q)	19012	30440
Cg10062(E114D)	18999	30440
Cg10062(R70A)	18928	30440

Table 4.4 (cont'd)

Cg10062(R73A)	18928	30440
Cg10062(R70K)	18985	30440
Cg10062(R73K)	18985	30440
Cg10062(E114A)	18955	30440
Cg10062(H28A)	18947	30440
Cg10062(Y103A)	18921	29160
Cg10062(Y103F)	18997	29160
Cg10062(E114N)	18998	30440
Cg10062(E114S)	18971	30440
YdfG	28313	36130
MSAD	16464	8250
PTDH	36568	26600
RutE	24519	30440
TEV protease	27752	32290

The molecular weights and molar extinction coefficients were calculated using Protein Calculator v3.4.⁶

4.12. Polyacrylamide gel electrophoresis (SDS-PAGE)

Purified proteins were analyzed on 4–20% (w/v) acrylamide Mini-PROTEAN® TGX™ precast gels (Bio-Rad). Protein concentrations of samples were determined using Bradford reagent (cell lysates) or guanidinium-hydrochloride (purified protein), and stock solutions (300 ng μL^{-1}) were prepared in the buffer that the enzyme was stored in. Protein samples suitable for SDS-PAGE were prepared as follows. To 2x Laemmli Buffer (800 μL) containing SDS (4% w/v), glycerol (20% w/v), bromophenol blue (0.004% w/v) and 125 mM Tris-HCl (pH 6.8), 200 μL DTT (dithiothreitol) (2 M) was added to obtain a final concentration of 400 mM DTT. Then 2x Laemmli Buffer (50 μL) containing DTT was combined with protein sample (50 μL , 300 ng μL^{-1}) and incubated at 98 °C for 5 min. Samples were clarified by centrifugation (17,000 x g, 2 min).

Prepared protein samples (30 μ L) and Precision Plus Protein™ All Blue Ladder (10 μ L) (Bio-Rad) were loaded onto the precast gel. SDS-PAGE was carried out under denaturing conditions in running buffer (pH 8.3) containing SDS (0.1% w/v), L-glycine (192 mM) and Tris base (25 mM) at a constant voltage of 200 V. When the running dye was fully eluted (35-40 min), the gel was removed from the plastic housing and stained in a solution containing Coomassie Brilliant Blue R (0.1% w/v), MeOH (45% v/v) and AcOH (10% v/v) in water (1.5–2 h). The gels were subsequently destained in a solution containing MeOH (45% v/v) and AcOH (10% v/v) in water (6–12 h).

4.13. Distillation of acetylenecarboxylic acid (ACA)

Acetylenecarboxylic acid (ACA, 95%) (ACA) purchased from MilliporeSigma was purified by short-path distillation. Glassware was rinsed with acetone and oven-dried overnight. ACA was transferred to a tared round-bottom flask. The short-path distillation apparatus was set up using a short-path distillation head connected to a cow-type distilling receiver. The cow-type distilling receiver was connected to three round-bottom flasks for collection of distillates. An overhead thermometer was used to monitor the temperature of the distillate inside the short-path distillation head. The neck of the short-path distillation head was insulated using glass wool and was connected to a vacuum trap filled with indicating DRIERITE® desiccant. All glass joints were greased with vacuum grease to ensure a tight seal. A vacuum was applied using a vacuum pump and the round-bottom flask containing the ACA was gently heated with stirring using a hot water bath. The temperature was controlled using an IKA® ETS-D5 contact thermometer and an IKA® C-MAG HS 7 magnetic stirrer. The initial distillate was collected starting at 58 °C (overhead thermometer) until the temperature was steadily maintained. Purified ACA was collected in the second round-bottom flask connected to the cow-type distilling receiver (62 – 65 °C). The distilled ACA was transferred to glass vials (1 mL aliquots) inside a glove bag flushed with nitrogen and stored at -20 °C. Crude ACA and distillates collected from the short-path distillation were characterized by ¹H NMR (10 μ L sample in 690 μ L D₂O).

4.14. Bacterial strains and plasmids

Escherichia coli strains BL21(DE3), BL21 and DH5 α were obtained from Invitrogen (Table 4.5). Cells were grown at 37 °C in LB containing antibiotic as appropriate. *E. coli* RB791 *serA* was previously constructed by the Frost group at Michigan State University.^{7,8} Plasmids pET-21a(+), pET-28a(+), and pET-15b were obtained from Invitrogen and pBbA1a-RFP was purchased from Addgene.^{9,10} The plasmid construct expressing a His₇-tagged TEV protease (pMHT Δ 238) and *E. coli* BL21-CodonPlus(DE3)-RP were provided by Professor Heedeok Hong of Michigan State University.¹¹ Plasmid pKM6.240 was previously constructed by Dr. Kelly Miller using the pBbA1a-RFP vector.¹² Plasmid pET15b-12x was obtained from Addgene (#61699).¹³ The strains and plasmids used in this study are listed in Table 4.5.

Table 4.5. Strains and plasmids used in this study

Strain/Plasmid	Genotype/Description	Source
<i>E. coli</i> DH5 α	F ⁻ ϕ 80/ <i>lacZ</i> Δ M15 Δ (<i>lacZYA-argF</i>)U169 <i>recA1 endA1 hsdR17</i> (r _K ⁻ , m _K ⁺) <i>phoA supE44 λ thi-1 gyrA96 relA1</i>	Invitrogen
<i>E. coli</i> BL21(DE3)	F ⁻ <i>ompT hsdS_B</i> (r _B ⁻ , m _B ⁻) <i>gal dcm</i> (DE3)	Invitrogen
<i>E. coli</i> BL21	<i>fhuA2 [lon] ompT gal [dcm] ΔhsdS</i>	Invitrogen
<i>E. coli</i> RB791 <i>serA</i>	<i>E. coli</i> W3110 <i>lacL8^f serA::aroB</i>	Frost ^{8,9}
<i>E. coli</i> BL21-CodonPlus(DE3)-RP	F ⁻ <i>ompT hsdSx</i> (r _B ⁻ m _B ⁻) <i>dcm⁺ Tet^R gal λ</i> (DE3) <i>endA Hte [argU proL Cam^R]</i>	Agilent Technologies
<i>Corynebacterium glutamicum</i>	NCBI 10025	ATCC 13032
<i>Pseudomonas putida</i> KT2440	<i>rmo⁻ mod⁺</i>	ATCC 47054
<i>Pseudomonas denitrificans</i>	NCBI 9496	ATCC 13867
pET-21a(+)	Ap ^R , <i>lacI</i> , P _{T7} pMB1 replicon	Invitrogen
pET-15b	Ap ^R , <i>lacI</i> , P _{T7} pMB1 replicon	Invitrogen
pET-28a(+)	Kan ^R , <i>lacI</i> , P _{T7} pMB1 replicon	Invitrogen

Table 4.5. (cont'd)

pAS1.046	<i>P_{T7}cg10062</i> in pET-21a(+)	this study
pAS1.098	<i>cg10062(E114Q)</i> in pAS1.046	this study
pAS1.099	<i>cg10062(E114D)</i> in pAS1.046	this study
pAS2.096	<i>cg10062(R70K)</i> in pAS1.046	this study
pAS2.097	<i>cg10062(R73K)</i> in pAS1.046	this study
pAS2.098	<i>cg10062(Y103A)</i> in pAS1.046	this study
pAS2.099	<i>cg10062(Y103F)</i> in pAS1.046	this study
pAS2.100	<i>cg10062(E114N)</i> in pAS1.046	this study
pAS2.101	<i>cg10062(R70A)</i> in pAS1.046	this study
pAS2.102	<i>cg10062(R73A)</i> in pAS1.046	this study
pAS2.103	<i>cg10062(H28A)</i> in pAS1.046	this study
pAS2.104	<i>cg10062(E114A)</i> in pAS1.046	this study
pAS2.105	<i>cg10062(E114D-Y103F)</i> in pAS1.046	this study
pAS2.169	<i>cg10062(E114S)</i> in pAS1.046	this study
pMHTΔ238	Kan ^R , TEV protease expression vector	Hong ¹¹
pBbA1a-RFP	Ap ^R , <i>lacI^q</i> , <i>P_{trc}rfp</i> p15A replicon	Addgene
pKM6.240	<i>P_{trc}serA</i> in pBbA1a	Frost lab ¹²
pAS1.163	<i>P_{trc}cg10062(WT)</i> in pKM6.240	this study
pAS1.240	<i>cg10062(E114Q)</i> in pAS1.163	this study
pAS1.241	<i>cg10062(E114D)</i> in pAS1.163	this study
pAS2.084	<i>P_{T7}ydfG</i> in pET-21a(+)	this study
pAS2.141	<i>P_{T7}ydfG</i> in pET-15b	this study
pAS2.144	<i>P_{T7}rutE</i> in pET-15b	this study
pET15b-12x	<i>P_{T7}ptdh</i> in pET-15b	Addgene ¹³
pAS3.112	<i>P_{trc}cg10062(E114N) P_{trc}ydfG</i> in pBbA1a-RFP	this study

Table 4.5. (cont'd)

pAS5.003	Ap ^R , <i>P</i> _{J23102} <i>mmsR P_{mmsA}rfp</i> p15A replicon	this study
pAS5.005	<i>P</i> _{J23119} <i>cg10062</i> , <i>P</i> _{J23119} <i>ydfG</i> in pET-28a(+)	this study
pAS5.008	<i>P</i> _{J23119} <i>cg10062(E114N)</i> , <i>P</i> _{J23119} <i>ydfG</i> in pET-28a(+)	this study

4.15. Chapter One: Characterization of Cg10062 and variants

Reprinted with permission from Mathes Hewage, A.; Gavgani, H. N.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochemistry* 2021, 60 (51), 3879–3886.¹⁴ Copyright 2022 American Chemical Society.

4.15.1. Genes and plasmids

Plasmid pAS1.046

Plasmid pAS1.046 was synthesized by Genscript using the following modifications. The gene encoding wild type Cg10062 from *Corynebacterium glutamicum* was codon-optimized for expression in *E. coli* and modified to replace the stop codon with a TEV protease recognition site (ENLYFQG) (Figure 4.2).¹¹ The modified gene was cloned into pET-21a(+) at the NdeI and XhoI restriction sites at the 5' and 3' ends, respectively. This allowed expression of a His₆-tagged Cg10062 enzyme downstream of a T7 promoter from the resulting vector pAS1.046 (Figure 4.3). Restriction digestion of pAS1.046 with NdeI and XhoI resulted in DNA fragments of 0.47 kb and 5.4 kb. The gene sequence of codon-optimized Cg10062 was submitted to NCBI (accession number MZ369159).¹⁵ All other plasmids encoding Cg10062 or a variant of Cg10062 were prepared from pAS1.046 or pAS1.099 as the template for Q5 site-directed mutagenesis).

5' – **CATATG**CCGACCTACACCTGCTGGAGCCAACGCATTTCGTATTAGCCGTGAAG
 CGAAGCAACGCATCGCGGAAGCGATTACCGACGCGCACCATGAACTGGCGCAC
 GCGCCGAAGTACCTGGTGCAGGTTATTTTCAACGAAGTGGAGCCGGACAGCTATT
 TTATCGCGGCGCAGAGCGCGAGCGAGAACCACATTTGGGTTCAAGCGACCATCC
 GTAGCGGCCGTACCGAAAAGCAGAAAGAGGAACTGCTGCTGCGTCTGACCCAAG
 AGATCGCGCTGATTCTGGGTATCCCGAACGAGGAAGTGTGGGTTTACATTACCGA
 AATCCCGGGTAGCAACATGACCGAATATGGCCGTCTGCTGATGGAGCCGGGCGA
 GGAAGAGAAATGGTTCAACAGCCTGCCGGAGGGCCTGCGTGAGCGTCTGACCG
 AACTGGAGGGTAGCAGCGAA**GAGAACCTGTATTTTCAAGGCCTCGAG** – 3'

Figure 4.2. Nucleotide sequence of *cg10062* optimized for expression in *E. coli*. The NdeI and XhoI restriction sites are highlighted in yellow, and the TEV protease recognition site (ENLYFQG) is highlighted in blue.

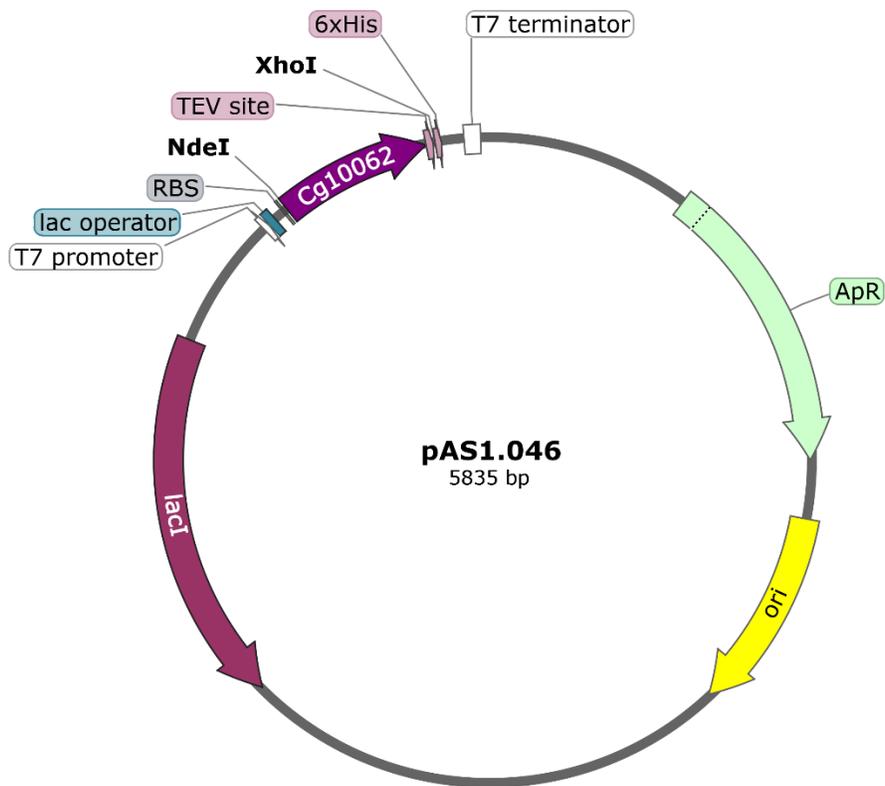


Figure 4.3. Plasmid map of pAS1.046

Plasmid pAS2.031

The plasmid encoding malonate semialdehyde decarboxylase (MSAD) from *Coryneform* bacterium strain FG41¹⁶ was codon-optimized for expression in *E. coli* and synthesized using the same modifications (Figure 4.4) as described for wild type Cg10062, and resulted in plasmid pAS2.031 (Figure 4.5). Restriction digestion of pAS2.031 with NdeI and XhoI resulted in DNA

fragments of 0.43 kb and 5.4 kb. The gene sequence of codon-optimized MSAD was submitted to NCBI (accession number MZ369160).

5' – **CATATG**CCCTTAATCCGTATAGATCTTACCAGTGATCGTTTCGAGAGAGCAACGGC
 GGGCGATTGCTGATGCAGTCCATGACGCTTTAGTAGAAGTTTTAGCGATTCCGGCTC
 GTGATCGCTTCCAGATACTGACTGCGCACGATCCCTCTGATATTATAGCCGAAGATG
 CTGGACTTGGCTTTCAGCGGTCCCCCAGTGTAGTCATCATAACACGTCTTTACACAGG
 CAGGTAGAACTATTGAAACGAAACAGAGAGTATTTGCAGCGATAACAGAAAGTCTGG
 CTCCAATCGGTGTTGCAGGATCTGATGTTTTTATCGCAATCACCGAAAATGCACCCC
 ATGACTGGAGCTTTGGGTTTGGCAGTGCACAATATGTCACGGGTGAACTTGCGATTC
 CAGCCACTGGTGC GGCT**GAGAACCTGTATTTCAAGGCCTCGAG** – 3'

Figure 4.4. Codon-optimized nucleotide sequence of MSAD. The NdeI and XhoI restriction sites are highlighted in yellow, and the TEV protease recognition site (ENLYFQG) is highlighted in blue.

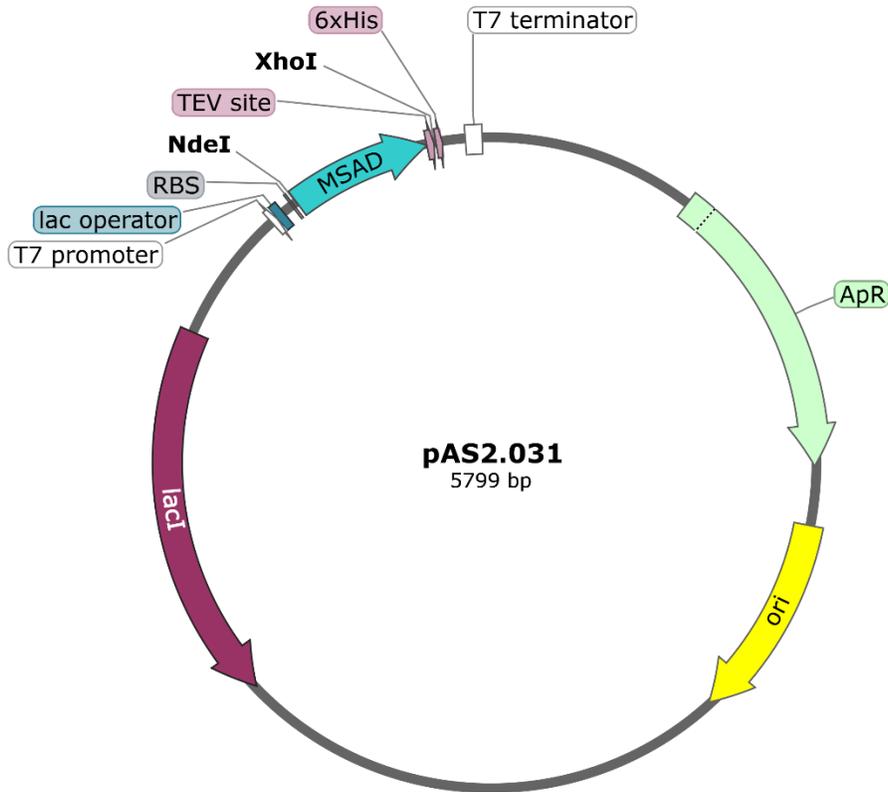


Figure 4.5. Plasmid map of pAS2.031

Plasmids pAS1.163, pAS1.240, pAS1.241

The gene encoding Cg10062 was cloned into pKM6.240 downstream of a *trc* promoter using the Bgl-Brick cloning^{10,11} technique (Figure 4.6). The *cg10062* gene was amplified from pAS1.046. EcoRI and BglII, and BamHI and XhoI restriction sites were inserted using the forward (AS004: 5'-ATTTTCAGAATTCAAAAAGATCTAAGAAGGAGATATACATATGCCGACCTACACCTGC-3') and reverse (AS005: 5'-AAATCTCTCGAGAAAGGATCCAAATTATTCGCTGCTACCCTCCAGTTCGGT-3') primers, respectively, to create the *cg10062* Bgl-Brick gene. Plasmid pKM6.240 was digested with EcoRI and BglII and *cg10062* was digested with EcoRI and BamHI. The two fragments were combined and incubated at room temperature with T4 DNA ligase (2 h) to obtain pAS1.163. Variants of pAS1.163 which include E114Q (pAS1.240) and E114D (pAS1.241) mutations were prepared from pAS1.163 using Q5 site-directed mutagenesis. Plasmid pAS1.240 was prepared using forward primer AS001 (5'-CAACATGACCCAGTATGGCCGTC-3') and reverse primer AS002 (5'-CTACCCGGGATTTTCGGTA-3'). Plasmid pAS1.241 was prepared using forward primer AS003 (5'-CAACATGACCGATTATGGCCGTC-3') and reverse primer AS004 (5'-CTACCCGGGATTTTCGGTA-3'). Digestion of pAS1.163, pAS1.240 and pAS1.241 with EcoRI and NcoI resulted in DNA fragments of 0.5 kb and 4.8 kb.

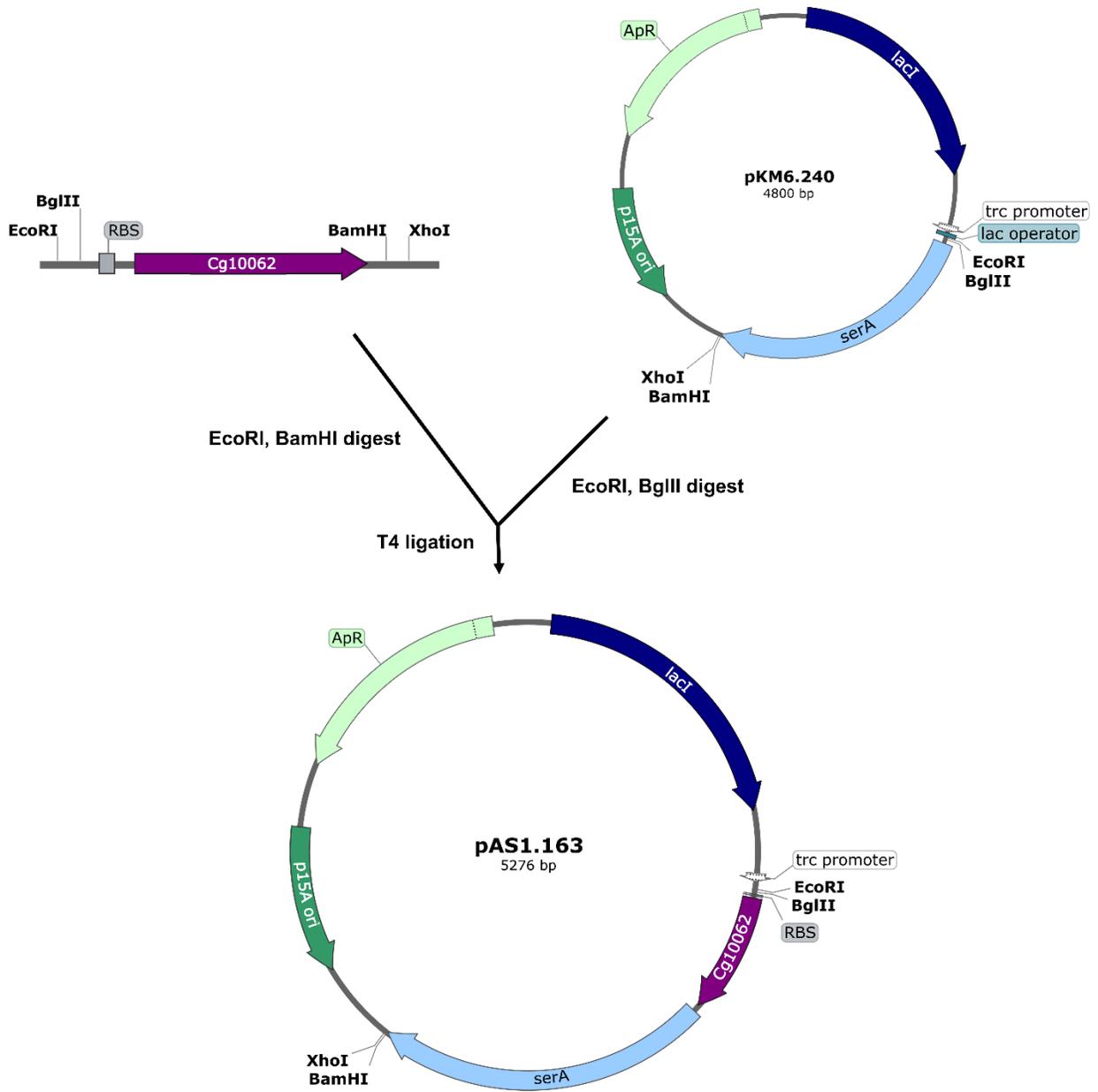


Figure 4.6. Construction of plasmid pAS1.163 expressing Cg10062. Plasmids pAS1.240 and pAS1.241 expressing E114Q and E114D variants, respectively were prepared using Q5 site-directed mutagenesis of pAS1.163.

4.15.2. Expression and purification of TEV protease

Plasmid pMHTΔ238 encodes a His₇-tagged TEV protease.^{11,17} The expression and purification of this enzyme was carried out using methods described previously by Blommel and Fox.¹⁷ *E. coli* BL21-CodonPlus(DE3) RP was used as an expression host. The purified protein was stored in 100 mM sodium phosphate pH 8, 20% glycerol at -20 °C.

4.15.3. His₆-tag removal by TEV protease

TEV protease was used for removal of His₆-tag from proteins, where appropriate. The His₆-tagged protein of interest was purified via affinity chromatography and fractions containing protein were combined and concentrated (200 μL) in 50 mM sodium phosphate, pH 8, 0.5 mM EDTA and 1 mM DTT and reacted with TEV protease (His₆-tagged protein: TEV molar ratio = 1:50) at 4 °C on a rocking platform (10 rpm) for 12 h. The enzyme mixture was re-purified via affinity chromatography and the protein was collected from the flowthrough. The protein of interest was concentrated in 100 mM sodium phosphate, pH 8.

4.15.4. Q5 site-directed mutagenesis

NEB Q5 site-directed mutagenesis kit was used to generate all Cg10062 variants described here. Unless otherwise specified, pAS1.046 was used as a parent template. Q5 site-directed mutagenesis was completed in three steps. Step 1 included PCR from parent template (Table 4.3) using the components listed in Table 4.6. The primers used to generate each Cg10062 variant are listed in Table 4.7. Step 2 was Kinase, Ligase and DpnI (KLD) treatment (Table 4.8) of the PCR product from step 1. Following KLD treatment, the sample from step 2 was purified using a Monarch® PCR and DNA Cleanup Kit to remove buffer salts. The final step was transformation of electrocompetent *E. coli* DH5α with the KLD product to isolate the plasmid with the desired modification. Sanger sequencing was used to confirm the accuracy of the modified sequence.

Table 4.6. Components of a routine Q5 site-directed mutagenesis PCR

Component	Volume (μ L)	Final Concentration
Plasmid DNA	varies	1–5 ng
10X Buffer*	2	1X
Restriction enzyme 1	1	10 U
Restriction enzyme 2 (optional)	1	10 U
Deionized water	up to 20 μ L	-

Table 4.7. Primers used to generate Cg10062 variants

Cg10062 variant	Primer	Sequence ^a
E114Q	AS001	CAACATGACCC <u>AGT</u> TATGGCCGTC
	AS002	CTACCCGGGATTTTCGGTA
E114D	AS003	CAACATGACCC <u>GATT</u> TATGGCCGTC
	AS004	CTACCCGGGATTTTCGGTA
E114N	AS026	CAACATGACCA <u>ACT</u> TATGGCCGTCTG
	AS025	CTACCCGGGATTTTCGGTA
H28A	AS012	TACCGACGCGGCATGAACTGGCGCACG
	AS013	ATCGCTTCCGCGATGCGT
R70A	AS014	AGCGACCATC <u>GCG</u> AGCGGCCGTAC
	AS015	TGAACCCAAATGTGGTTC
R73A	AS018	CCGTAGCGGC <u>GCG</u> ACCGAAAAGC
	AS019	ATGGTCGCTTGAACCCAA
R70K	AS016	AGCGACCATC <u>AAA</u> AGCGGCCGTA
	AS017	TGAACCCAAATGTGGTTCTC
R73K	AS020	CCGTAGCGGC <u>AAA</u> ACCGAAAAGC
	AS019	ATGGTCGCTTGAACCCAA
Y103A	AS021	AGTGTGGGTT <u>GCG</u> ATTACCGAAATCCCGGG
	AS022	TCCTCGTTCGGGATACCC

Table 4.7. (cont'd)

Y103F	AS023	AGTGTGGGTTTTATTACCGAAATCCC
	AS022	TCCTCGTTCGGGATACCC
E114A	AS024	CAACATGACCGCGTATGGCCGTCTG
	AS002	CTACCCGGGATTTTCGGTA
E114D-Y103F ^b	AS023	CAACATGACCGCGTATGGCCGTCTG
	AS002	CTACCCGGGATTTTCGGTA
E114S	AS027	CAACATGACCTCTTATGGCCGTCTGC
	AS025	CTACCCGGGATTTTCGGTA

^aCodons that introduced mutations are underlined. ^bPlasmid pAS1.099 served as a template.

Table 4.8. KLD treatment of PCR product in Q5 site-directed mutagenesis

Component	Volume (μL)	Final Concentration
PCR Product	1	-
2X KLD Reaction Buffer	5	1X
10X KLD Enzyme Mix	1	1X
Deionized water	3	-

^aKLD reaction incubated at room temperature for 5 min.

4.15.5. Kinetic characterization of Cg10062 and variants

Steady-state kinetics were carried out at 25 °C using a Molecular Devices SpectraMax® iD3 multi-mode microplate reader and Shimadzu UV2600 spectrophotometer. All assays were carried out in triplicate. The enzyme activity of Cg10062 and variants were measured using a coupled enzyme assay described by Whitman and coworkers,^{7,18,19} with some modifications. MSAD (malonate semialdehyde decarboxylase) is used to convert MSA (malonate semialdehyde) formed by Cg10062 and/or variants to acetaldehyde. The reduction of acetaldehyde (formed by either MSAD or Cg10062 and variants) by NADH-dependent alcohol dehydrogenase (ADH) was monitored by following the change in absorbance of NADH at 340 nm ($\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$) (Figure 4.7).

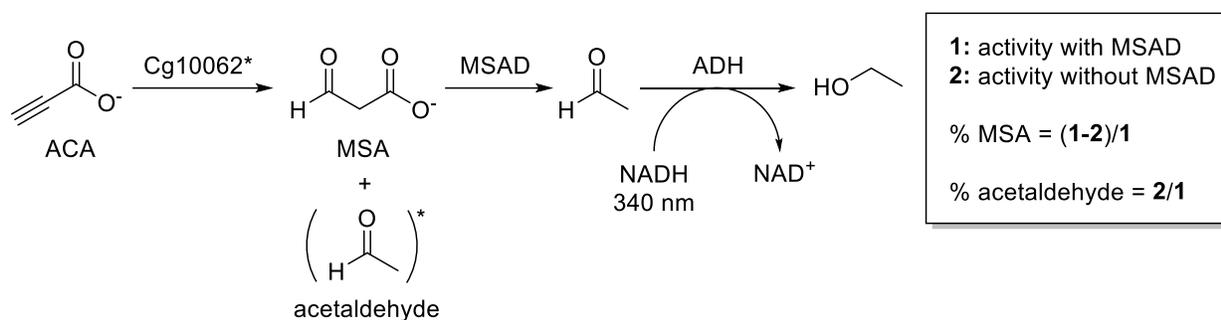


Figure 4.7. The coupled enzyme assay used for Cg10062 and variants

Kinetic assays used to quantify relative hydratase and hydratase/decarboxylase activities were carried out at 25 °C in 100 mM sodium phosphate, pH 8.0, and all stocks solutions (except ADH) were prepared in 100 mM sodium phosphate, pH 8.0. The final volume for each assay was 200 μL . Stock solutions of lyophilized ADH were prepared for use by addition of deionized water per manufacturer instructions. The initial screens used to determine hydratase or hydratase/decarboxylase activity contained the components as listed in Table 4.9. ACA stock solutions were prepared in volumetric flasks by diluting the appropriate volume of ACA in sterile 100 mM sodium phosphate, pH 8 and adjusting the pH back to 8 by addition of 10 N NaOH. Excess quantities of coupling enzymes MSAD and ADH were added to ensure that Cg10062 or its variants were rate-limiting. The amount of Cg10062 or variant used was varied to observe measurable activity. Thus, rates obtained from these initial assays were used only for determination of relative product ratios for each enzyme.

Table 4.9. Components used in the Cg10062 coupled enzyme assay

Component	Volume (μL)	Final Concentration/Units
Cg10062 or variant	varies	0.025 – 0.5 mg mL^{-1}
MSAD (4 mg mL^{-1} ; 30 U mg^{-1})	10	0.2 mg mL^{-1} (1.2 U)
ADH (1 mg mL^{-1} ; 300 U mg^{-1})	40	0.2 mg mL^{-1} (12 U)
ACA (5 mM, pH 8)	20	0.5 mM
NADH (6.7 mM)	10	0.34 mM

Table 4.9 (cont'd)

100 mM sodium phosphate, pH 8	up to 200 μ L	-
-------------------------------	-------------------	---

Specific activities ($\mu\text{mol min}^{-1} \text{mg}^{-1}$, U mg^{-1}) of Cg10062 variants with measurable activity were determined using Equation 4.1. Since protein concentrations varied, rates obtained from these experiments were not used to directly compare enzyme activities and were only used to establish product ratios for each variant as illustrated in Figure 4.7. The specific activities of MSAD and ADH were also calculated using Equation 4.1. MSAD specific activity was determined using the Cg10062 coupled enzyme assay and is described in Section 4.14.5. The specific activity of ADH (300 U mg^{-1}) was provided by the manufacturer.

Equation 4.1. Calculation of enzyme specific activity using a UV-based assay

$$\text{Specific Activity} \left(\frac{\text{U}}{\text{mg}} \right) = \text{Rate} \left(\frac{\text{mAU}}{\text{min}} \right) \times \frac{\text{AU}}{1000 \text{ mAU}} \times \frac{\text{mol cm}}{6220 \text{ L}} \times \frac{1}{b \text{ cm}} \times \frac{\text{mL}}{x \text{ mg}} \times \frac{z \text{ mL}}{y \text{ mL}} \times \frac{10^6 \mu\text{mol}}{1 \text{ mol}} \times \frac{1 \text{ L}}{1000 \text{ mL}}$$

$$b = \text{path length (cm)}; x = [\text{enzyme}]_{\text{stock}} (\text{mg mL}^{-1});$$

$$y = \text{volume of enzyme}_{\text{stock}} \text{ used (mL)}; z = \text{final assay volume (mL)}$$

Variants that showed no measurable activity were not characterized further. The most active Cg10062 variants were selected for further characterization and determination of kinetic parameters. Initial rates (5-15 s) of these variants with varying ACA concentrations (1 – 5000 μM) were plotted to fit the Michaelis-Menten model to determine kinetic parameters (k_{cat} and K_m) using OriginPro.

4.15.6. Specific activity of MSAD

The specific activity of MSAD was measured using the same coupled enzyme assay as for Cg10062 (Figure 4.7). Assays were carried out in triplicate (1 mL) at 25 °C using a Shimadzu UV2600 spectrophotometer. To ensure that MSAD was rate-limiting, Cg10062(E114N) and ADH were used in excess. Cg10062(E114N) was used to generate MSA *in situ*. The assays were set up with all the components listed in Table 4.10, except ACA and MSAD. Assays were initiated by

addition of ACA, and MSAD was added after a delay of 30 s, to allow for MSA formation from ACA.

Table 4.10. Components used in the MSAD coupled enzyme assay.

Component	Volume (μL)	Final Concentration/Units
Cg10062(E114N) (30 mg mL^{-1} ; 5 U mg^{-1})	10	0.3 mg mL^{-1} (1.5 U)
MSAD (4 mg mL^{-1})	varies	$0.02 - 0.04 \text{ mg mL}^{-1}$
ADH (1 mg mL^{-1} ; 300 U mg^{-1})	100	0.1 mg mL^{-1} (30 U)
ACA (100 mM, pH 8)	10	1 mM
NADH (27 mM)	10	0.27 mM
100 mM sodium phosphate, pH 8	up to 1000 μL	-

4.15.7. ^1H NMR characterization of Cg10062 and variants

Cg10062 and variants were incubated with ACA and products were characterized ^1H NMR spectroscopy. NMR was used only for product identification and not quantitation. NMR spectra were obtained on a 500 MHz Varian NMR spectrophotometer and analyzed using MestreNova 14.2.0. (H)wet1D was used for solvent suppression since all assays were carried in aqueous buffer. Spectra were obtained using standard acquisition parameters which include the following: (H)wet1D spectral width of -2 to 14 ppm; transient number 64 and delay d_1 10 s. DMSO- d_6 (δ 2.49) was added to enable a lock signal and TSP (3-(trimethylsilyl)propionate-2,2,3,3- d_4 , sodium salt) (δ -0.21 (s, 9H)) was used as an internal standard. The resonances at δ 3.5 (s, 4H) corresponds to ethylene glycol, which was added for long-term storage of protein at -20°C .

Stock solutions were prepared in 100 mM sodium phosphate, pH 8. A 5 M solution of ACA, pH 8, was prepared in a 10.0 mL volumetric flask by dilution of ACA (3.5 g) with 2 mL of 100 mM sodium phosphate, pH 8. The pH was adjusted to 8 by addition of 10 N NaOH before adjusting the final volume to 10.0 mL. ACA (111 mM, 20 μL of 5 M stock) was added to 830 μL of 100 mM sodium phosphate, pH 8. Reactions were initiated by addition of enzyme (50 μL of a 4.8 mg mL^{-1}

stock) and incubated at 25 °C. Aliquots (150 μ L) were removed and quenched by addition of 5 M H_2SO_4 (2 μ L). Time points were taken immediately after initiation of the reaction (0 h) and 1 h after initiation of the reaction. Samples were microcentrifuged to remove precipitated protein and clarified aqueous sample (100 μ L) was combined with TSP (10 mM, 70 μ L of a 100 mM stock), and $\text{DMSO-}d_6$ (95% atom D, NMR grade) (30 μ L). The final volume was adjusted to 700 μ L by addition of 100 mM sodium phosphate, pH 8, for NMR spectroscopy.

The resonance at δ 2.91 (s, 1H) corresponds to ACA. Resonances at δ 9.50 (t, J = 2.9 Hz, 1H), δ 3.20 (d, J = 2.9 Hz 2H), and δ 5.13 (t, J = 5.5 Hz, 1H), δ 2.30 (d, J = 5.5 Hz, 2H) correspond to MSA and its hydrate, respectively. Resonances at δ 9.47 (q, J = 3.0 Hz, 1H), δ 2.03 (d, J = 3.0 Hz 3H), and δ 5.05 (q, J = 5.2 Hz, 1H), δ 1.12 (d, J = 5.5 Hz, 3H) correspond to acetaldehyde and its hydrate, respectively.

4.15.8. Protein crystallization and structure determination

All protein crystallization and structural determination of Cg10062 and variants were carried out by Dr. Hadi Nayebi (Geiger group) and Katelyn Silva (Draths/Geiger groups). Cg10062 and variants were purified and prepared as stock concentrations of 18 mg mL^{-1} in 10 mM Tris- SO_4 , pH 8 for crystallography experiments. Detailed conditions and experimental methods for protein crystallography were previously described.¹⁴

4.15.9. Growth of *E. coli* on ACA

E. coli RB791 *serA* was transformed with the plasmids pKM6.240 (control), pAS1.163 (expressing Cg10062), pAS1.240 (expressing Cg10062(E114Q)) and pAS1.241 (expressing Cg10062(E114D)) and the resulting cells were plated onto LB/Ap plates to obtain single colonies. The culturing experiments described here were carried out in duplicate. Four types of M9 media were prepared as described in Section 4.2: M9, M9/glu, M9/ACA and M9/glu/ACA. The media was supplemented with M9/ACA (pH 7.2) to a final concentration of 1 g L^{-1} . The pH adjusted M9/ACA was prepared by diluting 0.5 g ACA with M9 salts in a 100 mL volumetric flask. The pH was adjusted to 7.2 using 10 N NaOH and the sample was filtered through a 0.45 μm syringe filter

resulting in a final concentration of 5 g L⁻¹ ACA in M9 salts. To achieve a final concentration of 1 g L⁻¹ in the cultures, M9/ACA (10 mL) was added to M9 media (containing no glucose) (40 mL). IPTG (1 mM) was added to each media.

Single colonies of RB791 *serA*/pKM6.240, RB791 *serA*/pAS1.163, RB791 *serA*/pAS1.240 and RB791 *serA*/pAS1.241 were inoculated into M9 media described above. The cultures were incubated at 37 °C for 48 h and 1 mL of each culture was used to inoculate freshly prepared 5 mL media and incubated further for 48 h at 37 °C. The culturing was carried out for 12 cycles.

4.15.10. Growth of *E. coli* on ACA

Four types of M9 plates were prepared using low EEO agarose (M9, M9/glu, M9/ACA and M9/glu/ACA) using methods described in Section 4.2 with some modifications. A 2X M9 salts solution (1 L) was prepared by addition of Na₂HPO₄ (12 g), KH₂PO₄ (6 g), NH₄Cl (2 g) and NaCl (1 g) to water, followed by sterilization. The pH adjusted stock solutions of 2XM9/ACA was prepared by diluting 4 g of ACA with 2XM9 salts in a 100.0 mL volumetric flask. The pH was adjusted to 7.2 using 10 N NaOH and filtered through a 0.45 µm filter, resulting in a final concentration of 40 g L⁻¹. The low EEO agarose was autoclaved separately (1.5 g per 50 mL deionized water). Combining the low EEO agarose and 2X M9 salts ensured a final 1X concentration of the M9 salts in the plates. IPTG (1 mM) was added to each plate. The *serA* gene on each plasmid ensure plasmid maintenance without the need for antibiotic.

M9 plates with no carbon source were prepared by combining 2XM9 salts (50 mL) with low EEO agarose (50 mL). M9/glu plates were prepared by combining 2XM9 salts (50 mL), low EEO agarose (50 mL) and 20% (w/v) glucose (2 mL). M9/ACA plates were prepared by combining 2XM9 salts (40 mL), low EEO agarose (50 mL), and 2XM9/ACA (10 mL). M9/glu/ACA plates were prepared by combining 2XM9 salts (40 mL), low EEO agarose (50 mL), 20% (w/v) glucose (2 mL) and 2XM9/ACA (10 mL). Individual colonies of RB791 *serA*/pKM6.240, RB791 *serA*/pAS1.163,

RB791 *serA*/pAS1.240 and RB791 *serA*/pAS1.241 were replicate-plated onto freshly prepared plates. The plates were incubated at 37 °C for one week and observations were recorded.

4.15.11. Growth of *E. coli* expressing Cg10062 on aldehyde indicator plates

E. coli RB791 *serA* strains expressing Cg10062 and variants were grown on variations of M9 containing *para*-rosaniline²⁰ (*ros*): M9/*glu*, M9/*ACA*, M9/*glu/ros*, M9/*ACA/ros*, M9/*glu/ACA* and M9/*glu/ACA/ros*. The plates were prepared using methods described in Section 4.14.9 with some modifications. L-Serine (40 mg L⁻¹) was added to the media for plasmid maintenance. The plates were prepared the day prior to the replicate-plating and wrapped in foil.²⁰ NaHSO₃ (75 mg) was dissolved in up to 1 mL of deionized water in a 1 mL volumetric flask to prepare a stock solution (50 mg mL⁻¹). *p*-rosaniline (25 mg) was dissolved in up to 10 mL ethanol in a 10 mL volumetric flask to prepare a stock solution (2.5 mg mL⁻¹). The M9/*glu*, M9/*ACA* and M9/*glu/ACA* plates were prepared as described in Section 4.2. The M9/*glu/ros*, M9/*ACA/ros* and M9/*glu/ACA/ros* plates (100 mL) were further supplemented with the stock solutions of *p*-rosaniline (2 mL) and NaHSO₃ (0.5 mL).

Individual colonies of RB791 *serA*, RB791 *serA*/pKM6.240, RB791 *serA*/pAS1.163, RB791 *serA*/pAS1.240 and RB791 *serA*/pAS1.241 were replicate-plated onto each plate, wrapped in foil, and incubated at 37 °C. Observations from each plate were recorded over a period of one week.

4.16. Chapter Two: Synthesis of 3-Hydroxypropionic Acid

Reprinted with permission from Draths, K.; Geiger, J. H.; Mathes Hewage, A. N. S.; Gavvani, H. N. Synthesis of 3-Hydroxypropionic Acid via Hydration of Acetylenecarboxylic Acid. Patent pending, filed on April 21, 2022. Manuscript in preparation.

4.16.1. Genes and plasmids

Plasmid pAS2.141

The *ydfG* gene was amplified from *E. coli* W3110 genomic DNA using the following primers. The forward primer (5'–TGGTGCCGCGCGGCAGCCAAGAAAACCTGTATTTTCAGAT

GATCGTTTTAGTAACTGG–3’) includes a TEV protease recognition site (bold) at the 5’ end.¹¹ The reverse primer used was 5’–TCGGGCTTTGTTAGCAGCCGTTACTGACGGTGGACATTCAGTCC–3’. Regions homologous to the *E. coli* genome are underlined in both primers. pET-15b were digested with BamHI and NdeI. The *ydfG* amplicon was inserted into the digested pET-15b vector using a NEB® HiFi DNA assembly kit following the protocol described by the manufacturer. The resulting plasmid, pAS2.141 (Figure 4.8), was used to express YdfG with an N-terminal His₆ tag. Digestion of pAS2.141 with EcoRI and NcoI resulted in DNA fragments of 1.1 kb and 5.3 kb.

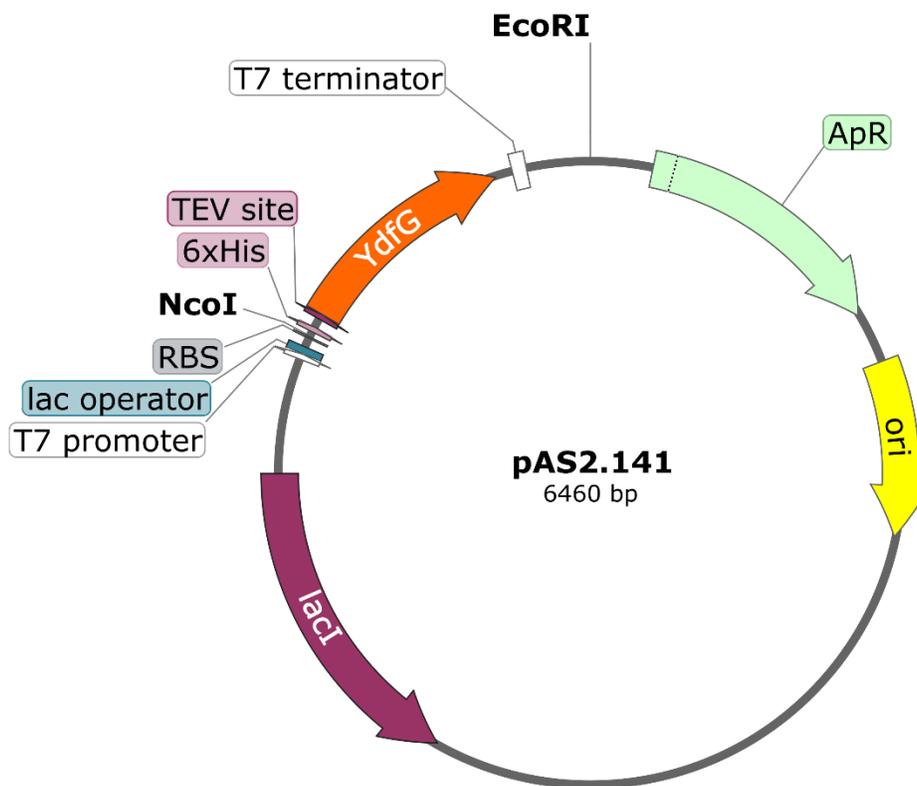


Figure 4.8. Plasmid map of pAS2.141

Plasmid pAS2.144

The *rutE* gene was amplified from *E. coli* W3110 genomic DNA using the same approach described for the pAS2.141 construct. The forward primer (5’–TGGTGCCGCGCGGCAGCCAAG **AAAACCTGTATTTTCAGATGAACGAAGCCGTTAGCCC**–3’) includes a TEV protease recognition site (bold) at the 5’ end.¹² The reverse primer used was 5’–TCGGGCTTTGTTAGCAG

CCGTTACAACAGCCCGCAG–3'. Regions homologous to the *E. coli* genome are underlined in both primers. The resulting plasmid pAS2.144 (Figure 4.9) enabled expression of RutE with an N-terminal His₆ tag. Digestion of pAS2.144 with EcoRI and NcoI resulted in DNA fragments of 1 kb and 5.3 kb.

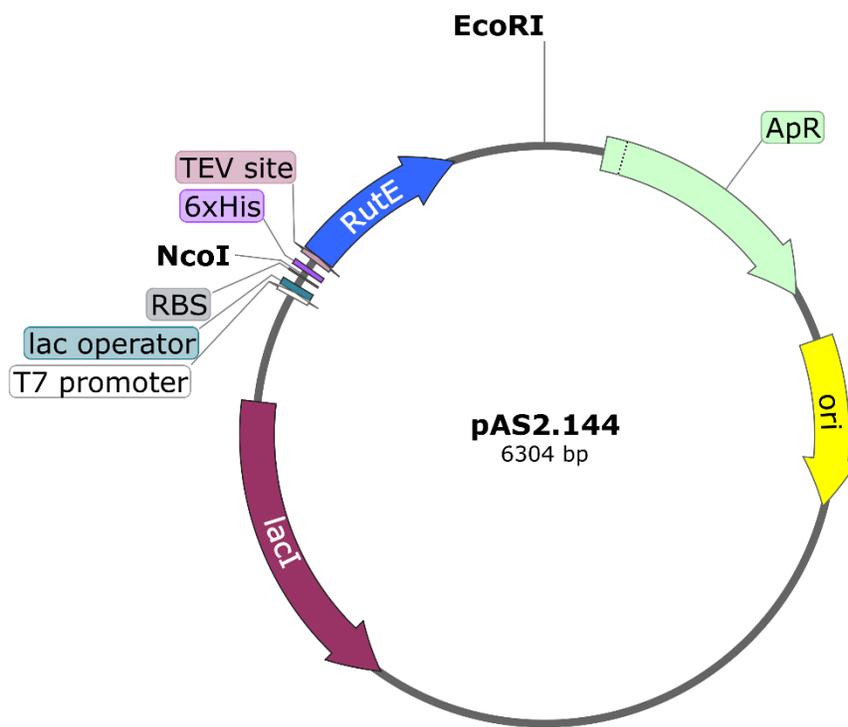


Figure 4.9. Plasmid map of pAS2.144.

Plasmid pAS2.084

The *ydfG* gene was amplified from *E. coli* W3110 genomic DNA using primers with NdeI (5'–GCGCATAATAC**CATATG**ATCGTTTTAGTAACTGGAGCAACGG–3') and XhoI (5'–GCATTAAAC**CTCGAG**CTGACGGTGGACATTCAGTCCG–3') restriction sites at the 5' and 3' positions, respectively. The restriction sites (bold) and regions homologous to the *E. coli* genome (underlined) are indicated in each primer. The amplicon was cloned into pET-21a(+) as described for pAS1.046 (Section 4.14.1) to enable expression of C-terminal His₆-tagged YdfG. Digestion of pAS2.084 (Figure 4.10) with NdeI and XhoI resulted in DNA fragments of 0.7 kb and 5.4 kb.

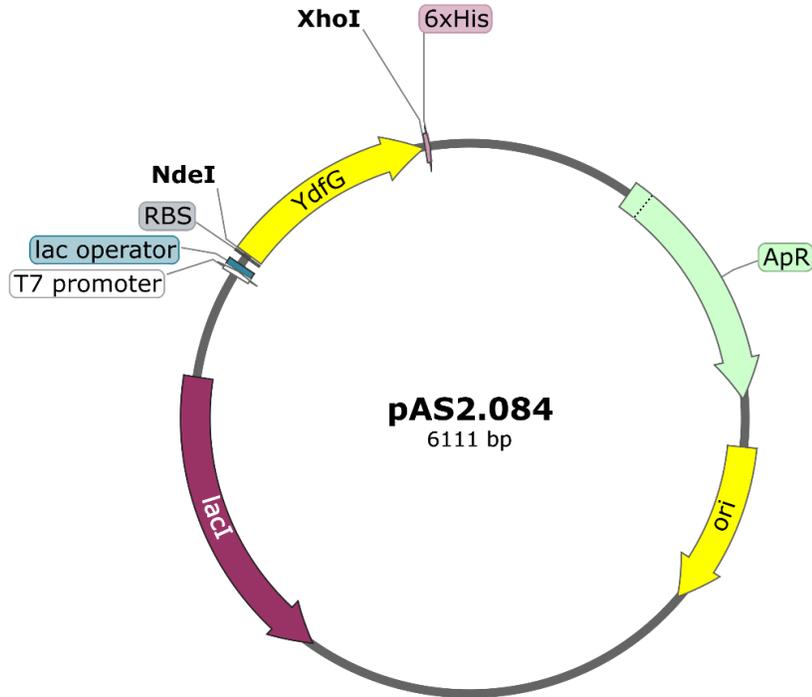


Figure 4.10. Plasmid map of pAS2.084

Plasmid pET15b-12x

The plasmid expressing the engineered phosphite dehydrogenase PTDH-12x Opt was purchased from Addgene (#61699) (Figure 4.11).²¹ Mutations D13E, M26I, V71I, E130K, Q132R, Q137R, I150F, E175A, Q215L, R275Q, L276Q, I313L, V315A, A319E, A325V, E332N and C336D were confirmed by sequencing. DNA sequencing also confirmed the *ptdh-12x opt* gene was cloned between NdeI and BamHI sites. Digestion of pET15b-12x with NdeI and BamHI resulted in DNA fragments of 1 kb and 5.7 kb.

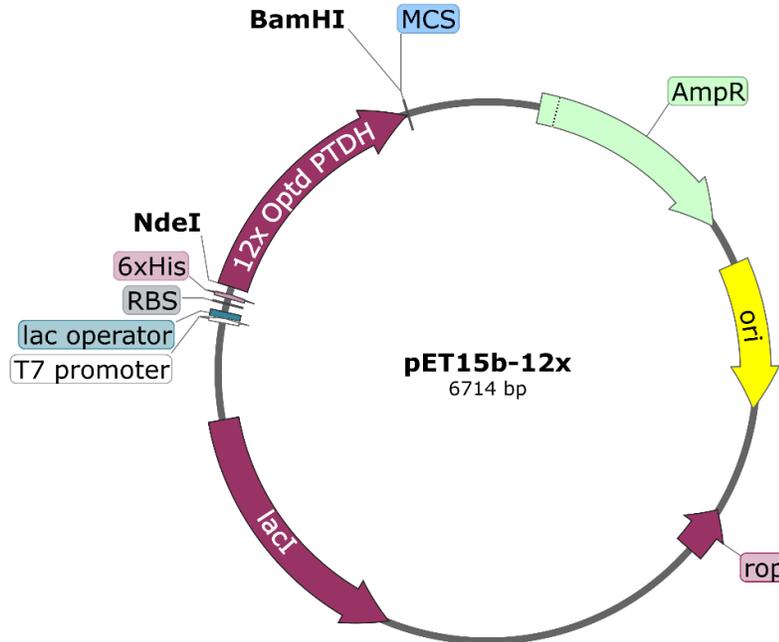


Figure 4.11. Plasmid map of pET15b-12x

Plasmid pAS3.112

For *in vivo* studies, the genes expressing Cg10062(E114N) and YdfG were cloned into pBbA1a-RFP, such that each gene was expressed from its own *trc* promoter optimally placed for gene expression. The plasmid (Figure 4.12) was constructed using NEB HiFi DNA Assembly. Forward primer (5'–AGATCTTTTAAGAAGGAGATATACATATGCCGACCTACACCTG–3') and reverse primer (5'–AATTGCGTTGCGCTCATTGCTGCTACCCTC–3') were used to amplify *cg10062(E114N)* from pAS2.100. Regions homologous to pAS2.100 are underlined. Forward primer (5'–AGGAGATATACATATGATCGTTTTAGTAACTGGAGC–3') was used with reverse primer (5'–TTATTTGATGCCTGGAGATCCTTACTCACTGACGGTGGACATTC–3') to amplify *ydfG* from pAS2.084. Regions homologous to pAS2.084 are underlined. Forward primer (5'–AGCAGCGAATGAGCGCAACGCAATTAATGTAAG–3') was used with reverse primer (5'–TAAAACGATCATATGTATATCTCCTTCTTAAAAGATCTTTTG–3') to amplify of *P_{trc}* from pBbA1a-RFP. Regions homologous to pBbA1a-RFP are underlined. Plasmid pBbA1a-RFP was digested with NdeI and XhoI. The digested vector was combined with the *cg10062(E114N)*, *ydfG* and *P_{trc}*

amplicons to construct pAS3.112 according to the standard protocol provided by the manufacturer.

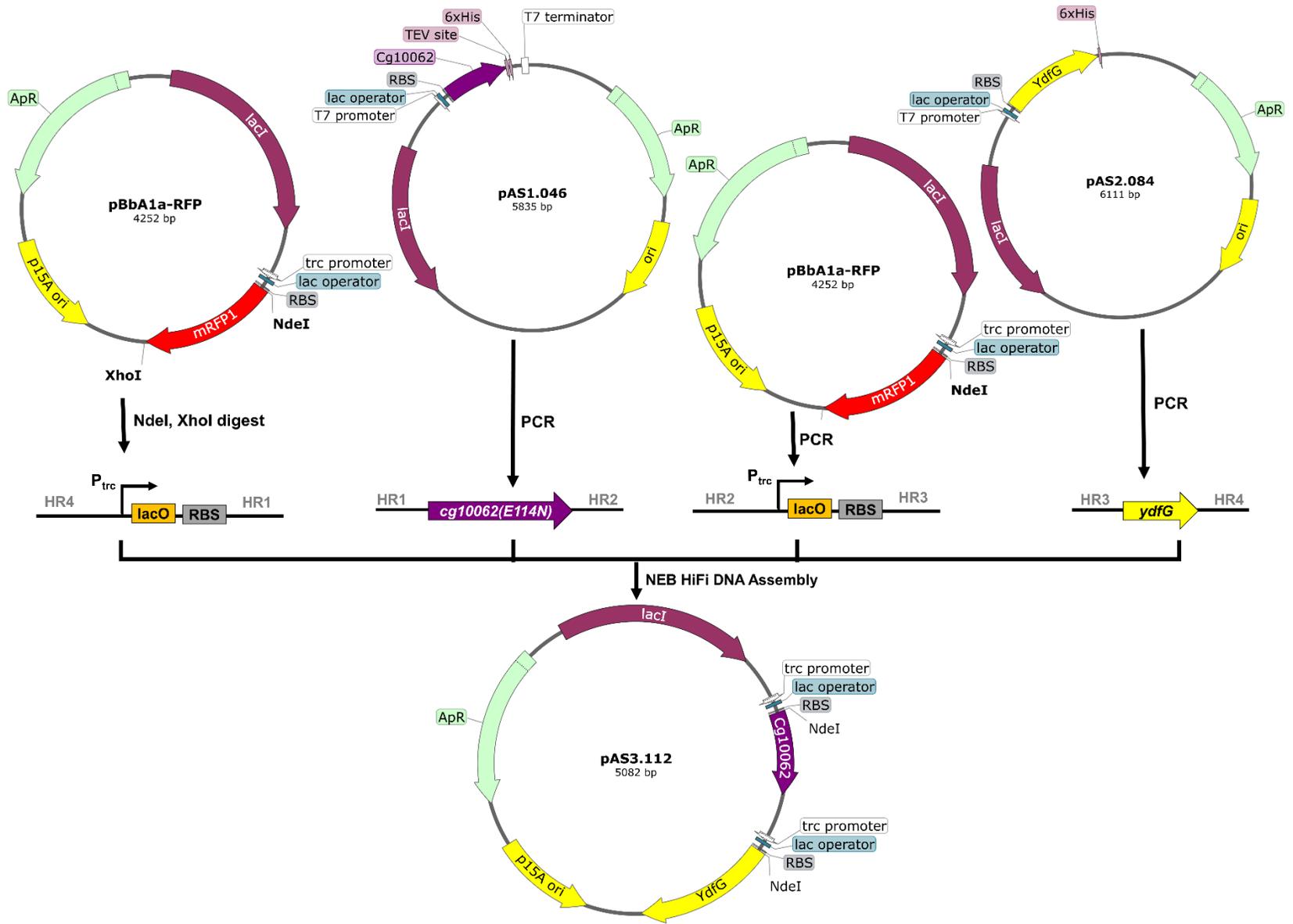


Figure 4.12. Construction of pAS3.112 using NEB HiFi DNA Assembly. (HR: homologous region)

4.16.2. Purification and characterization of YdfG and RutE

His₆-tagged YdfG and RutE were purified from BL21(DE3)/pAS2.141 and BL21(DE3)/pAS2.144, respectively. Cell culturing, protein expression and purification utilized procedures described previously (section 4.10.) Separately, FPLC fractions containing YdfG and RutE were combined and concentrated (200 μ L) in 50 mM sodium phosphate, pH 8, 0.5 mM EDTA and 1 mM DTT and reacted with TEV protease (YdfG/RutE: TEV molar ratio = 1:50) at 4 °C on a rocking platform (10 rpm) for 12 h. The enzyme mixtures were re-purified on a HisTrap FF 1 mL column and native YdfG and RutE were collected from the flowthrough. YdfG and RutE were concentrated in 100 mM sodium phosphate, pH 8. To initially screen for activity with MSA as a substrate, RutE and YdfG were characterized using the coupled enzyme assay shown in Figure 4.13. All assays were carried out in triplicate at 25 °C in 100 mM sodium phosphate, pH 8, in a final volume of 200 μ L, unless otherwise specified. All stock solutions prepared for the assays were prepared in 100 mM sodium phosphate, pH 8. A 10 mM stock solution of *cis*-3-chloroacrylic acid (cCA) was prepared by diluting 53 mg of cCA with 100 mM sodium phosphate, pH 8 up to 50.0 mL in a volumetric flask. The components used in the assay are listed in Table 4.11. and First, MSA was generated *in situ* from the *cis*-CaaD-catalyzed hydration of cCA, and the assays were initiated with the addition of YdfG/RutE. Since the standard conditions used in this study (growth at 37 °C, followed by 1 mM IPTG induction at 30 °C) resulted in poor expression of RutE, modified conditions (growth and 1 mM IPTG induction at 25 °C) were used to attempt improved expression.

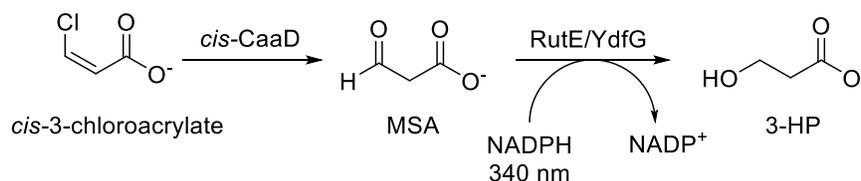


Figure 4.13. RutE/YdfG coupled enzyme assay for initial screening

Table 4.11. Components used in the RutE/YdfG coupled enzyme assay

Component	Volume (μL)	Final Concentration/Units
<i>cis</i> -CaaD (2 mg mL^{-1} , 7.5 U mg^{-1})	10	0.1 mg mL^{-1} (0.15 U)
RutE/YdfG (0.1 mg mL^{-1})	varied	$0.005 - 0.05 \text{ mg mL}^{-1}$
<i>c</i> CA (10 mM)	10	0.5 mM
NADPH (6 mM)	10	0.3 mM
100 mM sodium phosphate, pH 8	up to $200 \mu\text{L}$	-

4.16.3. YdfG coupled enzyme assay

YdfG was characterized using a coupled enzyme assay (Figure 4.14), where Cg10062(E114N) was used to generate MSA *in situ*, from the hydration of ACA. All assays were carried out in triplicate at $25 \text{ }^\circ\text{C}$ in 100 mM sodium phosphate, pH 8, in a final volume of $200 \mu\text{L}$, unless otherwise specified. All stock solutions prepared for the assays were prepared in 100 mM sodium phosphate, pH 8. All the components listed in Table 4.12 except ACA and YdfG, were combined in microfuge tubes and equilibrated for 15 mins at $25 \text{ }^\circ\text{C}$. The reactions were incubated with ACA, allowing *in situ* formation of MSA by Cg10062(E114N) (for 30 s) prior to initiation of assay by addition of YdfG. The specific activity of YdfG determined using the equation 4.1 when saturated with MSA.

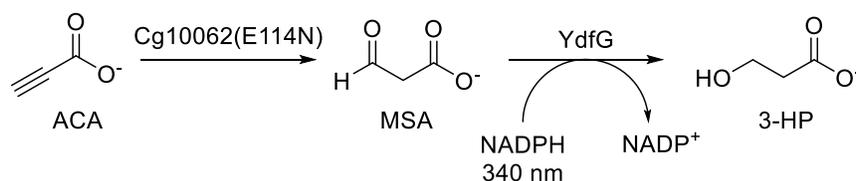
**Figure 4.14. YdfG coupled enzyme assay**

Table 4.12. Components used in the YdfG coupled enzyme assay

Component	Volume (μL)	Final Concentration/Units
Cg10062(E114N) (30 mg mL ⁻¹ , 5 U mg ⁻¹)	10	1.5 mg mL ⁻¹ (1.5 U)
YdfG (0.1 mg mL ⁻¹)	10	0.005 mg mL ⁻¹
ACA	Varied	10–2000 μM
NADPH (6 mM)	10	0.3 mM
100 mM sodium phosphate, pH 8	up to 200 μL	-

4.16.4. ¹H NMR characterization of YdfG

Cg10062(E114N) was used to generate MSA *in situ* from ACA hydration. (H)wet1D was used for solvent suppression since all assays were carried in aqueous buffer. All spectra were obtained using the following standard acquisition parameters: (H)wet1D with a spectral width of -2 to 14 ppm; 64 scans, d1 = 10 s. DMSO-*d*₆ (δ 2.49) was used as a lock signal and TSP (3-(trimethylsilyl)propionate-2,2,3,3-*d*₄, sodium salt) (δ -0.21 (s, 9H)) was used as an internal standard. To determine the products of YdfG-catalyzed reduction of MSA, ACA pH 8 (20 mM, 14 μL of a 1 M stock) was combined with YdfG (20 μL of a 6 mg mL⁻¹ stock), TSP (10 mM, 70 μL of a 100 mM stock), and DMSO-*d*₆ (95% atom D, NMR grade) (30 μL). The volume was adjusted to 680 μL by addition of 100 mM sodium phosphate, pH 8. The reaction was initiated with the addition of Cg10062(E114N) (0.15 U). ¹H NMR spectra were obtained after incubating the samples at 25 °C for 1 h. The resonance at δ 2.91 (s, 1H) corresponds to ACA. Resonances at δ 3.58 (t, *J* = 6.6 Hz, 2H) and δ 2.23 (t, *J* = 6.6 Hz, 2H) correspond to 3-HP.

4.16.5. Kinetic characterization of PTDH

PTDH activity was measured using the assay shown in Figure 4.15. All assays were carried out in triplicate at 25 °C in 100 mM sodium phosphate, pH 8 in a final volume of 200 μL , unless otherwise specified. All stock solutions were prepared in 100 mM sodium phosphate, pH 8. A 2 mM sodium phosphite stock solution was prepared by dissolving a 0.216 g of the solid in a

5 mL volumetric flask with water (pH adjustment was not necessary). The components of assay are listed in Table 4.14, and the assays were initiated with the addition of sodium phosphite. The specific activity (U mg⁻¹) of PTDH when saturated with sodium phosphite was calculated using the equation 4.1 (section 4.15.5). The Michaelis Menten kinetics of PTDH was determined using the same components listed in Table 4.13, where the sodium phosphite concentration ranged between 10–1000 μM.

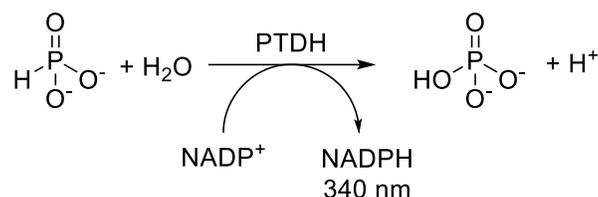


Figure 4.15. Enzyme assay used to determine measure PTDH activity

Table 4.13. Components used in the PTDH enzyme assay

Component	Volume (μL)	Final Concentration/Units
sodium phosphite (2 mM)	varied	10–1000 μM
PTDH (1 mg mL ⁻¹)	10	0.05 mg mL ⁻¹
NADP ⁺ (6 mM)	10	0.3 mM
100 mM sodium phosphate, pH 8	up to 200 μL	-

4.16.6. Buffer dependence of Cg10062(E114N), YdfG and PTDH

The activities of Cg10062(E114N), YdfG and PTDH were measured using four different buffers: 100 mM citrate-phosphate, 100 mM sodium phosphate, 100 mM Tris-Cl and 100 mM sodium carbonate/bicarbonate for pH ranges 3.6–5.6, 6.0–8.0, 7.0–9.2 and 9.2–9.6, respectively. All buffers were prepared according to standard protocols provided by the MilliporeSigma Buffer Reference Center.²¹

The pH dependence of each enzyme was studied using the respective enzyme assay used for kinetic characterization as described previously in this chapter. All pH studies were

carried out in triplicate (1 mL) on a Shimadzu UV2600 spectrophotometer at 25 °C to ensure that the final pH of each assay remained unchanged with the addition of assay components. All stock solutions were prepared in 100 mM sodium phosphate pH 8, unless otherwise specified, and the assays were carried out in the respective buffers for each pH. For Cg10062(E114N) and PTDH assays, all components except the substrates, ACA and sodium phosphate, respectively, were combined and prepared in 1 mL microfuge tubes and incubated at 25 °C for 15 min prior to addition of substrate. For the YdfG assays, all components except ACA and YdfG were combined in 1 mL microfuge tubes and incubated at 25 °C for 15 min. The assays were incubated with ACA and then initiated by addition of YdfG after a 30 s delay. The components used for each assay are listed in Table 4.14.

Table 4.14. Components used in pH dependence assays

Component	Volume (μL)	Final Concentration/Units
Cg10062(E114N) (1 mg mL⁻¹, 5 U mg⁻¹)	10	0.01 mg mL⁻¹(0.05 U)
MSAD (4 mg mL ⁻¹ , 30 U mg ⁻¹)	10	0.04 mg mL ⁻¹ (1.2 U)
ADH (1 mg mL ⁻¹ , 300 U mg ⁻¹)	40	0.04 mg mL ⁻¹ (12 U)
ACA (100 mM, pH 8)	10	1 mM
NADH (27 mM)	10	0.27 mM
Cg10062(E114N) (30 mg mL ⁻¹ , 5 U mg ⁻¹)	10	0.3 mg mL ⁻¹ (1.5 U)
YdfG (1 mg mL⁻¹, 20 U mg⁻¹)	10	0.01 mg mL⁻¹(0.2 U)
ACA (100 mM, pH 8)	10	1 mM
NADPH (24 mM)	10	0.24 mM
PTDH (5 mg mL⁻¹, 1 U mg⁻¹)	10	0.05 mg mL⁻¹ (0.05 U)
sodium phosphite (1 M)	10	0.01 mM
NADP ⁺ (24 mM)	10	0.24 mM

The enzyme being characterized in each assay is bolded.

4.16.7. *In vitro* synthesis of 3-HP with cofactor regeneration

All *in vitro* syntheses of 3-HP described in this study were carried out on a 1 mL scale. All stocks solutions were prepared in 100 mM sodium phosphate, pH 8. Ethylene glycol (20% w/v) was used as a cryoprotectant for all enzyme stock solutions. For reactions beginning with 100 mM ACA, three reactions containing 100 mM, 1 mM and 0.1 mM NADP⁺ (0.1 eq., 0.01 eq. and 0.001 eq.) were carried out. The components used for these three reactions are shown in Table 4.15. Preliminary tests using 1 molar equivalent of NADP⁺ indicated inhibition of Cg10062(E114N) and was not pursued further for *in vitro* 3-HP synthesis. Each reaction was carried out in duplicate at 25 °C with constant slow mixing on a Thermo Scientific digital compact rocking shaker (20 rpm). All reactions were initiated by the addition of ACA. Samples from each reaction were quenched at indicated timepoints for analysis by HPLC and ¹H NMR, as described in Sections 4.15.9 and 4.15.10. The conversion of 500 mM ACA to 3-HP was also demonstrated using the assay components as shown in Table 4.16.

Table 4.15. Conversion of 100 mM ACA to 3-HP with cofactor regeneration

Sample	ACA (mM)	NADP ⁺ (mM)	Na ₂ HPO ₃ (mM)	Cg10062(E114N) (U)	YdfG (U)	PTDH (U)
0.1 eq.	100	10	150	2	15	15
0.01 eq.	100	1	150	2	15	15
0.001 eq.	100	0.1	150	2	15	15

Table 4.16. Conversion of 500 mM ACA with cofactor regeneration

Sample	ACA (mM)	NADP ⁺ (mM)	Na ₂ HPO ₃ (mM)	Cg10062(E114N) (U)	YdfG (U)	PTDH (U)
0.1 eq.	500	50	550	2	15	15
0.01 eq.	500	5	550	2	15	15
0.001 eq.	500	0.5	550	2	15	15

4.16.8. *In vivo* synthesis of 3-HP in *E. coli*

To assess 3-HP production in rich media, a single colony of BL21/pAS3.112 was inoculated into LB (5 mL) and incubated at 37 °C with shaking (200 rpm) for 12 h. The overnight culture was used to inoculate two 25 mL LB (125 mL baffled flasks) to an initial OD₆₀₀ of 0.05 grown at 37 °C (200 rpm). When the OD₆₀₀ reached 0.5, IPTG (1 mM) was added to one culture. Then, ACA (5 M, pH 7.2) was added to both cultures to a final concentration of 100 mM. Both cultures were then grown at 30 °C for 8 h. Aliquots (100 µL) of each culture were removed and centrifuged immediately following addition of ACA (t = 0 h) and 9 h later. Cell-free broth was analyzed by ¹H NMR, using methods described below.

To assess the formation of 3-HP in minimal media, a single colony of BL21/pAS3.112 was inoculated into M9/glu (5 mL) and incubated at 37 °C with shaking (200 rpm) for 12 h. The overnight culture was used to inoculate two 25 mL M9/glu (125 mL baffled flasks) to an initial OD₆₀₀ of 0.05 grown at 37 °C (200 rpm). When the OD₆₀₀ reached 0.5, IPTG (1 mM) was added to one culture. The cultures were returned to the incubated shaker at 30 °C for 12 h. The cells were harvested by centrifugation (17,000 x g, 10 min) and resuspended in sterile M9 salts (10 mL) to remove residual glucose. This step was repeated two additional times, and the cells were resuspended in M9 salts containing ACA (100 mM, pH 7.2). To the cells previously treated with IPTG, IPTG was again added to a final concentration of 1 mM. Cultures were returned to incubation (30 °C) and shaking for 72 h. Cell-free broth from each culture was analyzed by ¹H NMR using methods described below.

4.16.9. HPLC quantification of ACA and 3-HP

For *in vitro* 3-HP synthesis, samples were quantified by RI detection on an Agilent 1100 HPLC equipped with an Aminex HPX-87H column (300 x 7.8 mm, 9 µm particle size) with isocratic elution of 0.01 N H₂SO₄ (0.6 mL min⁻¹) at 25 °C. Calibration curves for ACA (Figure 4.15) and 3-HP (Figure 4.16) were prepared using standard solutions described as follows. ACA (FW 70.04) was purchased from MilliporeSigma and purified via short-path distillation. An ACA solution (100

mM) was prepared by diluting 31 μL of ACA in 0.01 N H_2SO_4 in a volumetric flask (5 mL). Standard solutions of ACA (6.25, 12.5, 25, and 50 mM) were prepared by successive dilution of 100 mM ACA by addition of 0.01 N H_2SO_4 in a volumetric flask (1 mL). Standard solutions of 3-HP (FW 90.08) were prepared using a 30% (w/v) commercial solution (3.6 M) (CAS: 503-66-2) purchased from TCI chemicals. The concentration of the commercial solution was confirmed by ^1H NMR using TSP (10 mM, 70 μL of 100 mM stock), $\text{DMSO-}d_6$ (30 μL), 3-HP (10 μL of 30% (w/v) solution) in up to 700 μL of 100 mM sodium phosphate pH 8. For the preparation of a 500 mM 3-HP standard, 695 μL of 30% (w/v) 3-HP was diluted with 0.01 N H_2SO_4 in a volumetric flask (5 mL). Standard solutions of 3-HP (5, 10, 50, 100, 200, and 400 mM) were prepared by successive dilution of the 500 mM stock in volumetric flasks (1 mL) using 0.01 N H_2SO_4 .

Aliquots (100 μL) were removed from a reaction at the indicated time and quenched by addition of 5 M H_2SO_4 (2 μL) and the final volume of the sample was then adjusted to 500 μL by addition of 0.01 N H_2SO_4 . A 5-fold dilution (from 100 μL to 500 μL) afforded samples with 3-HP concentrations within the calibration curve. Samples were prepared for HPLC analysis using Whatman Mini-UniPrep® G2 syringeless filters with a glass microfiber membrane. The run time for each sample was 15 min. ACA and 3-HP eluted from the column at 11.5 min and 13.5 min, respectively.

4.16.10. ^1H NMR analysis of *in vitro* and *in vivo* 3-HP synthesis

For *in vitro* syntheses of 3-HP, ^1H NMR spectra were obtained at the beginning ($t = 0$ h) and end of each assay. (H)wet1D was used for solvent suppression since all assays were carried in aqueous buffer. All spectra were obtained using the following standard acquisition parameters: (H)wet1D with a spectral width of -2 to 14 ppm; 64 scans, $d_1 = 10$ s. $\text{DMSO-}d_6$ (δ 2.49) was used as a lock signal and TSP (3-(trimethylsilyl)propionate-2,2,3,3- d_4 , sodium salt) (δ -0.21 (s, 9H)) was used as an internal standard. Aliquots (150 μL) of each reaction were quenched by addition of 5 M H_2SO_4 (2 μL) and centrifuged, (17,000 $\times g$, 5 min) to remove precipitated protein. Clarified

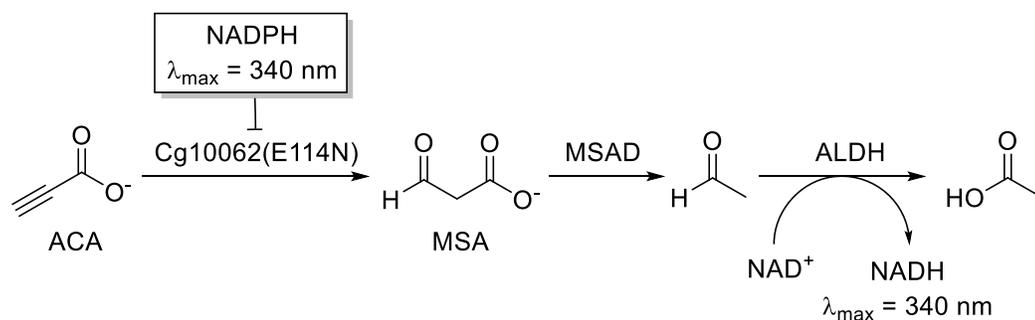
samples (100 μL) were combined with $\text{DMSO-}d_6$ (25 μL), TSP (10mM, 70 μL of a 100 mM stock), and volume adjusted to 700 μL by addition of 100 mM sodium phosphate, pH 8.

For *in vivo* 3-HP syntheses, ^1H NMR spectra were obtained at the beginning ($t = 0$ h) and end of each experiment. Aliquots of the cell culture supernatant (100 μL) were combined with $\text{DMSO-}d_6$ (30 μL), TSP (10mM, 70 μL of a 100 mM stock), and the volume was adjusted to 700 μL by addition of 100 mM sodium phosphate, pH 8.

For *in vivo* syntheses, the resonance at δ 2.91 (s, 1H) corresponds to ACA. Resonances at δ 3.58 (t, $J = 6.6$ Hz, 2H) and δ 2.23 (t, $J = 6.6$ Hz, 2H) correspond to 3-HP at pH 8. For *in vitro* syntheses, since samples were quenched with acid, the final pH of each NMR sample was less than 2, resulting in downfield-shifting of all resonances. The resonance at δ 3.27 (s, 1H) corresponds to ACA and resonances at δ 3.66 (t, $J = 6.6$ Hz, 2H) and δ 2.42 (t, $J = 6.6$ Hz, 2H) correspond to 3-HP.

4.16.11. Cg10062(E114N) inhibition by NADP(H)

The coupled enzyme assay used for kinetic characterization of Cg10062 and variants was modified by replacing the alcohol dehydrogenase (ADH) with aldehyde dehydrogenase (ALDH) (Figure 4.17). The PicoProbe™ NADH Fluorometric Assay Kit used to quantify the moles of NADH formed by ALDH in each sample. The NADH assay kit consisted of a fluorescent probe (PicoProbe™) and a cycling enzyme that can be used quantify NADH in a given sample.²² The quantification of NADH in the samples was carried out according to manufacturer's instructions.



Step 1: degrade NAD^+ at 60 C (30 min)

Step 2: Addition of PicoProbe and NADH cycling enzyme

Step 3: Quantify ex/em = 535/587 nm

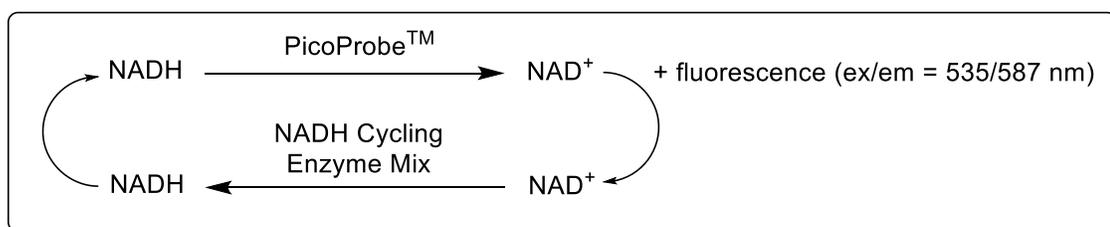


Figure 4.16. Modified coupled enzyme assay for monitoring Cg10062(E114N) inhibition by NADPH using PicoProbe™.

A calibration curve for NADH was prepared using the PicoProbe™ assay kit according to manufacturer's instructions as described below. Samples were prepared in a black solid bottom 96-well microplate. The NADH standard provided with the kit was reconstituted in 200 μL DMSO to prepare a 1 mM NADH standard. The NADH standard was diluted 100-fold by the addition of 990 μL of deionized water to 10 μL of 1 mM NADH standard. 10 μL of the resulting 10 μM NADH standard was diluted 100-fold by the addition of 990 μL deionized water. The final NADH standard solution (100 nM) was used to generate the calibration curve using the volumes shown in Table 4.17. The final volume of the NADH solutions were adjusted to 50 μL using NADH extraction buffer. A reaction mix (100 μL) containing NADH cycling buffer (96 μL), NADH cycling enzyme mix (2 μL) and PicoProbe™ (2 μL) was prepared and added to each well (100 μL) to ensure a final assay volume of 150 μL . The samples were incubated at 25 °C for 30 min and the fluorescence was measured (ex/em = 535/587 nm) using the Molecular Devices iD3 multi-mode microplate reader. Fluorescence emitted by the sample with no NADH was subtracted from all

other samples to account for background fluorescence. The resulting standard curve was used for quantification of NADH in the assays.

Table 4.17. PicoProbe™ components used for NADH standard curve preparation.

100 nM NADH (μL)	NADH extraction buffer (μL)	NADH _{final} (pmol)	[NADH] _{final} (nM)
0	50	0.0	0.0
4	46	0.4	2.7
8	42	0.8	5.3
12	38	1.2	8.0
16	34	1.6	10.7

The final reaction volume was adjusted to 150 μL by addition of a reaction mix (100 μL) containing NADH cycling buffer (96 μL), NADH cycling enzyme (2 μL) and PicoProbe (2 μL).

To determine the K_i of NADPH, four sets of samples (in triplicate) with varying concentrations of NADPH were set up for Michaelis Menten kinetics. A master mix containing all the assay components except 100 mM sodium phosphate pH 8 buffer (Table 4.18) was prepared for each set of assays. All assays had a final volume of 200 μL and were prepared in 0.5 mL microcentrifuge tubes and equilibrated at room temperature for 30 min prior to addition of ACA. The first set of samples contained varying ACA concentrations without any NADPH. The three remaining set of assays were identical to the first in terms of ACA concentrations but contained three different NADPH concentrations (10 mM, 20 mM and 30 mM). A stock solution of NADPH (100 mM) was prepared by dissolving 249 mg of NADPH tetrasodium salt ($\text{FW} = 833 \text{ g mol}^{-1}$) in up to 3 mL of 100 mM sodium phosphate pH 8 in a volumetric flask. All assays were initiated with the addition of ACA. Samples were quenched after 20 s using cold acetonitrile (ACN, 600 μL) and mixed thoroughly using a vortexer to ensure protein precipitation, resulting in a 4-fold dilution. All samples were incubated at 60 °C for 30 min to degrade NAD⁺ present in the samples and further diluted 15-fold by combining 10 μL of each sample with 140 μL of NADH extraction buffer. The resulting samples were further diluted 100-fold by combining 10 μL of each sample with 990 μL

of NADH extraction buffer. The dilution factor of the resulting samples relative to the initial assays was 6000. The samples (50 μL) were then incubated with the reaction mix (100 μL) containing NADH cycling enzyme (2 μL), PicoProbe™ (2 μL) and NADH cycling buffer (96 μL) in a 96-well black plate. The plate was sealed with foil and incubated at 25 °C for 30 min. Fluorescence was measured (ex/em = 535/587 nm) using the Molecular Devices SpectraMax iD3 multi-mode plate reader. Fluorescence emitted by a sample containing no NADH (triplicate) was subtracted from all other samples to account for background fluorescence.

Table 4.18. Components used in the Cg10062 coupled enzyme assay

Component	Volume (μL)	Final Concentration/Units
Cg10062)(E114N) (0.1 mg mL ⁻¹ ; 5 U mg ⁻¹)	20	0.01 mg mL ⁻¹ (0.01 U)
MSAD (4 mg mL ⁻¹ ; 30 U mg ⁻¹)	10	0.2 mg mL ⁻¹ (1.2 U)
ALDH (5 mg mL ⁻¹ ; 20 U mg ⁻¹)	20	0.1 mg mL ⁻¹ (2 U)
ACA (pH 8)	varies	mM
NAD ⁺ (6.7 mM)	10	0.34 mM
NADPH	varies	0, 10, 20 or 30 mM
100 mM sodium phosphate, pH 8	up to 200 μL	-

The IC₅₀ of NADPH on Cg10062(E114N) was determined using the same coupled enzyme assay (Figure 4.17) with 10 different concentrations of NADPH (0, 10, 20, 30, 40, 45, 50, 55, 60 and 65 mM) and a fixed concentration of ACA (500 μM). The concentrations of all other components listed in Table 4.18 remained unaltered. A control sample containing all components except NAD⁺ was used to account for background fluorescence. All assays had a final volume of 200 μL (triplicate) and followed the same methods described above for NADH quantification using the PicoProbe™ assay kit.

4.17. Chapter Three: 3-HP Biosensor for evolution of Cg10062(E114N)

This section describes the experimental methods used for the development of a biosensor which detects 3-HP for the evolution of Cg10062(E114N).

4.17.1. Genes and plasmids

Plasmid pAS5.003

The reporter plasmid pAS5.003 was constructed using an NEB HiFi DNA Assembly reaction. The *mmsR* gene (H681_13430)²³ was amplified from *Pseudomonas denitrificans* using primers AS090 and AS091 (Table 4.19). The *mmsA* promoter (*mmsR-mmsA* intergenic region: 130 bp)²³ was amplified from *P. denitrificans* using primers AS092 and AS093. The 105 bp DNA fragment AS094 containing the J23119 promoter²⁴ and RBS was synthesized by Integrated DNA Technologies (Figure 4.19). The vector pBbA1a-RFP was digested with BglII and Sall. The vector digest and PCR amplicons were purified using by gel extraction. The HiFi DNA assembly (Figure 4.20) was carried out using the recommended molar ratios (pBbA1a-RFP: AS094: *mmsR*: *mmsA* = 1: 7: 1: 7) using manufacturer's instructions.

5' – TTTCCCCGAAAAGTGCCACCTGACGTCGACTTGACAGCTAGCTCAGTCCTAGGT
ATAATGCTAGCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATAC – 3'

Figure 4.17. The DNA fragment containing promoter J23119. The promoter and RBS are highlighted in blue and yellow, respectively.

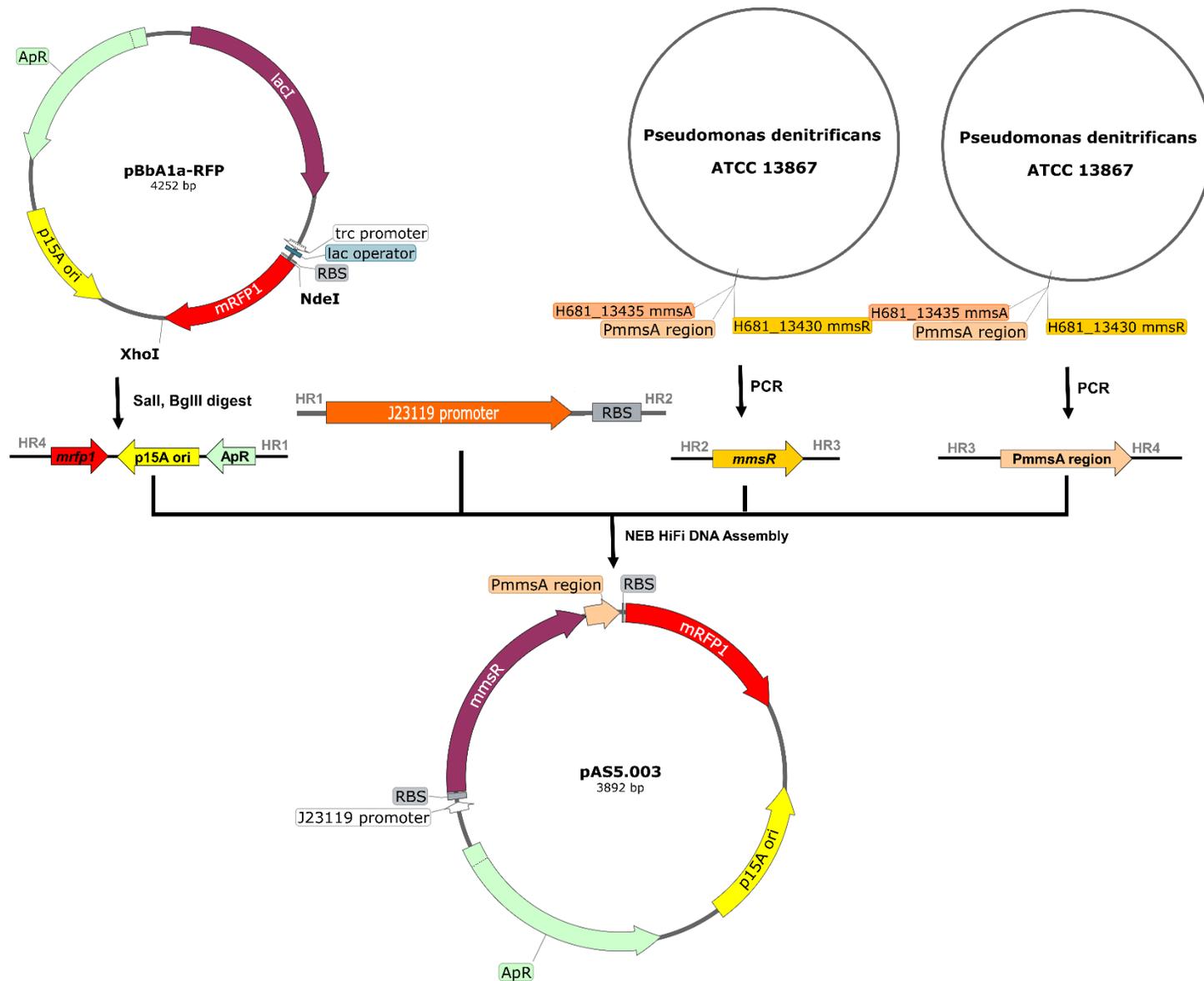


Figure 4.18. Construction of pAS5.003 using NEB HiFi DNA Assembly. (HR: homologous region)

Plasmid pAS5.005

Plasmid pAS5.005 was constructed using an NEB HiFi DNA Assembly reaction. The *cg10062* gene was amplified from pAS1.046 using primers AS116 and AS117 (Table 4.19). The *ydfg* gene was amplified from pAS2.084 using primers AS118 and AS121. A stop codon was introduced at the 3' end of both genes by inclusion in primers AS117 and AS121. The 149 bp DNA fragments J23119(YdfG) and J23102 containing the J23119 promoter and RBS for YdfG, and J23102 promoter for Cg10062(E114N) were synthesized by Integrated DNA Technologies (Figure 4.21). The vector pET-28a(+) was digested with BglII and EcoRV. The vector digest and PCR amplicons were purified using by gel extraction. The HiFi DNA assembly (Figure 4.22) was carried out using the recommended molar ratios (pET-28a(+): J23102: *cg10062*: J23119(YdfG): *ydfg* = 1: 7: 2: 7: 1) using manufacturer's instructions.

(A) 5' – TTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCG
TATCCCACTACCGAGATGATATC **TTGACAGCTAGCTCAGTCCTAGGTA****CTGTGCT**
AGCTCTAGAAATAATTTTGTTTAACTTTAAG **AAGGAG**ATATA– 3'

(B) 5' – CGTCTGACCGAACTGGAGGGTAGCAGCGAATGACTCGAGATCCGGCTGCT
AACAAAGCCCGAAAGGAAGCTGAGT **TTGACAGCTAGCTCAGTCCTAGGTA****TAAT**
GCTAGCTCTAGAAATAATTTTGTTTAACTTTAAG **AAGGAG**ATATA– 3'

Figure 4.19. The DNA fragments synthesized for pAS5.005. (A) fragment J23102 and (B) fragment J23119(YdfG). The promoters and RBS are highlighted in blue and yellow, respectively.

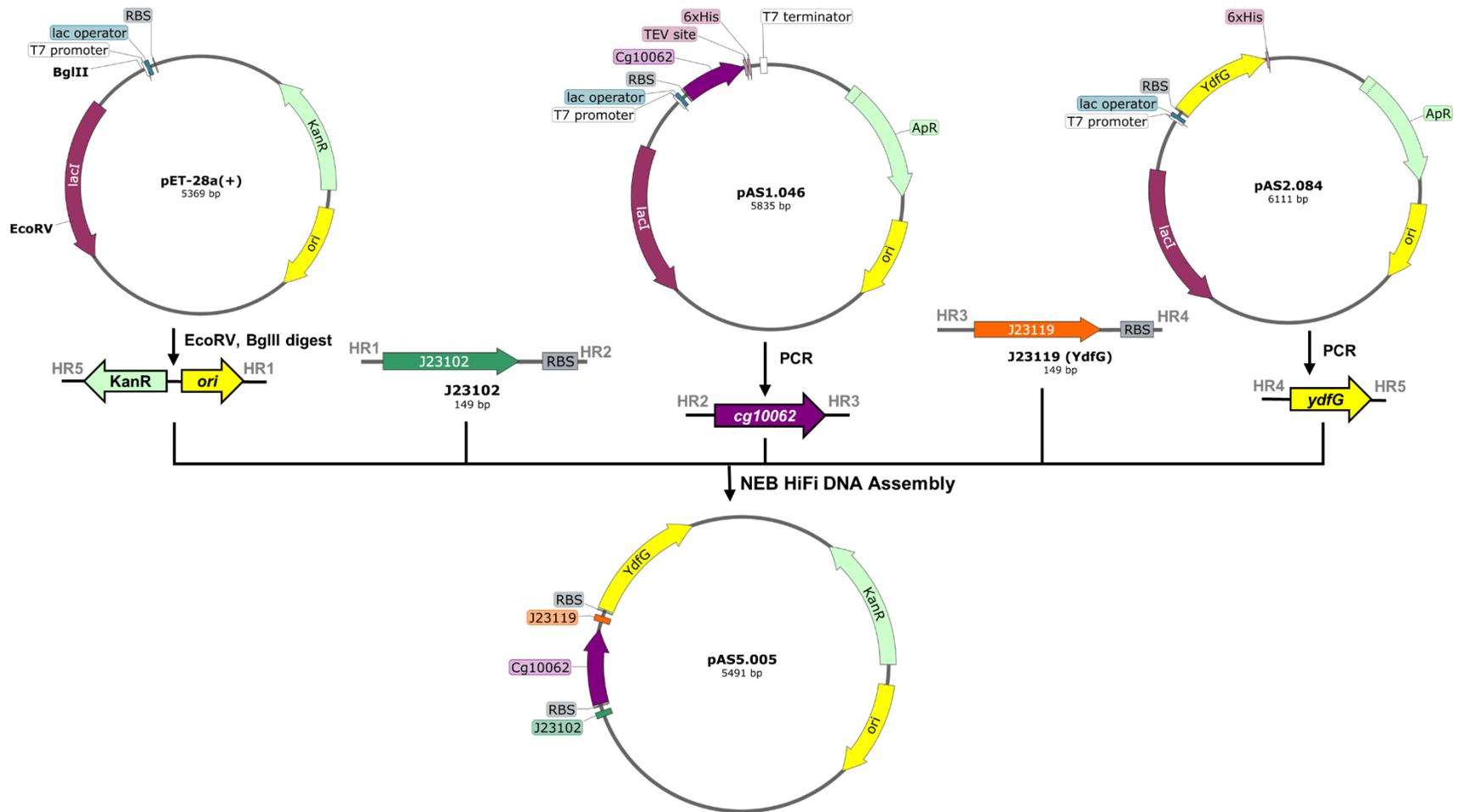


Figure 4.20. Construction of pAS5.005 using NEB HiFi DNA Assembly.

Plasmid pAS5.008

Plasmid pAS5.008 was constructed using the same methods described for pAS5.005 with some modifications. The *cg10062(E114N)* gene was amplified from pAS2.100 using primers AS116 and AS117 (Table 4.19), where a stop codon was introduced using the reverse primer. The assembly of plasmids pAS5.005 and pAS5.008 using the NEB HiFi DNA Assembly were unsuccessful. No transformants carrying the correct sequence for either plasmid was identified. Only a single transformant DH5 α carrying pAS5.008 with several mutations was isolated. The J23102 promoter had been inserted upstream of *ydfG* in place of the intended J23119.²⁴ The J23102 promoter upstream of *cg10062(E114N)* contained four deletions. Custom primers were used to construct pAS5.005 and pAS5.008 using the mutated plasmid via a second round of HiFi DNA Assembly (Figure 4.23). The *cg10062* gene was amplified from pAS1.046 where the correct J23102 sequence and *SpeI* restriction site, and a stop codon and *BglIII* site were inserted using primers AS132 and AS133, respectively. The PCR was repeated with pAS2.100 for amplification of *cg10062(E114N)*. Primers AS130 and AS131 were used to amplify the remaining backbone including the J23102 upstream of *ydfG*, *kan^R* and *ori*. For the final cloning of pAS5.005, the *cg10062* amplicon was incubated with the backbone vector according to manufacturer's instructions in a 2:1 molar ratio. The assembly was repeated with the *cg10062(E114N)* amplicon to obtain pAS5.008.

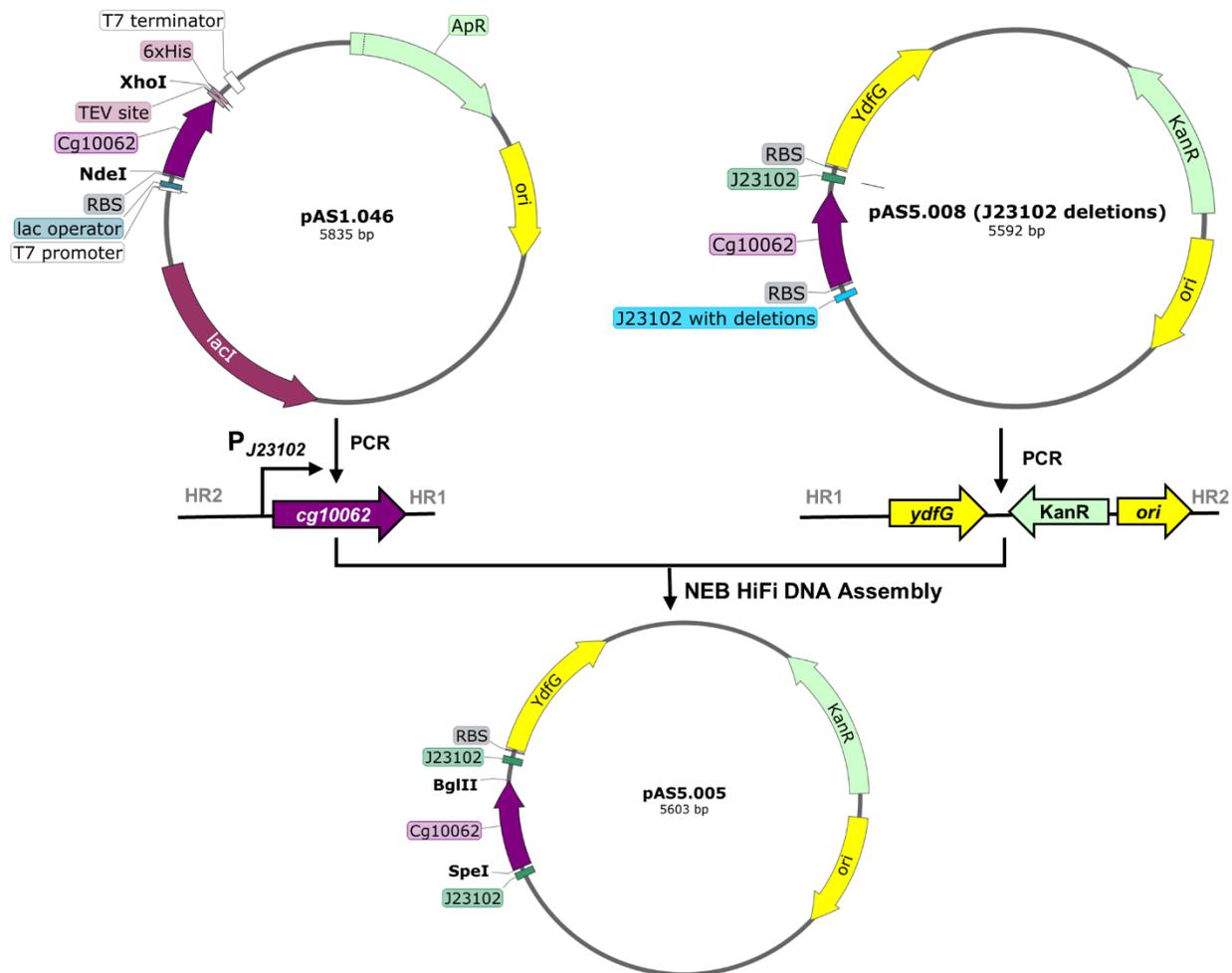


Figure 4.21. Final construction of pAS5.005 using NEB HiFi DNA Assembly

Table 4.19. Primers used in chapter three of this study

Primer	Sequence
AS090	ACTTTAAGAAGGAGATATACATGGACTGGGACAACCTG
AS091	TGAGGCTGACCTACATTCCCCGCCCATC
AS092	GGAATGTAGGTCAGCCTCAGCGCACCT
AS093	TGTATATCTCCTTCTTAAACGGGATTCTCCAGCCTG
AS116	AATTTTGTTTAACTTTAAGAAGGAGATATACATATGCCGACCTACACC
AS117	TCATTCGCTGCTACCCTCCAG
AS118	AATTTTGTTTAACTTTAAGAAGGAGATATACATATGATCGTTTTAGTAACTGG
AS121	AGTCGTATTAATTTTCGCGGGATCGATCACTGACGGTGGACATTCCAG

Table 4.19 (cont'd)

AS130	TCTTAGAAGTAAGCGACATGAGGGTAGCAGCGAATGAC
AS131	AGGACTGAGCTAGCTGTCAATCTCGGTAGTGGGATACGACG
AS132	TTGACAGCTAGCTCAGTCCTAGGTAAGTACTGTGCTAGCATATAGGAGGACTAGT ATGCCGACCTACACCTGC
AS133	CATGTCGCTTACTTCTAAGATCTTCATTCGCTGCTACCCTCCAG

4.17.2. Preparation of *E. coli* 3HPC, 3HPC-WT and 3HPC-E114N

Electrocompetent *E. coli* MG1655 were transformed with reporter plasmid pAS5.003 and plated on LB/Ap plates. The resulting transformants were used to prepare electrocompetent MG1655/pAS5.003 (designated *E. coli* 3HPC). Electrocompetent MG1655/pAS5.003 (3HPC) was then transformed with pAS5.005 (containing *cg10062* and *ydfG*) and the resulting MG1655/pAS5.003/pAS5.005, designated *E. coli* 3HPC-WT, plated onto LB/Ap/Kan plates. Similarly, the transformation of MG1655/pAS5.003 was repeated with pAS5.008 (containing *cg10062(E114N)* and *ydfG*) to obtain MG1655/pAS5.003/pAS5.008 (designated *E. coli* 3HPC-E114N).

4.17.3. Determination of RFP fluorescence as a function of IPTG concentration

A single colony of MG1655/pBbA1a-RFP was inoculated into LB/Ap (5 mL). Cells were grown in a shaker (200 rpm) at 37 °C for 12 h. The overnight culture was used to inoculate M9/glu (5 mL), to an initial OD₆₀₀ of 0.05. The cells were incubated at 37 °C with shaking until the OD₆₀₀ reached 0.5–0.7. Aliquots of the culture (190 µL) were transferred to a Corning® costar black clear-bottomed 96-well plate. IPTG was added to each well to a final concentration of 0, 0.01, 0.1 and 1 mM by addition of 0, 0.2, 2 and 20 mM IPTG stock solutions (10 µL). Following IPTG addition, each sample well contained identical sample volumes (200 µL). Each sample was prepared in triplicate. The plate was sealed with a MicroAmp™ Optical Adhesive film. Wells containing M9 salts (200 µL) were used as plate blank. OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 h.

4.17.4. Determination of RFP fluorescence as a function of 3-HP concentration

Single colonies of MG1655/pAS5.003 (3HPC), MG1655/pAS5.003/pAS5.005 (3HPC-WT) and MG1655/pAS5.003/pAS5.008 (3HPC-E114N) were inoculated into LB (5 mL). The 3HPC culture contained ampicillin. The 3HPC-WT and 3HPC-E114N cultures contained both ampicillin and kanamycin. The cultures were incubated in a shaker (200 rpm) at 37 °C for 12 h. The cells from the overnight cultures were pelleted by centrifugation (17,000 rpm, 1 min) and resuspended in M9 salts to remove residual rich media. The cells were pelleted once more, and the supernatant was discarded. The wash step was repeated twice more. The cells were then resuspended in 2 mL M9 salts and used to inoculate 10 mL M9 media containing glucose and antibiotic, to a final OD₆₀₀ of 0.1. Aliquots of the cultures (180 µL) were transferred to a Corning® costar 96-well black clear-bottomed plate sealed with a Genetix BreatheSeal film and incubated in a shaker (200 rpm) at 37 °C until the OD₆₀₀ reached 0.5–0.7. Since the cell growth was slow under these conditions, in later experiments, the cells were allowed to reach an OD₆₀₀ of 0.5–0.7 in the culture tubes and then transferred (180 µL) into the 96-well plates. A pH-adjusted 1 M stock solution of 3-HP was prepared in a volumetric flask (5 mL) by dilution of 1.39 mL of 30% (w/v) 3-HP in M9 salts. The pH was adjusted to 7 using 10 N NaOH. This was used to prepare 3-HP standards of 0, 5, 10, 25, 50, 100, 250, 500, 750 and 1000 mM concentrations in volumetric flasks (1 mL) using M9/glu media. The 3-HP standards (20 µL each) were then added to the wells to obtain final concentrations of 0, 0.5, 1, 2.5, 5, 10, 25, 50, 75 and 100 mM, respectively. Following addition of 3-HP, the final volume of each sample cell was identical (200 µL). Each well was prepared in triplicate with a total of 90 wells. The plate was sealed with a MicroAmp™ Optical Adhesive film and incubated in the shaker (200 rpm) at 37 °C. The OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 6, 12, and 24 h. Three wells filled with 200 µL M9 salts were used as the plate blank.

4.17.5. Determination of RFP fluorescence as a function of ACA concentration

The experiments described above for varying 3-HP concentrations were repeated for the same concentrations of ACA. The OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 h.

4.17.6. Determination of RFP fluorescence with increased 3-HP and ACA concentrations

The biosensor was also tested with higher concentrations of 3-HP and ACA relative to those mentioned above. The cell cultures were prepared using the same methods described above with some modifications. The cultures of 3HPC, 3HPC-WT and 3HPC-E114N were grown until an OD₆₀₀ of 0.5–0.7 and aliquots of the cultures were transferred (170 µL) to a Corning® costar black clear-bottomed 96-well plate. Stock solutions of 3-HP and ACA (2 M, pH 8) were prepared in volumetric flasks (5 mL) by dilution in M9 salts. The pH was adjusted to 7 using 10 N NaOH. This was used to prepare ACA and 3-HP standards of 0, 0.5, 0.66, 0.83, 1, 1.17, 1.33, 1.66, and 2.0 M concentrations in volumetric flasks (1 mL) using M9/glu media. The prepared standards (30 µL each) were added to the wells to obtain final concentrations of 0, 75, 100, 125, 150, 175, 200, 250 and 300 mM, respectively. Following addition of ACA or 3-HP, the final volume in each sample well was identical (200 µL). Each well was prepared in triplicate with a total of 81 wells each for ACA and 3-HP. The plate was sealed with a MicroAmp™ Optical Adhesive film and incubated in a shaker (200 rpm) at 37 °C. The OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 h. Three wells filled with 200 µL M9 salts were used as the plate blank.

4.17.7. ¹H NMR characterization of 3-HP biosensor

Single colonies of 3HPC, 3HPC-WT and 3HPC-E114N were inoculated into LB (5 mL). The 3HPC culture contained ampicillin. The 3HPC-WT and 3HPC-E114N cultures contained both ampicillin and kanamycin. The cultures were incubated in a shaker (200 rpm) at 37 °C for 12 h. The cells from the overnight cultures were pelleted by centrifugation (17,000 rpm, 1 min) and resuspended in M9 salts to remove residual rich media. The cells were pelleted once more and

supernatant discarded. The wash step was repeated twice more. The cells were then resuspended in 2 mL M9 salts and used to inoculate two 5 mL M9 media each containing glucose and antibiotic, to a final OD₆₀₀ of 0.1. One of the samples for each strain was supplemented with 50 mM ACA pH 8. The cultures were returned to the shaker (200 rpm) for 48 h. A sample (1 mL) was removed from each culture after 48 h. The cells were pellet by centrifugation (17,000 rpm , 1 min) and the supernatant (600 μ L) was combined with 100 mM TSP in 100 mM sodium phosphate pH 8 (10mM, 70 μ L of a 100 mM stock), DMSO-*d*₆ (95% atom D, NMR grade) (30 μ L). All spectra were obtained using the following standard acquisition parameters: (H)wet1D with a spectral width of -2 to 14 ppm; 64 scans, *d*₁ = 10 s. DMSO-*d*₆ (δ 2.49) was used as a lock signal and TSP (3-(trimethylsilyl)propionate-2,2,3,3-*d*₄, sodium salt) (δ -0.21 (s, 9H)) was used as an internal standard. The resonance at δ 2.91 (s, 1H) corresponds to ACA. Resonances at δ 3.58 (t, *J* = 6.6 Hz, 2H) and δ 2.23 (t, *J* = 6.6 Hz, 2H) correspond to 3-hydroxypropionate at pH 8.

REFERENCES

REFERENCES

- (1) Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual*; CSHL Press, 2001.
- (2) Miller, J. H. *Experiments in Molecular Genetics*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N. Y, 1972.
- (3) Keilhauer, C.; Eggeling, L.; Sahm, H. Isoleucine Synthesis in *Corynebacterium glutamicum*: Molecular Analysis of the *IlvB-IlvN-IlvC* Operon. *J. Bacteriol.* 1993, 175 (17), 5595–5603.
- (4) Studier, F. W.; Moffatt, B. A. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *J. Mol. Biol.* 1986, 189 (1), 113–130.
- (5) Gill, S. C.; von Hippel, P. H. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. *Anal. Biochem.* 1989, 182 (2), 319–326.
- (6) *Protein Calculator*. <http://protcalc.sourceforge.net/> (accessed 2022-04-21).
- (7) Frost, J. W.; Hansen, C. A. Synthesis of 1,2,3,4-Tetrahydroxybenzenes and 1,2,3-Trihydroxybenzenes Using Myo-Inositol-1-Phosphate Synthase and Myo-Inositol 2-Dehydrogenase. US6750049B1, June 15, 2004.
- (8) Li, K.; Mikola, M. R.; Draths, K. M.; Worden, R. M.; Frost, J. W. Fed-Batch Fermentor Synthesis of 3-Dehydroshikimic Acid Using Recombinant *Escherichia coli*. *Biotechnol. Bioeng.* 1999, 64 (1), 61–73.
- (9) Anderson, Jc.; Dueber, J. E.; Leguia, M.; Wu, G. C.; Goler, J. A.; Arkin, A. P.; Keasling, J. D. BglBricks: A Flexible Standard for Biological Part Assembly. *J. Biol. Eng.* 2010, 4 (1), 1.
- (10) Lee, T. S.; Krupa, R. A.; Zhang, F.; Hajimorad, M.; Holtz, W. J.; Prasad, N.; Lee, S. K.; Keasling, J. D. BglBrick Vectors and Datasheets: A Synthetic Biology Platform for Gene Expression. *J. Biol. Eng.* 2011, 5 (1), 12.
- (11) Kapust, R. B.; Tözsér, J.; Copeland, T. D.; Waugh, D. S. The P1' Specificity of Tobacco Etch Virus Protease. *Biochem. Biophys. Res. Commun.* 2002, 294 (5), 949–955.
- (12) Miller, Kelly. Semi-Synthesis of Aromatic Diacids and Biosynthesis of Kanosamine in *Escherichia coli*. Ph.D. Michigan State University - United States.
- (13) Johannes, T. W.; Woodyer, R. D.; Zhao, H. Efficient Regeneration of NADPH Using an Engineered Phosphite Dehydrogenase. *Biotechnol. Bioeng.* 2007, 96 (1), 18–26.
- (14) Mathes Hewage, A.; Nayebi Gavgani, H.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochemistry* 2021, 60 (51), 3879–3886.

- (15) *Synthetic construct tautomerase (Cg10062) gene, complete cds - Nucleotide - NCBI.* <https://www.ncbi.nlm.nih.gov/nuccore/MZ369159.1> (accessed 2022-08-03).
- (16) Guo, Y.; Serrano, H.; Poelarends, G. J.; Johnson, W. H.; Hackert, M. L.; Whitman, C. P. Kinetic, Mutational, and Structural Analysis of Malonate Semialdehyde Decarboxylase from *Coryneform* bacterium Strain FG41: Mechanistic Implications for the Decarboxylase and Hydratase Activities. *Biochemistry* 2013, 52 (28).
- (17) Blommel, P. G.; Fox, B. G. A Combined Approach to Improving Large-Scale Production of Tobacco Etch Virus Protease. *Protein Expr. Purif.* 2007, 55 (1), 53–68.
- (18) Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, W. H.; Whitman, C. P. Characterization of Cg10062 from *Corynebacterium Glutamicum*: Implications for the Evolution of *cis*-3-Chloroacrylic Acid Dehalogenase Activity in the Tautomerase Superfamily. *Biochemistry* 2008, 47 (31), 8139–8147.
- (19) Poelarends, G. J.; Johnson, W. H.; Murzin, A. G.; Whitman, C. P. Mechanistic Characterization of a Bacterial Malonate Semialdehyde Decarboxylase: Identification of a New Activity In The Tautomerase Superfamily. *J. Biol. Chem.* 2003, 278 (49), 48674–48683.
- (20) Conway, T.; Sewell, G. W.; Osman, Y. A.; Ingram, L. O. Cloning and Sequencing of the Alcohol Dehydrogenase II Gene from *Zymomonas mobilis*. *J. Bacteriol.* 1987, 169 (6), 2591–2597.
- (21) *Buffer Reference Center.* <https://www.sigmaaldrich.com/US/en/technical-documents/protocol/protein-biology/protein-concentration-and-buffer-exchange/buffer-reference-center#citric> (accessed 2022-08-19).
- (22) *PicoProbe™ NADH Fluorometric Assay Kit | K338 | BioVision, Inc.* <https://www.biovision.com/picoprobe™-nadh-fluorometric-assay-kit.html> (accessed 2022-05-25).
- (23) Nguyen, N. H.; Kim, J.-R.; Park, S. Development of Biosensor for 3-Hydroxypropionic Acid. *Biotechnol. Bioprocess Eng.* 2019, 24 (1), 109–118.
- (24) *Promoters/Catalog/Anderson - parts.igem.org.* <http://parts.igem.org/Promoters/Catalog/Anderson> (accessed 2022-06-30).