

SYNTHESIS OF COMMODITY CHEMICALS FROM CH₄ AND CO₂-DERIVED
ACETYLENECARBOXYLIC ACID: EXPLORATION OF TAUTOMERASE CATALYSIS

By

Katelyn Leigh Schwager Silva

A DISSERTATION

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

Chemistry—Doctor of Philosophy

2025

ABSTRACT

As an alternative to petroleum and biomass derived starting materials, this dissertation proposes utilization of CH₄ and CO₂-derived acetylenecarboxylic acid (ACA) as feedstock for commodity chemical production. These commercial building block chemicals such as 3-hydroxypropionic acid are being targeted for production via biocatalytic methods over petroleum and biomass derived methods. We look to biotechnology as an alternative route to produce these valued chemicals through enzyme biocatalysis. Here we evaluate the tautomerase enzyme *cis*-CaaD, which almost exclusively converts ACA to intermediate malonic semialdehyde (MSA). MSA can then be further be enzymatically converted to high priority target chemicals. We also compare *cis*-CaaD to the previously evaluated tautomerase Cg10062, which largely decarboxylates to acetaldehyde and produces minimal MSA. In addition, although *cis*-CaaD produces predominantly MSA, Cg10062 is more catalytically efficient. In our efforts to construct a highly catalytic MSA-forming *cis*-CaaD enzyme we used kinetic studies and protein crystallography to understand the *cis*-CaaD mechanism and gain a deeper understanding for the different product profiles these two enzymes produce. Previous efforts have shown the *in vitro* production of 3-hydroxypropionic acid from ACA. Here we look at *in vivo* production. Earlier experimental trials have shown that *E. coli* as well as many other bacterial species are not viable with ACA as a sole carbon source. However, the use of a tautomerase modified fusion protein allows the conversion of ACA to MSA in media followed by the uptake of MSA by *E. coli* cells. This will allow ongoing *in vivo* biocatalytic conversion of MSA to high priority chemicals.

This dissertation is dedicated to my sister, Jaclyn. Remember that you are capable of incredible things. I love you very much.

ACKNOWLEDGEMENTS

I would like to thank my advisors Dr. Karen Draths and Dr. Jim Geiger. I am incredible grateful for the opportunity to have learned from you both over these years. I have been able to have learned broad skill sets from having both of you as mentors that I never would have thought possible. The amount of knowledge I have accumulated under your guidance will carry me throughout my career. I am also eternally grateful for not only the academic support but the personal support that you have provided. Your kindness and humanity will never be forgotten. I would also like to thank my committee members Dr. Heedeok Hong and Dr. David Weliky for their guidance throughout this doctorate program.

I am grateful for the colleagues and friends I have met in this program. I would like to express my sincere thanks for the advice from Dr. Amaya Sirinimal, Courtney Bingham, Dr. Hadi Nayebi, and Dr. Bismarck Amaniampong during my years in the program. Thank you for passing down your knowledge to me. I would also like to thank Dr. Katie Kwiatkowski, Dr. Yasheen Jadidi, Nada Al-Ahmad, Noor Saber, Sophia Rose, and Esther Lee for your assistance.

Finally, I would like to thank my family for their support. I'm grateful for my mother, Evelyn, who encourages me to do what is right and is always willing to come running when I need her. I'm thankful to my father, Lee, who taught me a strong work ethic and instilled my ambitious nature. Thank you to my grandparents, who crossed mountains, volcanoes, rivers, and oceans so that I might have opportunities such as this. And to my sister, Jaclyn, thank you for being sunshine in the dark. Thank you to Juniper and Winona for the joy you have given me. There are many others who have helped me on my journey, and they are all so deeply appreciated. Most importantly, thank you God.

TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF TABLES.....	xiii
LIST OF ABBREVIATIONS	xv
CHAPTER ONE: New Strategies for Commodity Chemical Production.....	1
1.1. Caveats of Current Production of Commodity Chemicals	2
1.2 The Tautomerase Superfamily.....	16
1.3 <i>cis</i> -CaaD and Cg10062 Function and Purpose	29
REFERENCES	33
CHAPTER TWO: Comparison of <i>cis</i> -CaaD and Cg10062 Catalyzed Transformation of Acetylenecarboxylic Acid	42
2.1. <i>cis</i> -CaaD Engineering.....	43
2.2. <i>cis</i> -CaaD Mechanism of Action.....	53
2.3 Cg10062 (E114N) Engineering.....	67
2.4 Cg10062 Mechanism.....	69
2.5 α -Carbon Shift Leads to Product Change	80
2.6 C-Terminal is Necessary for Activity	86
REFERENCES	91
CHAPTER THREE: Application of Cg10062 (E114N) to 3-Hydroxypropionic Acid Synthesis....	93
3.1. Cg10062 (E114N) Use in the Production of Commodity Chemicals	94
3.2 Short Term Adaptive Laboratory Evolution (STALE): <i>E. coli</i> Growth on Medias.....	96
3.3 A Novel Cg10062 (E114N) Fusion Protein	100
REFERENCES	106
CHAPTER FOUR: Experimental	108
4.1 Materials	109
4.2 Culture Media and Stock Solutions	110
4.3 Preparation and Transformation of Electro- and Chemically Competent <i>E. coli</i>	110
4.4 Isolation of Plasmid DNA.....	112
4.5 Restriction Digestion of DNA	113
4.6 DNA Sequencing	114
4.7 PCR Amplification	114
4.8 Protein Expression	115
4.9 Protein Purification	117
4.10 Protein Quantification	119
4.11 Short-Path Distillation of Acetylenecarboxylic Acid.....	120
4.12 Chapter Two: Comparison of <i>cis</i> -CaaD and Cg10062 Transformation of Acetylenecarboxylic Acid	121
4.13 Chapter Three: Application of Cg10062 (E114N) to 3-Hydroxypropionic Acid Synthesis	194
REFERENCES	206

LIST OF FIGURES

Figure 1.1. Production of acrylonitrile, acrylic acid, and acrylamide from propylene.....	4
Figure 1.2. Petroleum, biomass, and CH ₄ and CO ₂ feedstocks for chemical production.	5
Figure 1.3. Scheme for the production of ACA.	6
Figure 1.4. Pathway from greenhouse gasses CO ₂ and CH ₄ via substrate ACA using enzymes <i>cis</i> -CaaD and Cg10062 to produce MSA and further form high value chemicals.	7
Figure 1.5. BASF synthesis of acetylene.....	7
Figure 1.6. Plasma arc production of acetylene.	8
Figure 1.7. Production of ACA from acetylene.....	8
Figure 1.8. High value chemicals from 3HP.....	10
Figure 1.9. 3HP synthesis from glycerol via CoA-dependent and independent pathways.	12
Figure 1.10. 3HP synthesis from glucose via the malonyl-CoA pathway.....	14
Figure 1.11. Synthesis of 3HP from glucose via the β -alanine pathway.....	14
Figure 1.12. The tautomerase superfamily sequence similarity network. The gray subgroup indicates sequences from the tautomerase superfamily that have not yet been assigned to a specific subgroup.	17
Figure 1.13. Mechanism of 4-OT utilization of 2-hydroxymuconate as a substrate for the tautomerization to 2-oxo-3-hexenedionate and 2-oxo-4-hexenedionate.....	19
Figure 1.14. Tautomerase family enzymes (A) 4-OT from <i>B. lata</i> (PDB: 6OGM), (B) CHMI from <i>E. coli</i> (PDB: 1OTG) and (C) MIF from <i>P. yoelii</i> (PDB: 3GAD) colored by chain.	23
Figure 1.15. Tautomerase family enzyme MSAD from <i>C. bacterium</i> (PDB:3MJZ) (A) with crystal structure colored by chains and (B) active site residues.....	24
Figure 1.16. 1,3-Dichloropropene degradation pathway in <i>P. pavonacea</i>	25
Figure 1.17. Tautomerase family enzyme CaaD from <i>P. pavonacea</i> (PDB: 3EJ3) colored by chains.....	26
Figure 1.18. Previously understood <i>cis</i> -CaaD mechanism.....	27
Figure 1.19. <i>cis</i> -CaaD (green, PDB: 9NFH) and Cg10062 (pink, PDB: 3N4G) crystal structure (A) trimers overlayed and (B) showing active site residues Pro-1, His-28, Arg-70, Arg-73, Tyr 103, and Glu-114.....	29
Figure 1.20. Scheme of the formation of MSA or ACH by <i>cis</i> -CaaD and Cg10062 and their wild-type product profiles.	30

Figure 1.21. <i>cis</i> -CaaD (green, PDB: 9NFH) and Cg10062 (pink, PDB: 3N4G) crystal structures showing active site residues Pro-1, His-28, Arg-70, Arg-73, Tyr-103 and Glu-114.	31
Figure 2.1. Coupled enzyme assay used in the evaluation of <i>cis</i> -CaaD and <i>cis</i> -CaaD variant activities.....	45
Figure 2.2. Product ratios of variants and <i>cis</i> -CaaD with (A) values shown in a table, (B) graphed product ratios for ACA substrate and (C) CCA substrate. (Those signified with – indicate too low activity to determine product ratios)	46
Figure 2.3. Steady state kinetics of (A) <i>cis</i> -CaaD (wild-type) with substrate ACA, (B) <i>cis</i> -CaaD (wild-type) with substrate CCA, (C) Cg10062 (wild-type) with ACA, (D) Cg10062 (wild-type) with CCA, (E) <i>cis</i> -CaaD (T32A) with CCA, (F) <i>cis</i> -CaaD (T34A) with ACA, and (G) <i>cis</i> -CaaD (T34A) with CCA.	48
Figure 2.4. Specific activities (A) of <i>cis</i> -CaaD (wild-type) and its variants with Cg10062 (wild-type) using substrates CCA and ACA. Differing substrate concentrations (B) were used to ensure saturation of substrate to enzyme.	50
Figure 2.5. ¹ H NMR of product MSA and hydrate formation by wild-type <i>cis</i> -CaaD with substrates (A) ACA and (B) CCA with time at 0 hours, 15 minutes, 30 minutes, 1 hour, 2 hours, and 16 hours. Other enzymes were evaluated similarly using ¹ H NMR (not shown).	52
Figure 2.6. Crystal structure of (A) apo <i>cis</i> -CaaD (E114D) and (B) apo <i>cis</i> -CaaD (E114N) with sulfate in the active site and (C) apo <i>cis</i> -CaaD (E114Q) with phosphate in the active site.	54
Figure 2.7. Crystal structure of apo <i>cis</i> -CaaD (Y103F) with sulfate in the active site (A). Crystal structure of apo <i>cis</i> -CaaD (T34A) with sulfate in the active site (B).	55
Figure 2.8. Crystal structures of ACA soaked <i>cis</i> -CaaD (E114D) at 2.5 Å showing (A) unreacted substrate and (B) covalent intermediate of 3-(<i>N</i> -prolyl)-acrylate.	56
Figure 2.9. Crystal structure of shifted orientation of (A) ACA and (B) MSA product in the active site of <i>cis</i> -CaaD (E114N) is shown at 2.3 Å.	56
Figure 2.10. Crystal structures of ACA soaked <i>cis</i> -CaaD (E114Q) at 1.9 Å crystals showing covalent intermediate of 3-(<i>N</i> -prolyl)-3-hydroxypropionate linked to catalytic Pro-1(A).	57
Figure 2.11. Crystal structures of ACA soaked <i>cis</i> -CaaD (H28A) with (A) acetate and (B) water in the active site. Crystal structures of ACA soaked <i>cis</i> -CaaD (Y103F) with (C) sulfate in the active site.	58
Figure 2.12. Crystal structures of ACA soaked <i>cis</i> -CaaD (T34A) with sulfate in the active site.	59
Figure 2.13. Crystal structure of wild-type <i>cis</i> -CaaD at 1.3 Å (A) with an acetate in the active site, (B) soaked with ACA substrate showing an unreacted ACA in the active site and 2.0 Å, and (C) the two aligned.	60
Figure 2.14. Crystal structures of CCA soaked <i>cis</i> -CaaD (E114D) at 1.8 Å showing (A) the covalent intermediate of (<i>N</i> -prolyl)-ethene linked to catalytic Pro-1 with sulfate and (B) sulfate alone.....	61

Figure 2.15 Crystal structures of CCA soaked <i>cis</i> -CaaD (E114Q) at 2.2 Å crystals showing (A) covalent intermediate of 3-(<i>N</i> -prolyl)-3-hydroxypropionate linked to catalytic Pro-1, (B) covalent intermediate of 3-(<i>N</i> -prolyl)-acrylate linked to catalytic Pro-1 and (C) (<i>N</i> -prolyl)-ethene covalent intermediate with sulfate.	62
Figure 2.16 Crystal structures of CCA soaked <i>cis</i> -CaaD (Y103F) with (A) three waters in the active site. Crystal structures of CCA soaked <i>cis</i> -CaaD (T34A) with (B) sulfate in the active site.	63
Figure 2.17 Mechanism of <i>cis</i> -CaaD (A) without and (B) with crystal structures showing intermediates trapped in the active site.	65
Figure 2.18. Mechanism of Cg10062.	71
Figure 2.19. Cg10062 (E114N) is shown with (A) trapped ACA in the active site, (B) 3-(<i>N</i> -prolyl)-acrylate (ACR) intermediate, (C) 3-(<i>N</i> -prolyl)-3-hydroxypropionate (HPA) intermediate, and (D) 3-(<i>N</i> -prolyl)-acrylate and ACA aligned. The acrylate (E) is in hydrogen bonding distance to the side chain of the Tyr-103 (2.1 Å), Pro-1 (2.9 Å), and Asn-114 (2.3 Å) residues and Leu-38 carbonyl main chain (3.1 Å). Arg-70 and Arg-73 are (F) farther away in the unreacted substrate ACA and (G) move closer within hydrogen bonding distance to the reacted intermediate 3-(<i>N</i> -prolyl)-3-hydroxypropionate forming a salt bridge.	73
Figure 2.20. 3-(<i>N</i> -prolyl)-acrylate (ACR) intermediate aligned with (A) 3-(<i>N</i> -prolyl)-3-hydroxypropionate (HPA) intermediate or (B) sulfate within the active site, respectively. The ordered water resembles the stage prior to water's nucleophilic attack.	75
Figure 2.21. Error prone PCR and activity assay to screen for highly active variants.	77
Figure 2.22. Induction of RFP by biosensor.	79
Figure 2.23. Lysate supplemented biosensor for Cg10062 (E114N).	80
Figure 2.24. Alignment of Y103F variants of Cg10062 (pink) and <i>cis</i> -CaaD (blue). There is a 1.4 Å shift in the α -carbon position of residue Glu-114 and around 5 Å shift in the loop at residue 120.	81
Figure 2.25. Hydrogen bonding for different enzymes.	83
Figure 2.26. The crystal structure of (A) truncated <i>cis</i> -CaaD that has a sulfate in the active site. <i>N</i> - <i>cis</i> -CaaD/Cg10062 shows (B) a chloride ion trapped in the active site. The C-terminal is disordered (C), only the first four amino acids of the Cg10062 C-terminal sequence provide defined electron density. An overlay of the two structures with wild-type <i>cis</i> -CaaD (blue), reveals (D) Arg-70 of truncated <i>cis</i> -CaaD flipped out of the active site upon disorder of the C-terminal.	87
Figure 2.27. Arg-70 residue movement for all <i>cis</i> -CaaD crystal structure when in apo form (red), ACA soaked (blue), and CCA soaked (green).	90
Figure 3.1. <i>E. coli</i> growth on different medias (A) M9/Glu, (B) M9/Gly, (C) M9/ACA, (D) M9/ACA with M9/Glu, (E) and M9/ACA with M9/Gly.	97

Figure 3.2. Malonic semialdehyde transport into the cell and further production of 3-hydroxypropionic acid.....	98
Figure 3.3. (A) Growth curve and (B) ¹ H NMR analysis of products of MG1655 <i>E. coli</i> cells when grown on M9/ACA media containing Cg10062 (E114N) lysate in media.	99
Figure 3.4. ¹ H NMR analysis of products of BL21(DE3) <i>E. coli</i> cells encoding MmsB for the production of ethanol when grown on M9/ACA media containing Cg10062 (E114N) lysate in media.....	100
Figure 3.5. Cg10062 (E114N) fusion protein conversion of ACA to MSA outside the cell and subsequent transfer of MSA into the <i>E. coli</i> cell.....	101
Figure 3.6. Cg10062 (E114N) fusion protein in the large-scale production of high value chemicals from CH ₄ and CO ₂ starting products.	101
Figure 3.7. pAIDA1 expression vector.....	102
Figure 3.8. AIDA/Cg10062 (E114N) fusion expression and transport through the cell membrane of <i>E. coli</i>	103
Figure 3.9. ¹ H NMR spectra of ACA conversion in media. No MSA is shown implicating that MSA is readily taken up by the cell. Ethanol indicates acetaldehyde conversion via anaerobic fermentation.	104
Figure 3.10. ACA depletion by cells with time.....	104
Figure 3.11. Increase in cell growth with time relative of ACA decline. (I) indicates trial 1 and (II) indicates trial 2.	105
Figure 4.1. SDS-PAGE of <i>cis</i> -CaaD (18.5 kDa), <i>cis</i> -CaaD (E114D) (18.5 kDa), <i>cis</i> -CaaD (E114N) (18.5 kDa), and <i>cis</i> -CaaD (E114Q) (18.5 kDa). Lane 1: protein ladder, lane 2: uninduced culture of <i>cis</i> -CaaD (wild-type), lane 3: induced culture of <i>cis</i> -CaaD (wild-type), lane 4: uninduced culture of <i>cis</i> -CaaD (E114D), lane 5: induced culture of <i>cis</i> -CaaD (E114D), lane 6: uninduced culture of <i>cis</i> -CaaD (E114N), lane 7: induced culture of <i>cis</i> -CaaD (E114N), lane 8: uninduced culture of <i>cis</i> -CaaD (E114Q), and lane 9: induced culture of <i>cis</i> -CaaD (E114Q). (Other variants not shown.) The band at 14 kDa depicts lysozyme used in cell lysis during SDS-PAGE preparation.....	117
Figure 4.2. FPLC purification of wild-type <i>cis</i> -CaaD. Other variants (not shown) were purified similarly.....	119
Figure 4.3. Plasmid map of pNJD1.002.....	123
Figure 4.4. Codon-optimized nucleotide sequence for expression of <i>cis</i> -CaaD in <i>E.coli</i> . Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.	123
Figure 4.5. Plasmid map of pKS2.0503.....	124

Figure 4.6. Codon-optimized nucleotide sequence for expression of <i>N-cis</i> -CaaD/C-Cg10062 in <i>E. coli</i> . Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.....	124
Figure 4.7. Plasmid map of pKS2.0504.....	125
Figure 4.8. Codon-optimized nucleotide sequence for expression of <i>N</i> -Cg10062/C- <i>cis</i> -CaaD in <i>E. coli</i> . Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.....	125
Figure 4.9. Plasmid map of pKS2.0615.....	126
Figure 4.10. Codon-optimized nucleotide sequence for expression of <i>tcis</i> -CaaD in <i>E. coli</i> . Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.	126
Figure 4.11. Plasmid map of pAS1.046.....	127
Figure 4.12. Plasmid map for pAS2.031.....	127
Figure 4.13. Plasmid map for pKK1.1025.....	128
Figure 4.14. Plasmid map of pAS5.003.....	129
Figure 4.15. Plasmid map of pAS5.005.....	130
Figure 4.16. Kinetic trend example.....	133
Figure 4.17. Michaelis-Menten kinetics of Cg10062 (E114N).	134
Figure 4.18. Steady state kinetics of (A) <i>N-cis</i> -CaaD/C-Cg10062 with CCA, (B) <i>N</i> -Cg10062/C- <i>cis</i> -CaaD with ACA, and (C) <i>N</i> -Cg10062/C- <i>cis</i> -CaaD with CCA.....	135
Figure 4.19. Crystal structure of <i>cis</i> -CaaD (wild-type) soaked with substrate showing an (A) unreacted ACA in the active site at 2.0 Å, (B) electron density before addition of substrate in active site, (C) trimer with one monomer in a lighter shade, (D) side view of trimer.	143
Figure 4.20. Crystal structure of wild-type <i>cis</i> -CaaD at 1.3 Å (A) with an acetate in the active site, (B) electron density before addition of acetate, (C) top view of trimer with one monomer in a lighter shade, (D) side view of trimer, and (E) aligned with the ACA-soaked structure showing ACA in the active site with acetate.	144
Figure 4.21. Crystal structure of apo <i>cis</i> -CaaD (E114D) with (A) sulfate shown in the active site, (B) a top view of the trimer with one monomer a lighter shade, and (C) a side view of the trimer.	147
Figure 4.22. Crystal structures of ACA soaked <i>cis</i> -CaaD (E114D) at 2.5 Å showing (A) unreacted substrate and (B) covalent intermediate of 3-(<i>N</i> -prolyl)-acrylate linked to catalytic Pro-1. The electron density is shown (C) before and (D) after the addition of 3-(<i>N</i> -prolyl)-acrylate into the active site. The trimer is shown with (E) a top view and (F) side view.	148

- Figure 4.23. Crystal structures of CCA soaked *cis*-CaaD (E114D) at 1.8 Å showing the covalent intermediate of (*N*-prolyl)-ethene linked to catalytic Pro-1 with sulfate (A) and (B) sulfate alone product. Electron density shows active site before the addition of ligand (C). The trimer is shown at a top view (D) with one monomer a lighter shade and at a side view (E). 150
- Figure 4.24. Crystal structure of apo *cis*-CaaD (E114N) is shown with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view. 152
- Figure 4.25. Crystal structure of shifted orientation of (A) ACA and (B) MSA product in the active site of *cis*-CaaD (E114N) is shown at 2.3 Å. Electron density is shown before the addition of (C) ACA and (D) MSA. The trimer is shown (E) at a top view with one monomer a lighter shade and (F) at a side view. 153
- Figure 4.26. Crystal structure of apo *cis*-CaaD (E114Q) shown with (A) phosphate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view. 156
- Figure 4.27. Crystal structures of ACA soaked *cis*-CaaD (E114Q) at 1.9 Å crystals showing (A) covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1. Electron density is shown (B) before the addition of 3-(*N*-prolyl)-3-hydroxypropionate. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view. 157
- Figure 4.28. Crystal structures of CCA soaked *cis*-CaaD (E114Q) at 2.2 Å crystals showing (A) covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1, (B) covalent intermediate of 3-(*N*-prolyl)-acrylate linked to catalytic Pro-1 and (C) (*N*-prolyl)-ethene covalent intermediate with sulfate. Electron density is shown before the addition of (D) 3-(*N*-prolyl)-3-hydroxypropionate, (E) 3-(*N*-prolyl)-acrylate and (F) (*N*-prolyl)-ethene with sulfate. The trimer is shown (G) at a top view with one monomer a lighter shade and (H) at a side view. 158
- Figure 4.29. Crystal structure of apo *cis*-CaaD (Y103F) is shown with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view. 161
- Figure 4.30. Crystal structures of ACA soaked *cis*-CaaD (Y103F) with (A) sulfate in the active site. Electron density is shown (B) before the addition of sulfate. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view. 162
- Figure 4.31. Crystal structures of CCA soaked *cis*-CaaD (Y103F) with (A) three waters in the active site. Electron density is shown (B) before the addition of waters. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view. 163
- Figure 4.32. Crystal structures of ACA soaked *cis*-CaaD (H28A) with (A) acetate and (B) water in the active site. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view. 164
- Figure 4.33. Crystal structure of apo *cis*-CaaD (T34A) with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view. 166

Figure 4.34. Crystal structures of ACA soaked <i>cis</i> -CaaD (T34A) shown with (A) sulfate in the active site (A). The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.....	167
Figure 4.35. Crystal structures of CCA soaked <i>cis</i> -CaaD (T34A) shown with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.....	168
Figure 4.36. Crystal structure of apo Cg10062 (E114N) with (A) sulfate in the active site. The two trimers are shown (B) at a top view and (C) at a side view.	169
Figure 4.37. Crystal structure of the first obtained structure of ACA-soaked Cg10062 (E114N) with (A) ACA in the active site. The single chain shown (B) at a top view and (C) at a side view. The second crystal structure of ACA-soaked Cg10062 (E114N) shows (D) a covalent intermediate of 3-(<i>N</i> -prolyl)-3-hydroxypropionate linked to catalytic Pro-1 and (E) a covalent intermediate of 3-(<i>N</i> -prolyl)-acrylate linked to catalytic Pro-1. The two trimers are shown (F) at atop view and (G) at a side view.	170
Figure 4.38. The crystal structure of truncated <i>cis</i> -CaaD has a (A) sulfate in the active site. The single chain shown (B) at atop view and (C) at a side view.	172
Figure 4.39. <i>N-cis</i> -CaaD/C-Cg10062 shows a (A) chloride ion trapped in the active site. The single chain shown (B) at atop view and (C) at a side view.....	174
Figure 4.40. Plasmid map of pAS2.100.....	195
Figure 4.41. Plasmid map of pAIDA1.....	196
Figure 4.42. Plasmid design and map of pKS8.0410.	198
Figure 4.43. Gibson assembly design to make plasmid pKS8.0411.	200

LIST OF TABLES

Table 1.1 Major reaction types catalyzed by enzymes in the tautomerase superfamily.	20
Table 2.1. Michaelis-Menten values for <i>cis</i> -CaaD (wild-type), <i>cis</i> -CaaD (T32A), and <i>cis</i> -CaaD (T34A) using ACA and CCA substrates. Activity for other variants had insufficient activity to accurately determine steady state kinetics.....	47
Table 2.2. Product profile of variants using ACA substrate.	68
Table 2.3. Michaelis-Menten kinetics for Cg10062 and variants.	69
Table 2.4. Distances between residue 114 and Pro-1 for crystal structures evaluated.	85
Table 2.5. Michaelis-Menten kinetics of chimera and truncated variants compared to wild-type <i>cis</i> -CaaD and Cg10062.....	89
Table 3.1. Bacterial species growth on different solid medias. X indicates growth on the respective media.....	95
Table 4.1. Components of a typical restriction digest.	113
Table 4.2. PCR protocol of parent template.....	115
Table 4.3. PCR conditions for Q5 site-directed mutagenesis.	115
Table 4.4. Parameters used for enzyme quantification.....	120
Table 4.5. Strains and plasmids used in Chapter Two of this study.....	121
Table 4.6. Primers used for Q5 site-directed mutagenesis.....	131
Table 4.7. KLD Treatment.....	132
Table 4.8 Error-prone PCR reaction components.....	137
Table 4.9. Error-prone PCR conditions.....	138
Table 4.10. Template DNA to add for desired mutation frequency.....	138
Table 4.11. T4 ligation conditions.	139
Table 4.12. Crystallographic soaking parameters for wild-type and variants of <i>cis</i> -CaaD and Cg10062 (E114N).....	175
Table 4.13. Crystallographic data collection and structure refinement statistics. ^a Apo and ^b ACA soaked, ^c CCA soaked. ACA, acetylenecarboxylic acid. HPA, 3-(<i>N</i> -prolyl)-3-hydroxypropionate. ACR, 3-(<i>N</i> -prolyl)-acrylate. ETY, (<i>N</i> -prolyl)-ethene. MSA, malonic semialdehyde. ACT, acetate.....	177
Table 4.14. Strains and plasmids used in Chapter Three of this Study.	194

Table 4.15. PCR reactions for the amplification of vector and insert in pAIDA/Cg10062(E114N) fusion protein design.	201
Table 4.16. Thermocycler settings for fusion protein design.	201
Table 4.17. HiFi Gibson assembly reaction components for fusion protein.....	201

LIST OF ABBREVIATIONS

ACA	acetylenecarboxylic acid
ACH	acetaldehyde
ACR	acrylate
ADCA	acetylenedicarboxylic acid
ADH	alcohol dehydrogenase
CCA	<i>cis</i> -3-chloroacrylic acid
CaaD	<i>trans</i> -3-chloroacrylic acid dehalogenase
<i>cis</i> -CaaD	<i>cis</i> -3-chloroacrylic acid dehalogenase
CoA	coenzyme A
Cm	Chloramphenicol
DNA	deoxyribonucleic acid
<i>d</i> ₆ -DMSO	deuterated dimethyl sulfoxide
Glu	glucose
Gly	glycerol
h	hours
HPA	3-(<i>N</i> -prolyl)-3-hydroxypropionate
IPTG	isopropyl- β -D-1-thiogalactopyranoside
kb	kilobase
kDa	kilo Dalton
LB	Luria-Bertani
M	molar
MA	malonic acid
mg	milligram
min	minute(s)
mL	milliliter

mM.....	millimolar
MIF	migration inhibitory factor
MmsB	3-hydroxyisobutyrate dehydrogenase
MSA.....	malonic semialdehyde
MSAD	malonate semialdehyde decarboxylase
NAD ⁺	oxidized nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
NADP ⁺	oxidized nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
OD	optical density
PCR	polymerase chain reaction
PDB	protein data base
PTDH.....	phosphite dehydrogenase
RBS	ribosome binding site
rt	room temperature
s.....	second(s)
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCA	tricarboxylic acid
TSP.....	3-(trimethylsilyl) propionic-2,2,3,3- <i>d</i> ₄ acid
U	unit; $\mu\text{mol min}^{-1}$
YdfG	NADP dependent 3-hydroxy acid dehydrogenase
Ynel	succinate semialdehyde dehydrogenase
3HP.....	3-hydroxypropionic acid
4-OT	4-oxocrotonate

CHAPTER ONE: New Strategies for Commodity Chemical Production

1.1. Caveats of Current Production of Commodity Chemicals

The identification of optimal carbon feedstocks for the production of high value building block materials is fundamental in the chemical industry. Traditionally, most building block chemicals, commodity chemicals, and value-added chemicals have been derived from petroleum-based carbon sources.^{1,2} Still, these conventional processes are associated with the generation of significant greenhouse emissions, which contribute to damaging effects on the environment and climate. In global efforts to prioritize sustainability, we see a notable shift towards embracing more environmentally and economically viable alternatives to petroleum-based feedstocks for chemical production. Petroleum remains the largest source of global chemical production comprising approximately 80% of the industry's output.¹ A vast majority of this petroleum is used for the production of chemicals and fuels. Petrochemicals derived from oil are used in the production of plastics, fertilizers, and other chemicals.² However, the synthesis of these products involves the release of toxic greenhouse gasses. Therefore, the reliance on nonrenewable fuels for chemical production is non-sustainable in the long term. Chemical production from petroleum is responsible for the emission of tons of greenhouse gases every year which proves detrimental to the environment and climate. The Haber-Bosch process which is used for ammonia synthesis in the production of synthetic fertilizer and other chemicals is an example of how chemical production impacts the environment. Here, hydrogen derived from natural gas is used in a method that releases more hydrogen and significant quantities of CO₂ into the atmosphere.³ The petroleum industry alone emits approximately 1.5 gigatons of CO₂ annually.⁴ These emissions contribute to worldwide climate change, health effects among the population, negative effects on the ecosystem, biodiversity and a decrease in food supply.⁵ The long-term health effects, climate effects, and non-sustainability that fossil fuels pose make petroleum a resource to draw back from.

Given the environmental and health concerns associated with petroleum-based chemical production, as well as the limited nature of fossil fuels, there is pressure to move to more

sustainable alternatives. As the automotive industry moves towards electric vehicles and improved fuel economies, the demand for gasoline is likely to decrease. With the reduction of fuel-based automobiles, there is less petroleum being processed, and therefore, the chemical industry will become less reliant on byproducts from the refining of petroleum to obtain transportation fuels. Consequently, the chemical industry will need to explore alternative sources for feedstocks, reduce its reliance on petroleum-based feedstocks, and investigate renewable and sustainable alternatives.⁶

Major commodity chemicals such as acrylonitrile, acrylic acid, and acrylamide are traditionally produced from petroleum. The market sizes for acrylonitrile, acrylic acid, and acrylamide in 2024 was 11.9 billion USD, 14.2 billion USD, and 3.8 billion USD, respectively, and each is expected to increase.⁷⁻⁹ Acrylonitrile is made from petroleum by extracting propylene through the refinement of petroleum (cracking), synthesizing ammonia using the Haber process, and then reacting with oxygen to form acrylonitrile in a process called the Sohio process.¹⁰ Acrylonitrile is essential in the production of acrylic fibers, carbon fibers, and nitrile rubber used in textiles, furnishings, the automotive industry, and consumer goods. Acrylic acid is formed from the oxidation of propylene (Figure 1.1). Similarly, propylene is obtained as a byproduct of petroleum refining during the cracking of hydrocarbons. Propylene is then oxidized to form acrolein using a vanadium oxide catalysis and then further oxidized to form acrylic acid using a molybdenum oxide catalyst.¹¹ Acrylonitrile is a precursor in the production of acrylamide. Acrylonitrile is hydrolyzed to form acrylamide under heat, pressure, and in the presence of a strong acid or base.¹² The production of these high value chemicals is reliant on petroleum, therefore investigations into their production by alternative and more sustainable means is of growing interest.

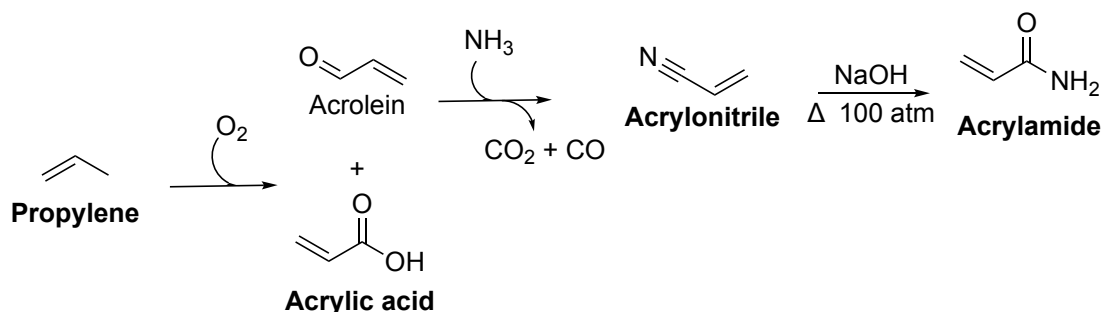


Figure 1.1. Production of acrylonitrile, acrylic acid, and acrylamide from propylene.

In a step away from petroleum, there has been a significant shift towards the use of biomass as a feedstock for chemical production (Figure 1.2). Biomass is organic material such as plants and waste that can be used as a renewable source of carbon for the synthesis of chemicals through processes such as fermentation, chemical catalysis, and enzymatic conversion.¹³ Biomass feedstocks are classified into two categories: first-generation and second-generation. First generation biomass production utilizes biomass that is edible such as crops like corn and sugarcane. First generation biomass feedstocks place the chemical industry in direct competition with the food industry as it utilizes feed that could be used as a nutritional source. Previous studies have shown the production of high value chemicals from first-generation biomass such as malonic acid from glucose.¹⁴⁻¹⁶ An increasing population requires an increase in food. Utilization of first generation biomass as a feedstock also creates competition for agricultural resources by using land that could otherwise be used for nutriment.^{17,18} Second generation biomass is non-edible crop material that includes lignocellulose and hemicellulose that compromises corn stover, bark, and solid food waste. These materials are more abundant and do not compete directly with food production, but their conversion into useful chemicals can be very lengthy and involves expensive processing that minimizes its benefit.¹⁹⁻²²

For this reason, CH_4 and CO_2 are being investigated as a feedstock for the production of chemicals. CH_4 and CO_2 can be captured from natural gas and biogas sources, such as from the

anaerobic digestion of organic waste.²³ These gasses can be used in biocatalysis by microbes to be converted to high value chemicals, presenting a promising route for chemical production and reducing dependence on fossil fuels.^{24,25} However, the direct use of CH_4 and CO_2 as substrates are problematic, as keeping the gasses in solution proves difficult. Additionally, enzymatic conversion of CH_4 and CO_2 is generally slow. The direct conversion of CH_4 and CO_2 has produced high value chemicals such as 3-hydroxypropionic acid.²⁵ However, the titer and yields from these processes remain low and the scaling up of these reactions has yet to be optimized.²⁷ Malonic acid, a key building block for the synthesis of flavors, fragrances, and pharmaceuticals, has been produced on an industrial scale, with companies such as Lygos establishing successful biosynthetic pathways for its production.²⁶⁻²⁹ 3-Hydroxypropionic acid is a Department of Energy-targeted chemical and serves as a precursor to important industrial chemicals such as acrylic acid, acrylamide, and acrylonitrile, which are widely used in textiles, superabsorbent polymers, and adhesives.²⁸ In efforts to produce high value chemicals like 3-hydroxypropionic acid on a larger scale, we look to intermediate chemicals that may be produced from CH_4 and CO_2 and subsequently converted at a high kinetic rate to high value chemicals through biocatalysis. This brings our attention to acetylenecarboxylic acid as a feedstock. The development of more efficient bioconversion processes, including the use of intermediate chemicals like acetylenecarboxylic acid, could help advance biocatalysis and biosynthetic pathways to enable the large-scale production of high-value chemicals by environmentally sustainable means.



Figure 1.2. Petroleum, biomass, and CH_4 and CO_2 feedstocks for chemical production.

1.1.1 Acetylenecarboxylic Acid Production from CH₄ and CO₂

The low rate of chemical production from direct CH₄ and CO₂ use as a feedstock has led to the exploration of intermediate chemicals derived from CH₄ and CO₂, such as formate and acetate, that may be more effective feedstocks. Both CH₄ and CO₂ can be captured from the anaerobic digestion of waste. Utilizing CH₄ and CO₂ in chemical production can reduce environmental gas emissions, avoid food competition posed by biomass-based methods, and promote sustainable renewable chemical production. We can produce value added chemicals without competing with the agricultural industry while simultaneously converting environmental greenhouse gases to useful chemicals. These gasses have previously been described in the conversion to valuable chemicals through methodologies such as thermal catalysis, electrocatalysis, photocatalysis, and photoelectrocatalysis.³⁰ Alternately, acetylenecarboxylic acid (ACA) is a novel intermediate that can be used for chemical production and is formed by the direct carboxylation of acetylene (Figure 1.3).³¹ Acetylene can be generated from CH₄ via dehydrodimerization.³²

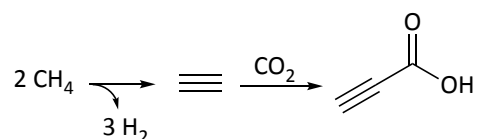


Figure 1.3. Scheme for the production of ACA.

Acetylene, a highly reactive and unstable gas, dangerous for large-scale transportation in its pure and condensed form, can be carboxylated to form ACA, a more stable liquid. Acetylene and CO₂ can be converted to ACA, a stable liquid. Where using food biomass lacks sustainability, other means of chemical production such as petroleum derived carbon sources are limited and pose other restraints. Biocatalysis, however, has many suitable methods. Utilizing enzymes for the biocatalytic production of high value or commodity chemicals is of growing interest. Figure 1.4 proposes a model for the enzymatic production of high value chemicals which will be discussed in this dissertation.

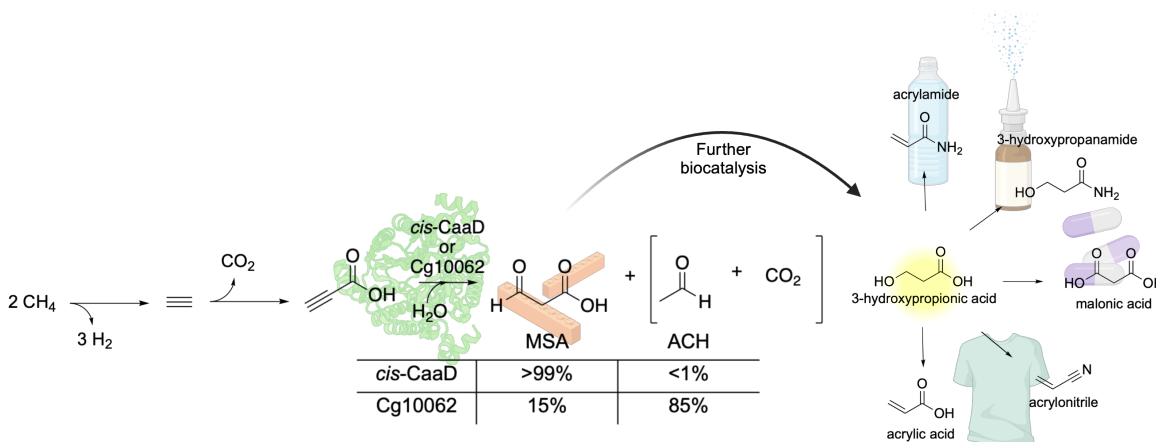


Figure 1.4. Pathway from greenhouse gasses CO_2 and CH_4 via substrate ACA using enzymes *cis*-CaaD and Cg10062 to produce MSA and further form high value chemicals.

Currently, there is no large-scale industrial synthesis of ACA. However, methods exist for the production of ACA derived from CH_4 and CO_2 . Acetylene can be produced from hydrogen, oxygen and CH_4 in a production developed by BASF (Figure 1.5).³³ However, this production leads to a high carbon footprint as most CH_4 is burned in this process. Other studies show that acetylene can be generated from CH_4 in high yield, with high selectivity and high carbon balance through dehydrodimerization. This process, although not used on an industrial scale, utilizes a plasma reactor by transporting CH_4 with H_2 as a carrier gas at supersonic speeds.³⁴⁻³⁶ The high temperature thermal plasma dimerizes CH_4 to acetylene with over 95% selectivity and 70% conversion (Figure 1.6).³⁶ Unreacted CH_4 is then recycled through the plasma reactor for further efficient acetylene conversion. However, acetylene in its gaseous phase is unstable, making it an unideal intermediate for biocatalysis. Thus, further carboxylation to ACA is necessary.

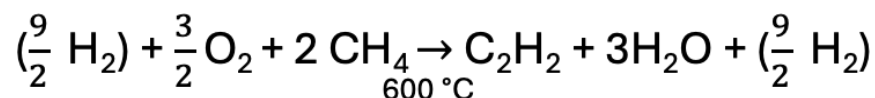


Figure 1.5. BASF synthesis of acetylene.

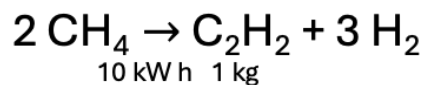


Figure 1.6. Plasma arc production of acetylene.

Carboxylation of terminal alkynes has been extensively studied using metal catalysis for the carboxylation of acetylene.³⁷⁻³⁹ Presently, ACA is synthesized by the oxidation of propargyl alcohol by a PbO₂ electrode.⁴⁰ However, acetylene may be carboxylated to form ACA with the use of DBU and TBD (Figure 1.7).³¹ ACA is more stable substrate that can therefore be used in the bioconversion to high value chemicals. Biocatalysis, through the use of engineered enzymes, offers a pathway for the conversion of ACA to high-value chemicals, such as 3HP.

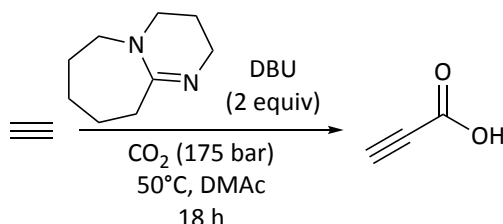


Figure 1.7. Production of ACA from acetylene.

1.1.2 Production of High Demand Chemicals

Microorganisms can be genetically engineered to produce a wide range of chemicals and other valuable products through biosynthetic pathways, offering a promising approach for sustainable chemical production.^{24,25,41-43} This has been demonstrated in bioengineered bacteria designed to produce 3HP.^{42,43} There has been successful demonstration of converting CH₄ and CO₂ into valuable chemicals such as 3-hydroxypropionic acid (3HP) and muconic acid enzymatically using engineered bacteria. An engineered *E. coli* strain was able to produce 15.8 g L⁻¹ of 3HP (yield of 0.71 g g⁻¹ yield) from chemically synthesized acetic acid from CO₂ and 11.2 g L⁻¹ of 3HP (0.55 g g⁻¹ yield) from syngas (a gas mixture of hydrogen and carbon monoxide)-derived acetic acid.²⁵ Alternately, an engineered strain of the methanotroph, *M. alcaliphilum*, gave lower yields of 2.7 mg muconic acid per gram of CH₄.²⁴ These engineered microorganisms, or microbial

cell factories, can use alternative chemical feedstocks, such as sugars, agricultural waste, gases like CO₂ or CH₄ or the chemicals produced from them, as their primary energy and carbon source for the production of high priority chemicals. This is in contrast to traditional chemical production methods that rely heavily on fossil fuel and petroleum derived sources, making microbial-based production a more sustainable alternative.

The integration of biotechnology and synthetic biology is used to engineer these microorganisms and enable them to produce bulk chemicals that are valuable to several industries. Through genetic modifications, microorganisms can be optimized to produce chemicals that are valuable to the, chemical, material, pharmaceutical, and biofuel industries. Specifically, biocatalysis is a highly efficient and selective alternative to conventional chemical production methods that often involve harsh conditions, non-renewable resources, and toxic solvents. By taking advantage of the metabolic capabilities of engineered microorganisms, biocatalysis offers a more sustainable pathway for producing high-priority chemicals with fewer environmental impacts and potentially lower energy consumption

The production of high priority chemicals through biocatalysis is a rapidly expanding area of scientific research, offering substantial advantages for sustainable and efficient manufacturing processes. 3-Hydroxypropionic acid (3HP), a key building block chemical, has gained attention from institutions like the U.S. Department of Energy as an ideal target for biocatalytic production from renewable feedstocks such as biomass or syngas.⁴¹ 3HP is sought out as an important precursor in the synthesis of various high value chemicals such as 1,3-propanediol, acrylonitrile, acrylamide, and acrylic acid (Figure 1.8).⁴²⁻⁴⁴ These chemicals are essentials for industries involved in adhesives, coatings, dyes, polymers, super absorbent polymers, and resins. Additionally, in 2008, Cargill and Novozymes announced a joint venture for the development and commercialization of bio-based acrylic acid through the dehydration of 3HP obtained from fermentation of renewable sugars making 3HP a valuable target chemicals.⁴⁵

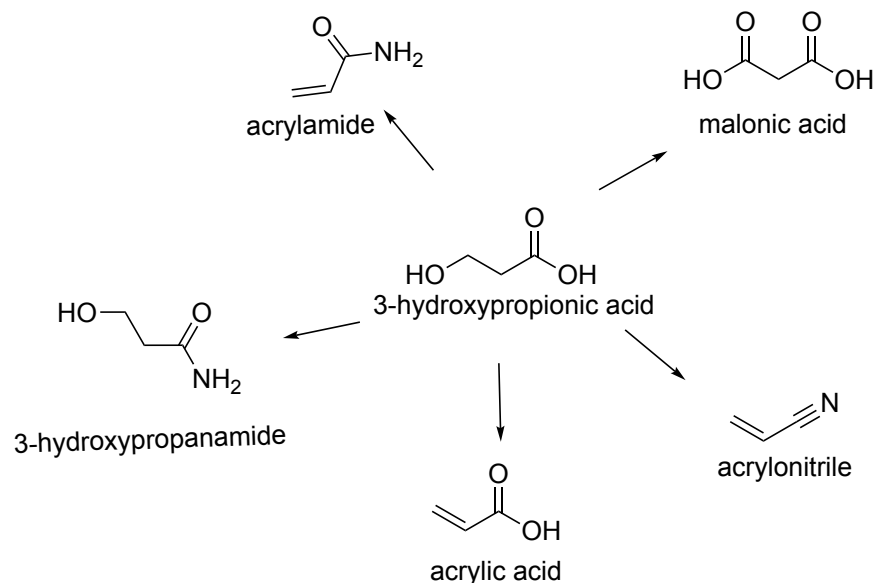


Figure 1.8. High value chemicals from 3HP.

Biocatalytic synthesis of commodity chemicals through enzymology is generally carried out through either *in vitro* or *in vivo* approaches. *In vitro* methods generally present high specificity, enzyme selectivity, and a reduced need for complex cellular design.⁴⁶ In many cases, *in vitro* synthesis using isolated enzymes or cell-free systems leads to low efficiency and poor product yield due to manipulation of reaction conditions that inhibit optimal enzyme function. By utilizing the full metabolic and altering the regulatory network of the microbe, *in vivo* production may be able to increase production of chemicals by the microorganism. However, small genetic manipulations may affect yields as well by altering the metabolism and efficacy of the enzyme or enzymes involved.

In addition, cost of production decreases with *in vivo* systems as a need to purify individual protein enzymes, an often costly and time-consuming process, is not required. *In vivo* processes also benefit from the ability to integrate multiple enzymatic steps into a single organism. This reduces the need for complex reaction cascades and optimizing reaction efficiency. Furthermore, microorganisms can often be grown on substrate feedstocks or evolved to grow on feedstocks enhancing the viability of this approach. Nevertheless, *in vivo* processes require further

refinement involving metabolic engineering and bioprocess optimization to enhance product formation, scalability and economic feasibility of using microbial production systems. While microbial biosynthesis is a promising alternative to traditional chemical production, optimization is required to achieve large scale success. Improving the efficiency of both *in vitro* and *in vivo* systems is important in securing a future of sustainable chemical production. While these systems offer advantages and the potential for optimization and integration of enzymatic pathways, their use has been demonstrated in literature as 3HP has been produced from biomass derived feedstocks such as glycerol and glucose through biocatalysis

Glycerol is a readily available and renewable carbon source, produced as a by-product in the biodiesel industry, generating around 100 kg of crude glycerol per ton of biodiesel produced.⁴⁷ It is also a significant by-product in bioethanol and soap production processes.⁴⁸ Often regarded as waste, glycerol's non-toxic nature and abundance make it an ideal substrate for high-titer 3HP production. There are two pathways for the production of 3HP from glycerol: CoA-dependent and CoA-independent.

The CoA-dependent pathway (Figure 1.9) for 3HP production from glycerol follows the metabolic pathway for 1,2-propanediol degradation involving enzymes in the *pdu* operon (native to the *Salmonella*, *Klebsiella* and *Lactobacillus* species) for propanediol utilization and allows for 3HP production.⁴⁹ The process begins with the coenzyme B₁₂-dependent glycerol dehydratase catalyzing the conversion of glycerol to 3-hydroxypropionaldehyde.⁵⁰ 3-Hydroxypropionaldehyde is then oxidized by propionaldehyde dehydrogenase to form 3-hydroxypropionyl-CoA. Next, 3-hydroxypropionyl-CoA is phosphorylated by phosphate propanoyltransferase and dephosphorylated by propionate kinase to yield 3HP. This 3HP synthesis is conserved in several strains, such as *Lactobacillus reuteri* and *Klebsiella pneumoniae*, making them economically viable due to their native ability to produce coenzyme B12, unlike recombinant *E. coli* strains that require exogenous coenzyme supplementation.⁵¹ However, 3HP production via this pathway is limited by the toxic accumulation of 3-hydroxypropionaldehyde.

To overcome this issue, a microbial consortium of *L. reuteri* producing 3-hydroxypropionaldehyde was co-cultured with recombinant *E. coli* strains overexpressing aldehyde dehydrogenase, which reduces 3-hydroxypropionaldehyde to 3HP. This approach resulted in titers of 125 g L⁻¹ of 3HP and 88 g L⁻¹ of 1,3-propanediol.⁵² While certain species with the native *pdu* operon can naturally produce 3HP, the 3-hydroxypropionaldehyde intermediate can also be converted to 1,3-propanediol by 1,3-propanediol reductase, leading to equimolar quantities of both 3HP and 1,3-propanediol, thus limiting the maximum yield of 3HP. However, the co-production of 1,3-propanediol alongside 3HP can be beneficial, as the NAD⁺ required for converting 3-hydroxypropionaldehyde to 3-hydroxypropionyl-CoA in the Pdu pathway can be regenerated by 1,3-propanediol reductase.⁵³

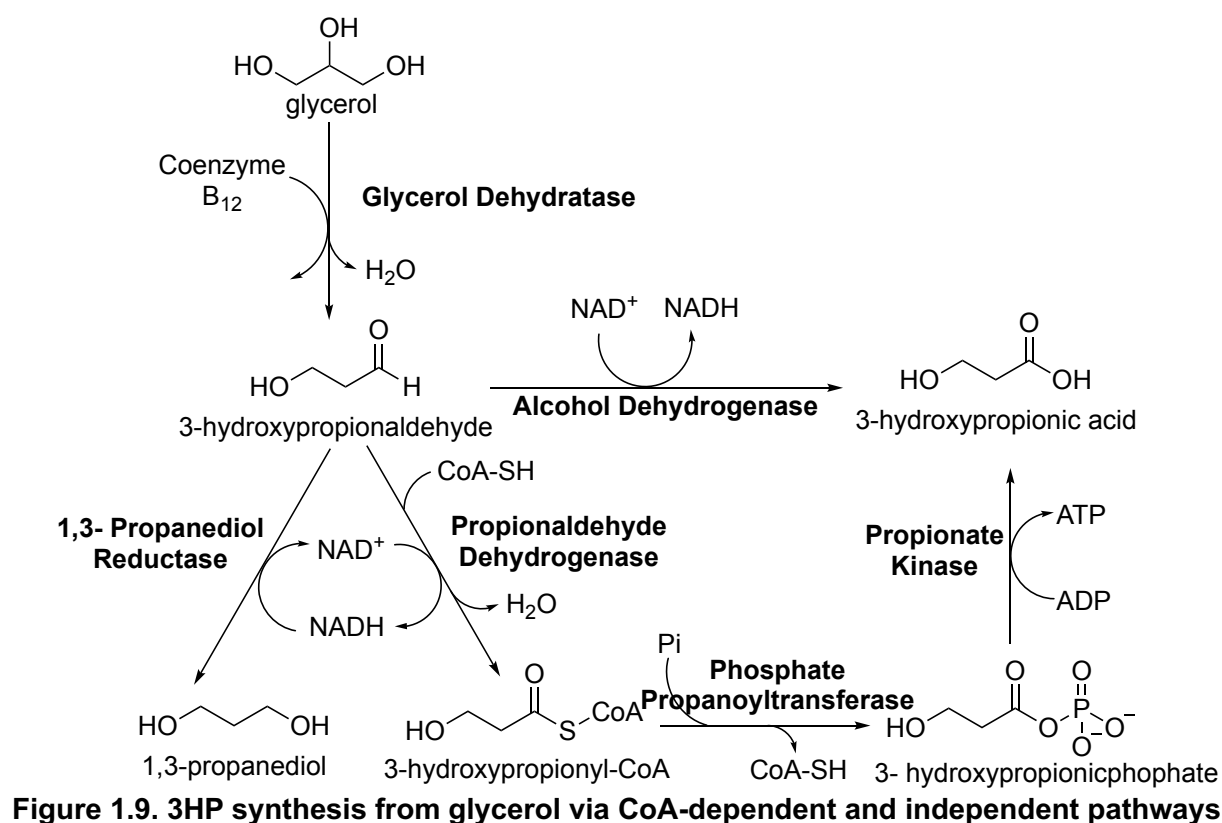


Figure 1.9. 3HP synthesis from glycerol via CoA-dependent and independent pathways.

3HP can also be synthesized through a CoA-independent pathway (Figure 1.9). This two-step pathway, involves the conversion of glycerol to 3HPA via coenzyme B₁₂-dependent glycerol

dehydratase, followed by oxidation to 3HP by an aldehyde dehydrogenase.⁵⁴ Production of up to 102 g L⁻¹ of 3HP has been achieved using *Klebsiella pneumoniae*.⁵⁵ This pathway is preferred over the CoA-dependent pathway as it allows 3HP production by overexpression of aldehyde dehydrogenase and does not require CoA.

In both CoA-dependent and CoA-independent pathways, maintaining redox balance and ensuring an adequate supply of NAD⁺ are crucial for efficient *in vivo* 3HP biosynthesis. NAD⁺ can be regenerated through oxidative glucose metabolism, the electron transport chain, or by increasing aeration in cell cultures. However, microbial optimization strategies must be carefully refined, as coenzyme B₁₂ can be inactivated by oxygen.⁵⁶ Additionally, glycerol dehydratase undergoes suicide-based inactivation by its substrate glycerol during its catalytic cycle, where coenzyme B₁₂ is not regenerated and an inactive form of the coenzyme remains bound to the enzyme, preventing further interaction with free coenzyme B₁₂.⁵⁷ However, studies show that glycerol dehydratase may be reactivated in the presence of ATP and Mg²⁺.⁵⁸

An alternative starting material for the production of 3HP is glucose which proceeds via two main pathways: the malonyl-CoA pathway and the β -alanine pathway. In the malonyl-CoA pathway, glucose is converted to acetyl-CoA via glycolysis. Rather than entering the tricarboxylic acid (TCA) cycle, acetyl-CoA is converted into malonyl-CoA by acetyl-CoA carboxylase. Acetyl-CoA carboxylase is involved in natural fatty acid metabolism of microorganisms. Malonyl-CoA is then converted into 3HP by first converting to malonate semialdehyde (MSA), by NADPH-dependent malonyl-CoA reductase (Figure 1.10). Malonic semialdehyde reductase, from *Chloroflexus aurantiacus*, has two catalytic domains that facilitate the conversion of malonyl-CoA into 3HP in two steps. The C-terminal domain catalyzes the formation of MSA, while the N-terminal domain reduces MSA to 3HP.⁵⁹ A microbial strain developed by OPX Biotechnologies, now owned by Cargill, that produces 3HP from glucose uses this malonyl-CoA pathway. However, a major limitation of this pathway is the requirement for two equivalents of NADPH for every mole of 3HP synthesized.

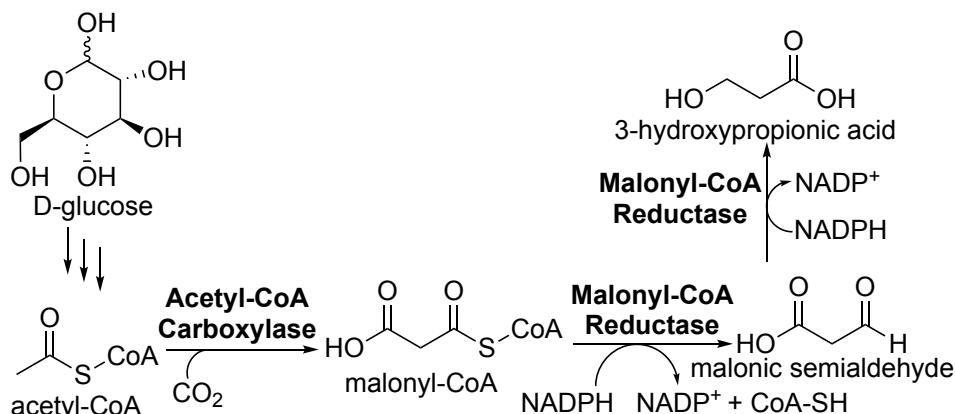


Figure 1.10. 3HP synthesis from glucose via the malonyl-CoA pathway.

The β -alanine pathway utilizes MSA, identical to that of the malonyl-CoA pathway. However, MSA in the β -alanine pathway is derived from the TCA cycle.⁶⁰ Earlier, aspartate is produced through the aspartate aminotransferase-catalyzed conversion of fumarate (Figure 1.11). β -Alanine is then formed by the decarboxylation of aspartate catalyzed by aspartate decarboxylase. The resulting β -alanine is converted into MSA by β -alanine pyruvate transaminase. Finally, 3HP is formed from MSA catalyzed by malonic semialdehyde reductase. High production of 3HP was reported in a recombinant *E. coli* strain at 31.1 g L⁻¹ using this pathway.⁶¹

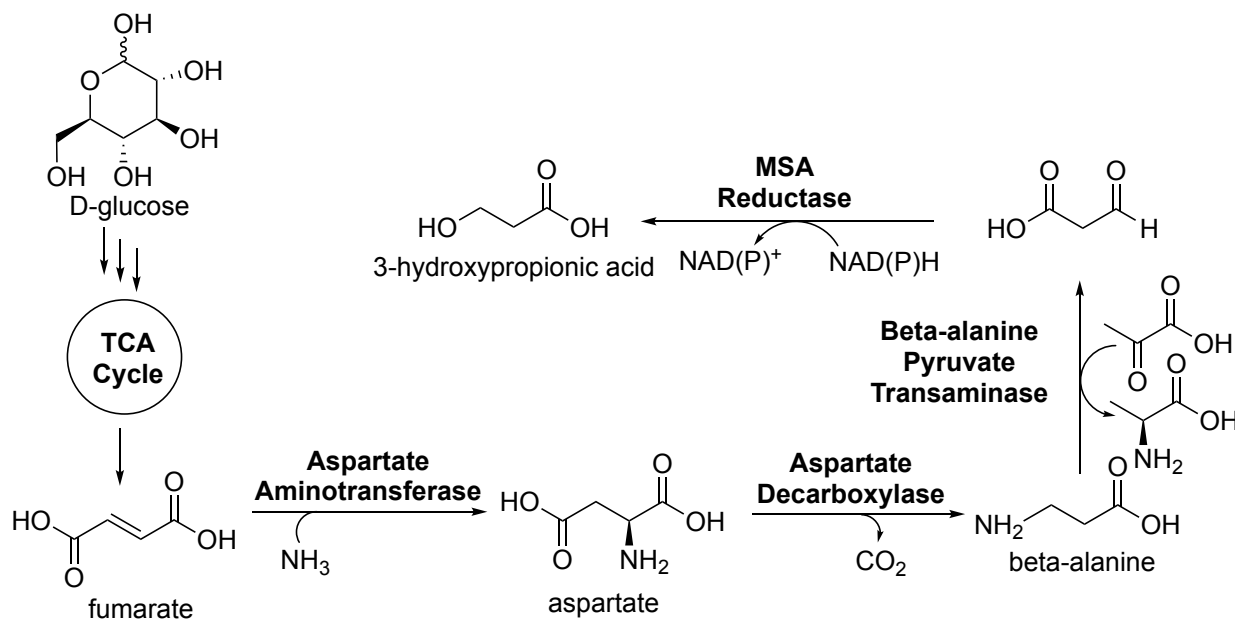


Figure 1.11. Synthesis of 3HP from glucose via the β -alanine pathway.

In 1958 Yamada et al. identified a crucial enzyme that was capable of converting acetylenedicarboxylic acid into pyruvate for further cellular metabolism.⁶² This study demonstrated how acetylenic compounds can be used by microorganisms and converted into a useful intermediate. From this we understand enzyme substrate specificity and the diverse reactions that enzymes can catalyze. This provides insight into the biocatalytic utilization of acetylenic compounds and their applications in chemical synthesis.

The development and optimization of genetically engineered microorganisms for the production of valuable chemicals through biocatalysis represents a promising path towards more sustainable and efficient chemical manufacturing. By leveraging renewable feedstocks microorganisms offer a significant advantage over traditional fossil fuel-based and biomass methods, reducing environmental impact and reliance on non-renewable resources. The potential of biocatalysis, particularly in the production of high-priority chemicals like 3HP, demonstrates the versatility and scalability of engineered microbes. However, further refinement is needed in both *in vitro* and *in vivo* systems to improve efficiency, yield, and economic feasibility. As research continues to advance, the integration of synthetic biology, metabolic engineering, and bioprocess optimization will play a crucial role in unlocking the full potential of microbial-based production systems, paving the way for a more sustainable future in chemical manufacturing. Building on the potential of engineered microorganisms for sustainable chemical production, enzymology plays a pivotal role in advancing biocatalysis as a more efficient and selective alternative to traditional chemical synthesis. The use of enzymes in biocatalytic processes allows for highly specific reactions under mild conditions, which contrasts sharply with the harsh temperatures, pressures, and toxic solvents often required in conventional methods. Enzymes act as natural catalysts that can accelerate reactions with remarkable precision, minimizing the need for complex reaction conditions and reducing energy consumption. Furthermore, advances in enzyme engineering and optimization are critical for overcoming challenges such as enzyme inactivation and improving overall process yields. Enzymology and biocatalysis proves ideal for the production of commodity

chemical and optimization of these processes remain central to developing more cost-effective, scalable, and sustainable biocatalytic processes.

1.2 The Tautomerase Superfamily

To further enhance the potential of biocatalysis in sustainable chemical production, understanding the specific properties and mechanisms of enzymes becomes crucial, particularly those from unique enzyme families such as the tautomerase superfamily. Enzymes from the tautomerase superfamily are characterized by a β - α - β structural motif and catalytic *N*-terminal proline.^{63,64} This structural and catalytic feature is highly conserved across the family, distinguishing them from other enzyme classes. Unlike many other enzymes, these tautomerase do not require metal ions or cofactors for catalytic activity, making them unique in their functionality. The tautomerase superfamily includes over 11,000 non-redundant sequences from diverse organisms.⁶⁵ The characteristic similarities, functional diversity, and evolutionary relationships define the subgrouping within the tautomerase superfamily (Figure 1.12). Based on characteristic networks within modern tautomerase superfamily members, they may have evolved from a short 4-oxalocrotonate tautomerase (4-OT) like ancient ancestor through gene duplication and fusion.⁶⁵ The distinctive β - α - β structure of each monomer begins at the *N*-terminal Pro-1 forming a β -strand, followed by an α -helix with a 3₁₀ helix, and finally a second parallel β -strand.⁶⁵ There is a final β -hairpin at the *C*-terminal which assists in complex formation and overall structural stability. The specific number of β - α - β folds within an enzyme sequence dictates whether the tautomerase functions as a homo- or hetero-hexamer or trimer.⁶⁶ Homo and hetero hexamers contain a single β - α - β unit per monomer, while trimers contain two fused β - α - β units joined by a linker region. For example, 4-OT functions as a homohexamer or heterohexamer depending on the species, while another member of the tautomerase superfamily, macrophage inhibitory factor (MIF), exists as a homotrimer.

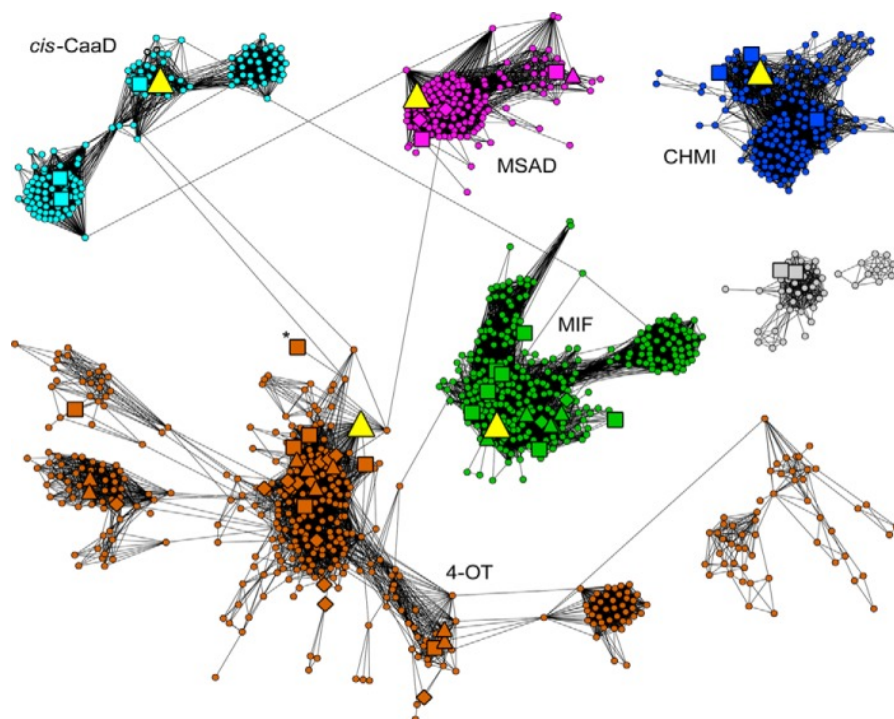


Figure 1.12. The tautomerase superfamily sequence similarity network. The gray subgroup indicates sequences from the tautomerase superfamily that have not yet been assigned to a specific subgroup.⁶⁵

The catalytic Pro-1 residue plays an essential role in activity and has been reported to function as a general base or acid catalyst depending on the pK_a of the active site proline.⁶⁵ This residue is a key component in the mechanism of the tautomerase enzymes and facilitates the various reactions they carry out. In the tautomerase superfamily, there are five general families that are separated by substrate specificity and the reaction they catalyze. Although tautomerase families may show differences in substrate and reaction specificity, these enzymes from different families may display cross-substrate reactivity, although with lower catalytic efficiency. The five families include 4-oxalocrotonate tautomerase (4-OT), 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI), macrophage inhibitory factor (MIF), *cis*- and *trans*- chloroacrylic acid dehalogenase (*cis*-CaaD and CaaD, respectively), and malonate semialdehyde decarboxylase (MSAD).^{67,68} 4-OT contains only one β - α - β fold and is a homohexamer. CHMI, MIF, *cis*-CaaD, and MSAD have two β - α - β fold per monomer and are homotrimers.

1.2.1 4-Oxalocrotonate

4-Oxalocrotonate tautomerase (4-OT) is a hexamer of 62 amino acids that catalyzes the enol-keto tautomerization of a pyruvoyl group (Figure 1.14A).⁶⁹⁻⁷¹ The native function of 4-OT enables it to act as a metabolic catalyst in the degradation of aromatic hydrocarbons. 4-OT operate without the use of cofactors, instead relying on its catalytic Pro-1 residue, which typically has a lower pK_a of about 6.4. Due to its interaction with neighboring residues in the active site, Pro-1 easily becomes deprotonated and acts as a general base to initiate catalysis.⁶³ 4-OT from *Pseudomonas sp.* is known to use 2-hydroxymuconate as a substrate for the tautomerization to 2-oxo-3-hexenedionate and 2-oxo-4-hexenedionate (Figure 1.13, Table 1.1).⁷² In addition, the Whitman research group proposed a mechanism for the catalysis of 4-OT from structures of *Pseudomonas pavonacea* in which Arg-39 acts as a general acid and interacts with the C-1 carboxylate of substrate during binding.⁷³ Arg-11 was also involved in binding the C-6 carboxylate group of the substrate. No crystal structures have shown a trapped intermediate within the active site to confirm the mechanism. However, crystal structures of 4-OT from *P. putida* soaked with irreversible inhibitor, 2-oxo-3-pentynoate, showed a covalent bond between Pro-1 and the C4 of 2-oxo-3-pentynoate.

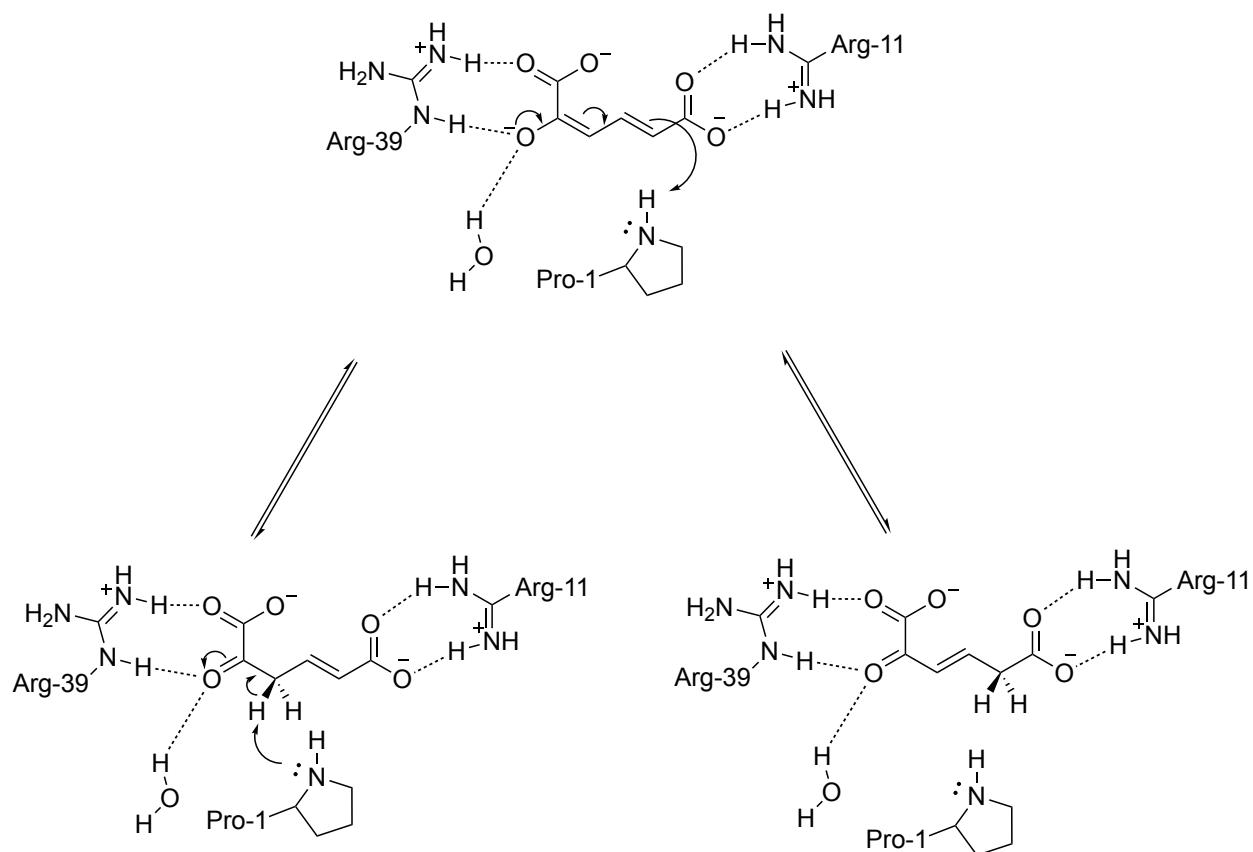


Figure 1.13. Mechanism of 4-OT utilization of 2-hydroxymuconate as a substrate for the tautomerization to 2-oxo-3-hexenedionate and 2-oxo-4-hexenedionate.

Table 1.1 Major reaction types catalyzed by enzymes in the tautomerase superfamily.

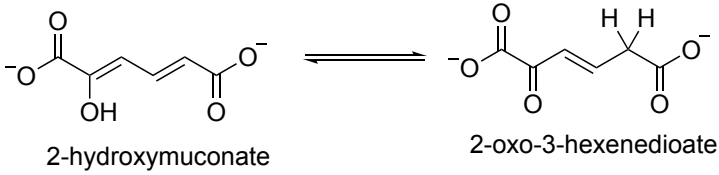
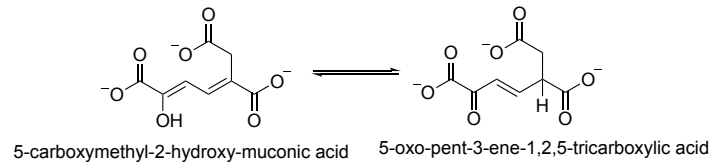
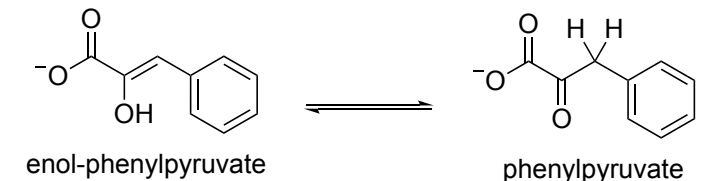
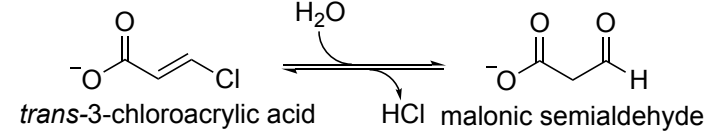
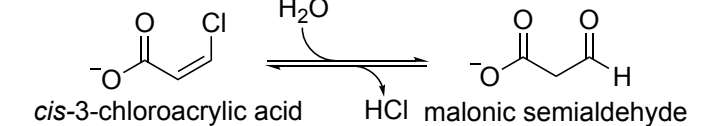
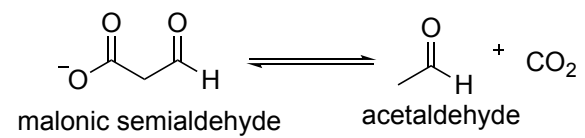
Enzyme Family	Typical Monomeric Length	Oligomeric State	Reaction Catalyzed
4-OT	62	Hexamer	 <p>2-hydroxymuconate \rightleftharpoons 2-oxo-3-hexenedioate</p>
CHMI	125	Trimer	 <p>5-carboxymethyl-2-hydroxy-muconic acid \rightleftharpoons 5-oxo-pent-3-ene-1,2,5-tricarboxylic acid</p>
MIF	114	Trimer	 <p>enol-phenylpyruvate \rightleftharpoons phenylpyruvate</p>
CaaD	α -subunit 75 β -subunit 70	Heterohexamer	 <p><i>trans</i>-3-chloroacrylic acid $\xrightleftharpoons[\text{HCl}]{\text{H}_2\text{O}}$ malonic semialdehyde</p>
<i>cis</i> -CaaD	149	Trimer	 <p><i>cis</i>-3-chloroacrylic acid $\xrightleftharpoons[\text{HCl}]{\text{H}_2\text{O}}$ malonic semialdehyde</p>

Table 1.1 (cont'd)

MSAD

129

Trimer



1.2.2 5-Carboxymethyl-2-hydroxymuconate Isomerase

5-Carboxymethyl-2-hydroxymuconate isomerase (CHMI) exists as a trimer and contains 125 amino acids per monomer (Figure 1.14B). Its metabolic pathway is in the degradation of muconate and catalyzes the tautomerization of 5-carboxymethyl-2-hydroxy-muconic acid into 5-oxo-pent-3-ene-1,2,5-tricarboxylic acid.^{65,74} This is then further broke down and incorporated into cellular metabolism. One such CHMI enzyme, muconate lactonizing enzyme is involved in ring cleavage and the metabolism of aromatic hydrocarbons. CHMI also has a large active site and is presumed to proceed with catalysis superficially involving proton transfer between Pro-1 and substrate.⁵⁸

1.2.3 Migration Inhibitory Factor

Similarly, macrophage migration inhibitory factor (MIF) also exists as trimer and is 114 amino acids in length (Figure1.14C).^{75,76} MIF is a multifunctional cytokine involved in immune system regulation by inhibiting macrophage migration, regulating inflammation and is a regulator of innate immunity involved in the formation of cancers and autoimmune diseases.⁷⁵ In addition, MIF is also known to have tautomerase activity on substrates such as phenylpyruvate.^{77,78} Its dual biological function has been carefully studied. No studies have reported that MIF catalysis proceeds by covalent interaction. However, it has been reported that Pro-1 serves as a catalytic base in tautomerase activity by Pro-1 nitrogen base extraction during catalysis with phenylpyruvate.⁷⁷ The low In addition, MIF is regulated by an allosteric site within a central solvent channel that affects conformation and activity.⁷⁶ This allows MIF to act as both a metabolic sensor and modulator, having the ability to regulate both immune and metabolic pathways by controlling tautomerase activity in responses to changes in the cellular environment. Additional studies have shown the C-terminal of the enzyme as critical for cytokine activity.⁷⁹

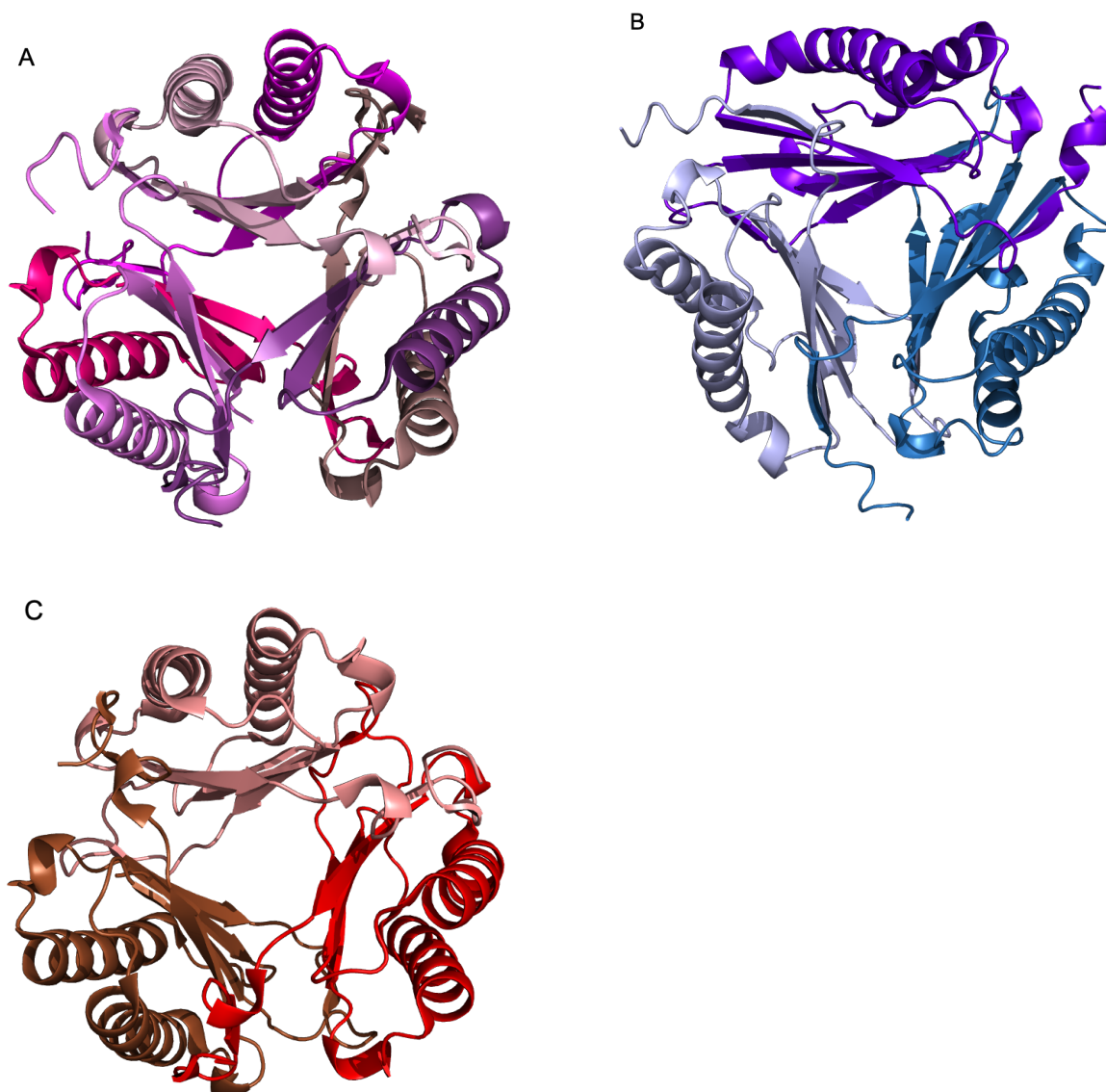


Figure 1.14. Tautomerase family enzymes (A) 4-OT from *B. lata* (PDB: 6OGM), (B) CHMI from *E. coli* (PDB: 1OTG) and (C) MIF from *P. yoelii* (PDB: 3GAD) colored by chain.^{69,74,75}

1.2.4 Malonate Semialdehyde Decarboxylase

Malonate semialdehyde decarboxylase (MSAD) from *Pseudomonas pavonacea* contains 129 amino acids and two β - α - β folds in a functional trimer (Figure 1.15). This enzyme is also a part of the degradation pathway of 1,3-dichloropropene, specifically in the conversion of MSA to ACH (Figure 1.16). The mechanism of MSAD is thought to be dependent on the high pK_a value of its

catalytic Pro-1 to facilitate decarboxylation through electrostatic interactions between Pro-1 and the substrate through hydrogen bonding.⁸⁰ For decarboxylase activity to occur the Pro-1 is at a pK_a of about 9.2 and polarizes the keto group. Pro-1, Asp-37, Arg-73, and Arg-75 are key residues involved in the active site for substrate catalysis. Arg-73 and Arg-75 are believed to position the carboxylate group into position within the active site and facilitate decarboxylation. The mechanism involves a hydrogen bonding water network between Pro-1 and Asp-37 that activates the carbonyl oxygen leading to an electron sink and consequential decarboxylation. Recent findings demonstrate covalent interactions between Pro-1 and substrate, but definitive mechanism remains unclear.⁸¹ In the crystallized structure of MSAD from *Corynebacterium bacterium* there is a glycine insertion that shifts the catalytic Arg-75 to residue 76. This species also lacks the active site residue Arg-73 (Figure 1.15). However, based on current findings of the *cis*-CaaD and Cg10062 mechanistic pathway, we propose that MSAD utilizes a covalent mechanism. MSAD assists in the kinetic product profile determination of both Cg10062 and *cis*-CaaD, providing an example of the multiple uses of these enzymes in bioengineering.

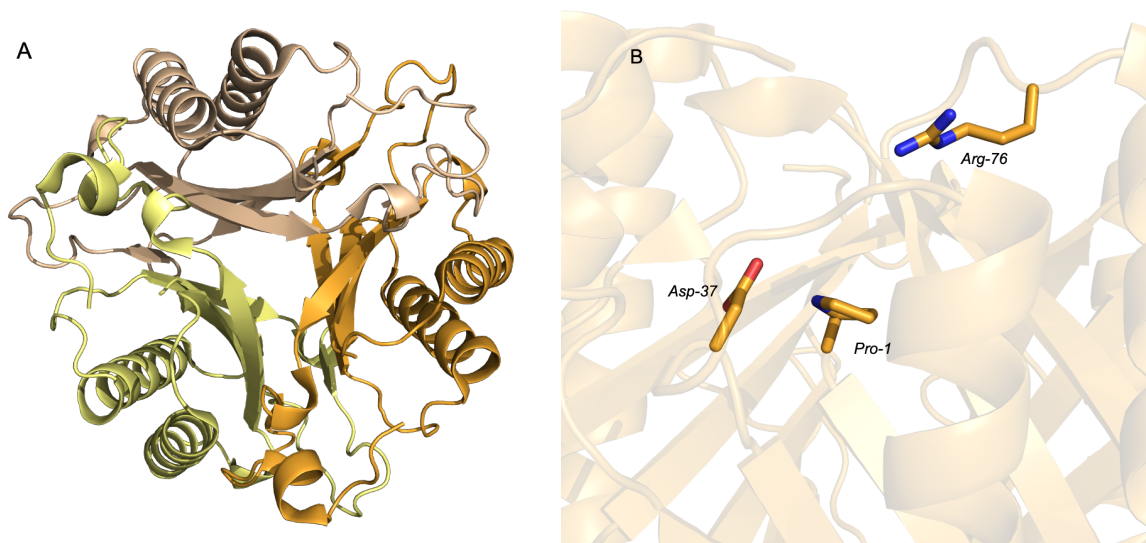


Figure 1.15. Tautomerase family enzyme MSAD from *C. bacterium* (PDB:3MJZ) (A) with crystal structure colored by chains and (B) active site residues.⁸⁰

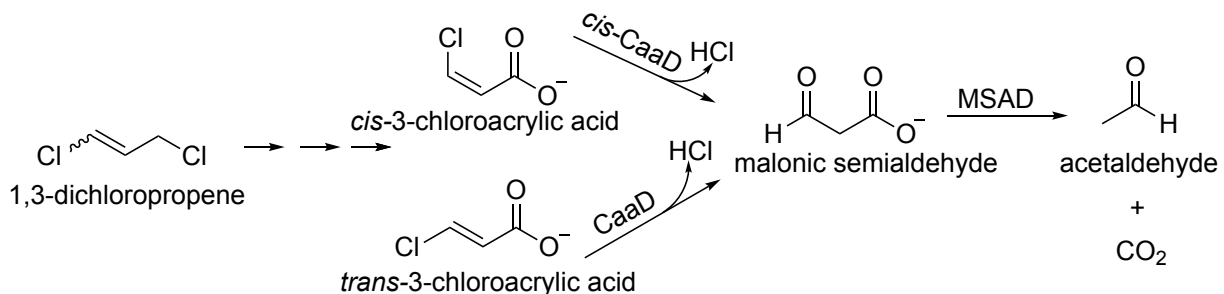


Figure 1.16. 1,3-Dichloropropene degradation pathway in *P. pavonacea*.

1.2.5. *trans*-3-Chloroacrylic Acid Dehalogenase

trans-CaaD (CaaD) was first identified in soil bacteria as part of the degradation pathway of 1,3-dichloropropene, an active ingredient in pesticides.^{65,82,83} CaaD catalyzes the conversion of *trans*-3-chloroacrylic acid to malonate semialdehyde (MSA) for further use in *Pseudomonas pavonacea*. Studies suggest CaaD catalyzes reactions with acetylene substrates as well.⁸⁴ CaaD contains one β - α - β fold yet it is a heterohexamer with only three active sites.⁸⁵ Structural studies show homology to that of 4-OT, suggesting that CaaD may have evolved from 4-OT through gene duplication and the evolution of an isoform, which is a similar protein that differs in genetic sequence.^{65,66,85,86} CaaD consist of two separate subunits, one containing 75 amino acids and the other with 70 amino acids, and their respective β - α - β folds (Figure 1.17). Each subunit retains the active site residues Pro-1, Arg-8, Arg-11, and Glu-52.^{85,86} The two arginine residues, Arg-8 and Arg-11, participate in substrate binding while Glu-52 activates a water molecule that facilitates the addition of water to the substrate. The catalytic Pro-1 of CaaD has a pK_a, of about 9.2 implicating that neighboring residue Glu-52 acts as the general base that activates the water molecule for attack on the substrate while Pro-1 acts as a general acid catalyst that provides protonation during tautomerization of the substrate.⁸¹



Figure 1.17. Tautomerase family enzyme CaaD from *P. pavonacea* (PDB: 3EJ3) colored by chains.⁸⁷

1.2.6 *cis*-3-Chloroacrylic Acid Dehalogenase

cis-CaaD which is a homotrimer in its functional form, contains 149 amino acids and two β - α - β in each monomer.^{88,89} Similar to CaaD, *cis*-CaaD catalyzes the dehydrohalogenation and tautomerization of *cis*-3-chloroacrylic acid to MSA. In addition to its native substrate, *cis*-CaaD has been shown to process alternative substrates such as allenes and acetylenes in *Pseudomonas pavonacea*, expanding its catalytic versatility.⁹⁰ A table depicting the tautomerase enzymes and the reactions they catalyze is shown in Table 1.1. Key residues involved in substrate active site include Pro-1, His-28, Arg-70, Arg-73, Tyr-103, and Glu-114, which play a critical role in substrate binding and catalysis.^{88,91} Of these residues, Tyr-103 is from a neighboring monomeric chain. Two other residues, Thr-32 and Thr-34, exist within an active loop and have previously drawn interest via their catalytic activity.^{92,93} The arginine residues and His-28 are thought to orientate binding of the substrate to the active site, while Glu-114 and Tyr-103 are suggested to be required to facilitate addition of a water molecule to the substrate. Previous mechanistic studies propose that the catalytic Pro-1, with a Pk_a of 9.3, acts as a general acid, donating a proton to the

substrate during tautomerization (Figure 1.18).⁹⁴⁻⁹⁶ However, no known mechanism for *cis*-CaaD has been supported by crystallographic evidence and confirmative intermediate trapping remains elusive.

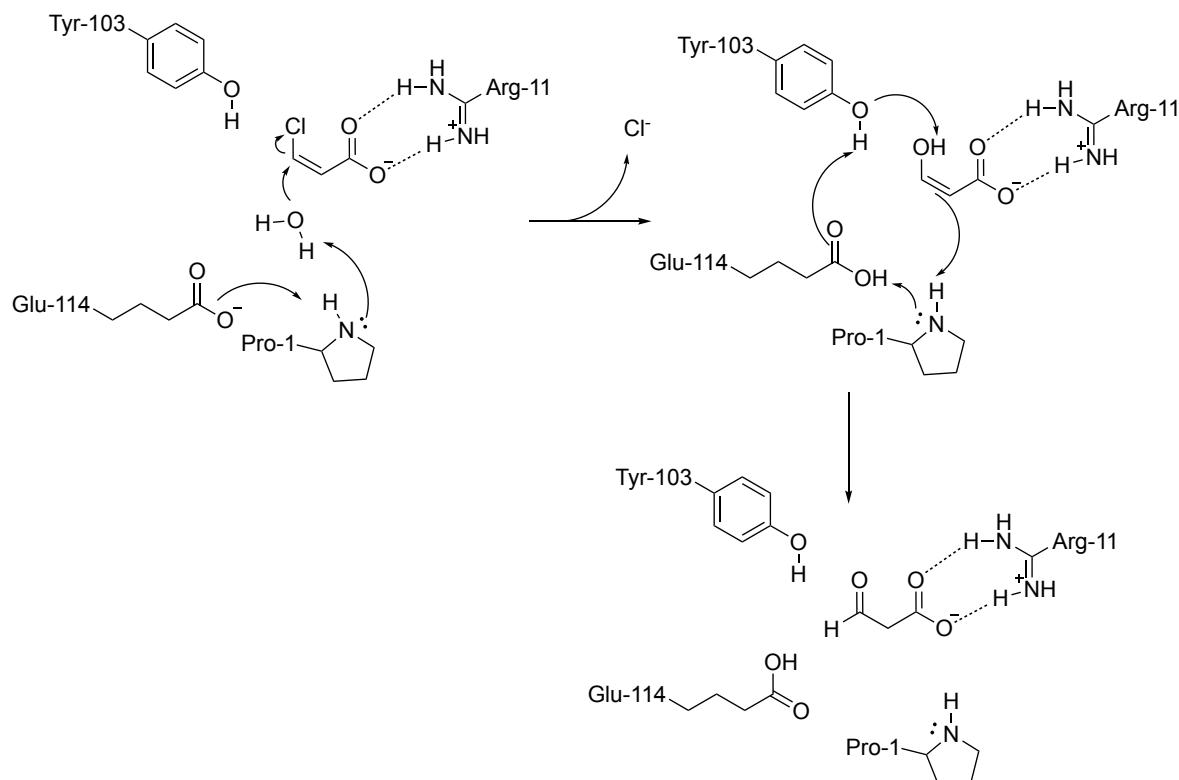


Figure 1.18. Previously understood *cis*-CaaD mechanism.⁹⁵

1.2.7 Cg10062

Previously we have explored the role of Cg10062 within the tautomerase superfamily. However, the exact origin and purpose of Cg10062 remains unclear. Cg10062, a homolog of *cis*-CaaD, is a homotrimer containing a characteristic β - α - β fold and *N*-terminal proline, consistent with the tautomerase superfamily. It is a homolog of *cis*-CaaD and shares 34% sequence identity and 54% sequence similarity.⁹⁷ Its monomer contains 147 amino acids and identical active site residues as that of *cis*-CaaD, Pro-1, His-28, Arg-70, Arg-73, Tyr-103, and Glu-114 (Figure 1.19). Where the native function of *cis*-CaaD in *Pseudomonas pavonacea* is known to be in the degradation of 1,3-dichloropropene, Cg10062 is not native to that process. Studies demonstrate

that Cg10062 is involved in the conversion of 2-oxo-3-pentynoate to acetopyruvate, but also shows activity with a variety of substrates including 2,3-dibutadienoate, 2-butyrate, *cis*-3-chloroacrylic acid (the native substrate of *cis*-CaaD), *trans*-3-chloroacrylic acid, and acetylenecarboxylic acid.^{90,98,99} Cg10062 exhibits low activity with *cis*-3-chloroacrylic acid suggesting a possible evolutionary link between Cg10062 and *cis*-CaaD, potentially arising from gene duplication followed by a fusion event producing our distinct protein oligomers within the tautomerase superfamily. This event may have produced a native isoform that subsequently led to the emergence of active enzymes. This may be the case of Cg10062 and *cis*-CaaD where Cg10062 was an active enzyme formed from an ancestral *cis*-CaaD isoform. Structural differences between *cis*-CaaD and Cg10062 include a 9-residue loop between the α -helix and second β -sheet in the β - α - β fold, where *cis*-CaaD reads RGLTGTQHF and Cg10062 reads HELAHAPKY.⁹⁹ Another structural difference is on a loop that directs towards the active site, which in *cis*-CaaD comprises residues Thr-32 and Thr-34 while the homologous position in Cg10062 are residues Ala-32 and Ala-34.⁹⁰ Previous studies describe the mechanism of both *cis*-CaaD and Cg10062 to be largely dependent on the pK_a of Pro-1, where Pro-1 acts as a proton donor and catalytic base. In the previously discussed *cis*-CaaD mechanism the direct attack of a water molecule on the substrate, followed by tautomerization, and finally decarboxylation and hydration yields product. Additionally in *cis*-CaaD, a covalently bound enamine has been shown which produces product either by hydrolysis or imine tautomerization.⁹⁴ This requires a pK_a ~9.3 for activity. The previously acknowledged mechanism for Cg10062 requires a Schiff base formation between Pro-1 and the substrate, water activation, and tautomerization to substrate.⁹⁷ However, based on crystallographic evidence of trapped covalent intermediates within the active site, the study presented in this dissertation suggests that both mechanisms of *cis*-CaaD and Cg10062 proceed by covalent catalysis.¹⁰⁰

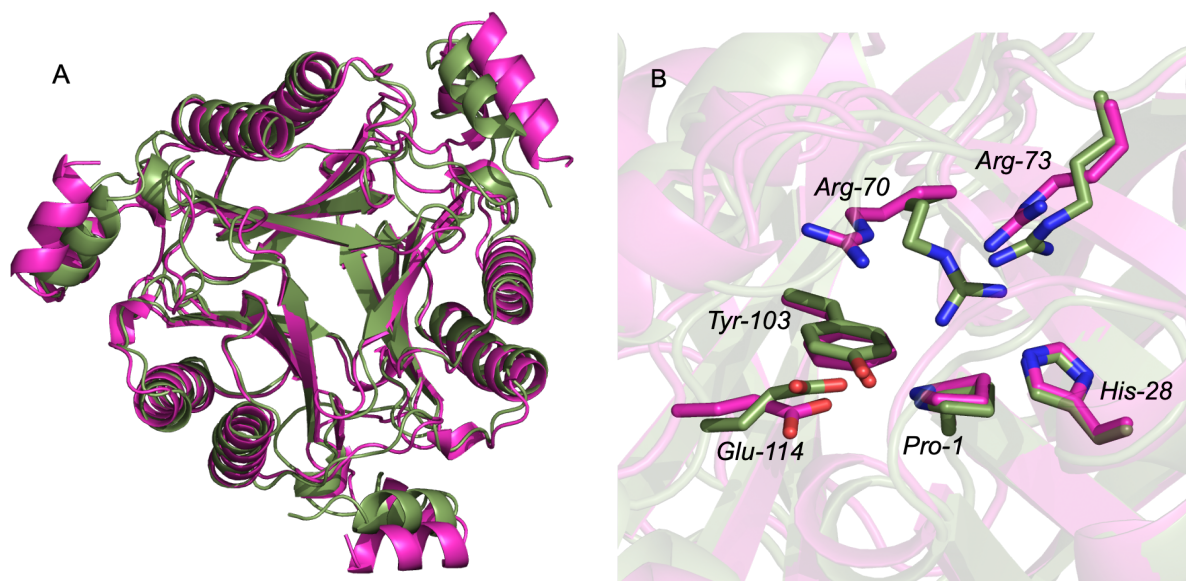


Figure 1.19. *cis*-CaaD (green, PDB: 9NFH) and Cg10062 (pink, PDB: 3N4G) crystal structure (A) trimers overlaid and (B) showing active site residues Pro-1, His-28, Arg-70, Arg-73, Tyr 103, and Glu-114.

1.3 *cis*-CaaD and Cg10062 Function and Purpose

A collaboration between the Geiger and Draths labs indicated a novel mechanism for Cg10062 using ACA as a substrate. In this Cg10062 pathway, nucleophilic attack by Pro-1 on ACA substrate leads to an enamine acrylate intermediate, which then tautomerizes to an iminium acrylate intermediate, which was formerly proposed by previous research groups.⁹⁰ Decarboxylation and hydration results in acetaldehyde (ACH) or sole hydration leads to 3-(*N*-propyl)-3-hydroxypropionate and MSA release.¹⁰⁰ However, both *cis*-CaaD and Cg10062 are able to utilize ACA as a substrate for the production of MSA, a precursor towards the production of value-added chemicals. MSA can be further converted into value added chemicals like malonic acid and HPA by other enzymes.¹⁰¹ *cis*-CaaD catalyzes the formation of MSA as the only observable product, while Cg10062 provides acetaldehyde (ACH, 85%) as the major product with low amounts of MSA (15%, Figure 1.20). However, Cg10062 is a more catalytically efficient enzyme when using ACA as a substrate. Variants of the active site residues, Pro-1, His-28, Arg-70, Arg-73, Tyr-103, and Glu-114, have been engineered to produce an enzyme that displays both

high conversion rates of ACA to MSA, and minimal production of ACH. Mutagenesis studies aimed at optimizing the conversion of ACA to MSA have emphasized the potential of Cg10062 and *cis*-CaaD in the biocatalysis for the production of high value chemicals from ACA. For these studies Cg10062 from *Corynebacterium glutamicum* and *cis*-CaaD from *Coryneform bacterium* designate FG41 were used. *C. glutamicum* is an organism increasingly used in biotechnology to develop high production strains and molecules such as amino acids.¹⁰² Both these species come from the *Coryneform* strain but are different particulate strains with unique features. MSAD, used in kinetic characterization studies, is also isolated from the unspecified FG41 designate. This common ancestral strain further indicates the evolutionary relationship between these enzymes. The enzymes originating from the same *Coryneform* genus have a similar mechanism, yet different designates, signifies evolutionary branching between *cis*-CaaD and Cg10062. Furthermore, there is a strong focus on Coryneforms for the large-scale production of chemicals, such as amino acids.¹⁰² Therefore, the utilization of tautomerases of designates from *Coryneform* are highly manipulatable and favorable in the biocatalytic production of high value chemicals.

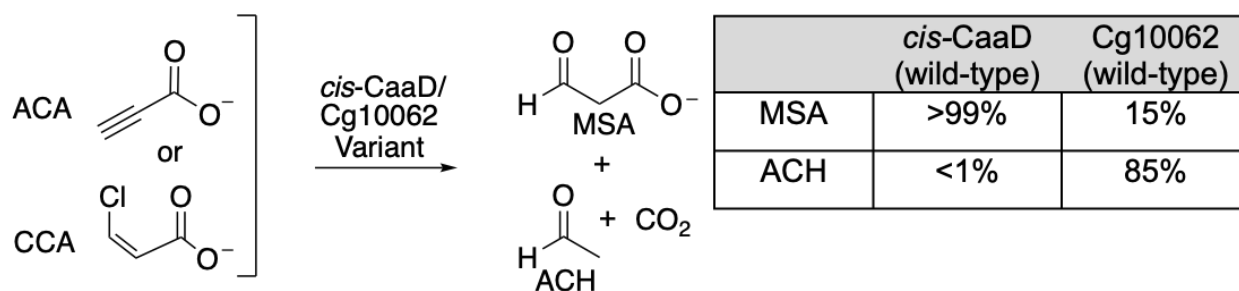


Figure 1.20. Scheme of the formation of MSA or ACH by *cis*-CaaD and Cg10062 and their wild-type product profiles.

Variants of *cis*-CaaD and Cg10062 were engineered to produce an enzyme that displays both high conversion rates of ACA to MSA, and minimal production of ACH. Both these wild type enzymes have identical active sites yet yield different products in their natural form (Figure 1.21). Structure analysis of these two enzymes gives an insight into the rational for the production of these different products.

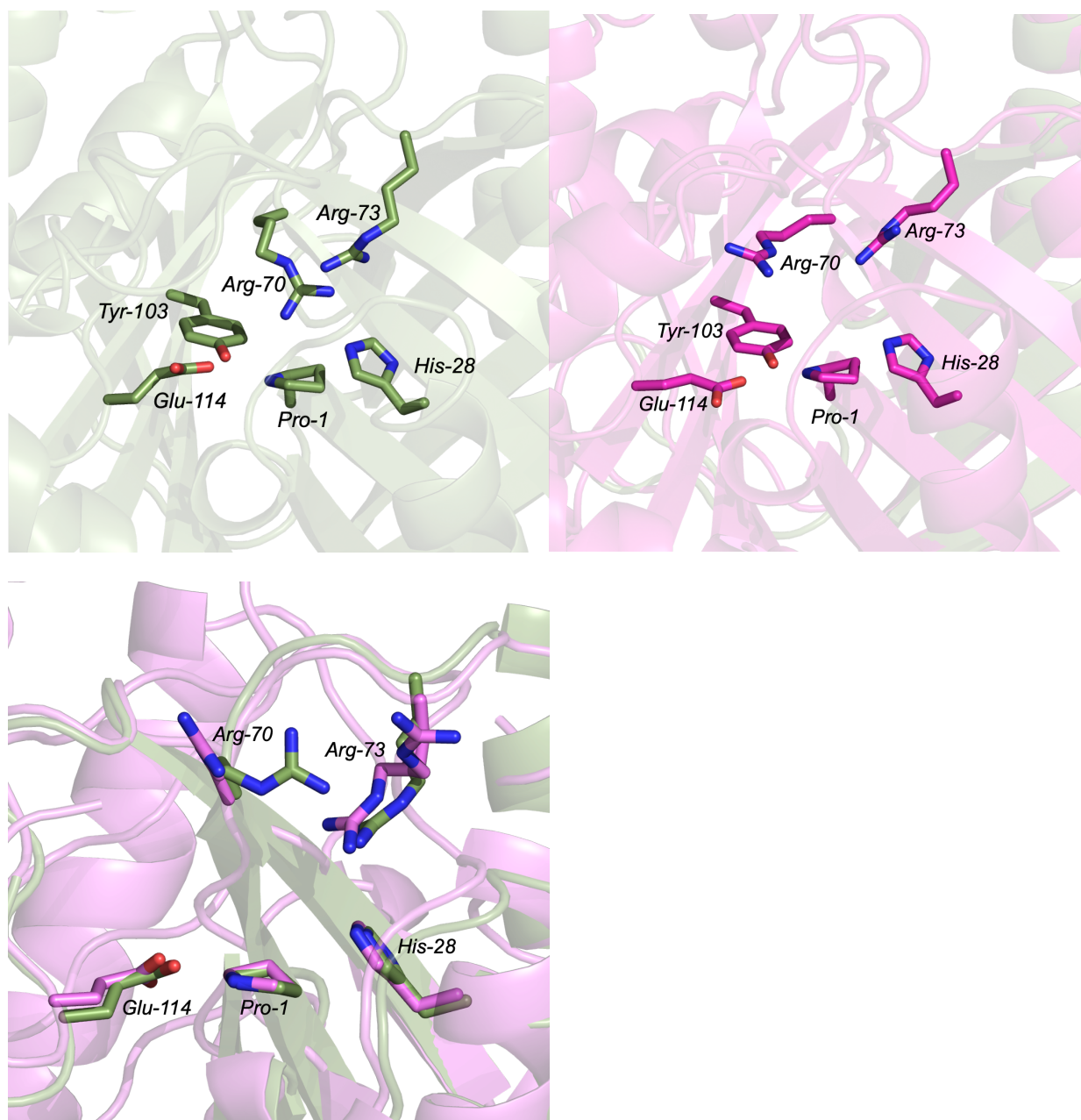


Figure 1.21. *cis*-CaaD (green, PDB: 9NFH) and Cg10062 (pink, PDB: 3N4G) crystal structures showing active site residues Pro-1, His-28, Arg-70, Arg-73, Tyr-103 and Glu-114.

1.3.1. Cg10062 (E114N) Towards the Production of 3-Hydroxypropionic Acid

ACA, a promising substrate in the biocatalysis of chemicals like 3HP, can be utilized by tautomerase such as *cis*-CaaD and Cg10062. These enzyme facilitate ACA's conversion into a valuable intermediate, MSA, that can then be converted to 3HP.^{101,104} In the identification of an

enzyme variant of either *cis*-CaaD or Cg10062 that resulted in high efficiency turnover of ACA into malonic semialdehyde (MSA), variant Cg10062 (E114N) was found.¹⁰¹ Cg10062 (E114N) retained high catalytic activity while converting ACA exclusively to MSA with undetectable quantities of acetaldehyde byproduct formation. Concurrent studies in the Draths lab have shown the role of Cg10062 (E114N) in the formation of 3HP *in vitro* as it is coupled with NADPH dependent enzyme YdfG from *E. coli* to produce 3HP.⁹⁵ This process brings incentive for biocatalytic synthesis of 3HP and other valuable chemicals from nonrenewable feedstocks

REFERENCES

1. Proved Reserves of Crude Oil and Natural Gas in the United States, Year-End 2022. <https://www.eia.gov/naturalgas/crudeoilreserves/> (accessed 2024-11-17).
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. Chen, S.; Perathoner, S.; Ampelli, C.; Centi, G. Chapter 2 - Electrochemical Dinitrogen Activation: To Find a Sustainable Way to Produce Ammonia. In *Studies in Surface Science and Catalysis*. 2019
4. Tickner, J.; Geiser, K.; Baima, S. Transitioning the Chemical Industry: The Case for Addressing the Climate, Toxics, and Plastics Crises. *Environ. Sci. Policy Sustainable Dev.* 2021, 63 (6), 4–15.
5. Vitro, I.; Models, E. V. *The impact of food bioactives on health*. 2015 <https://library.oapen.org/bitstream/handle/20.500.12657/28028/1001968.pdf?sequence=1#page=116> (accessed 2024-02-11).
6. April 22, 2024 Issue. Chemical & Engineering News. <https://cen.acs.org/magazine/102/10212.html> (accessed 2024-12-01).
7. Research Nester. Acrylonitrile Market Size & Share, Growth Trends 2037, 2024. <https://www.researchnester.com/reports/acrylonitrile-market/5973> (accessed 2025-02-07).
8. *Acrylic Acid Market Size, Share & Trends Analysis Report By Application (Acrylate Esters, Glacial Acrylic Acid), By End Use (Surfactants & Surface Coatings, Organic Chemicals), By Region, And Segment Forecasts, 2024 - 2030*. <https://www.grandviewresearch.com/industry-analysis/acrylic-acid-market> (accessed 2025-02-07).
9. *Acrylamide Market Size, Share, Growth*. <https://www.marketresearchfuture.com/reports/acrylamide-market-10300> (accessed 2025-02-07).
10. BP Chemicals Inc. The Sohio Acrylonitrile Process. *American Chemical Society Division of Chemistry and the office of Public Outreach*. 1996 <https://www.acs.org/content/dam/acsorg/education/whatischemistry/landmarks/acrylonitrile/sohio-acrylonitrile-process-commemorative-booklet-1996.pdf>. (accessed 2025-02-07).
11. Contractor, R. M.; Andersen, M. W.; Campos, D.; Hecquet, G.; Kotwica, R.; Pham, C.; Simon, M.; Stojanovic, M. Vapor Phase Oxidation of Propylene to Acrolein. 2002, 6437193. <https://patents.google.com/patent/US6437193B1/en> (accessed 2025-02-07)
12. Vanderkooi, W. N.; Jewett, G. L. Hydration of Acrylonitrile to Acrylamide. 1979, 4177210. <https://patents.google.com/patent/US4177210A/en> (accessed 2025-02-07)

13. Saini, R.; Osorio-Gonzalez, C. S.; Hegde, K.; Brar, S. K.; Magdouli, S.; Vezina, P.; Avalos-Ramirez, A. Lignocellulosic Biomass-Based Biorefinery: An Insight into Commercialization and Economic Standout. *Curr. Sustain./Renew. Energy Rep.* 2020, 7 (4), 122–136.
14. Peters, E. P.; Schlakman, G. J.; Yang, E. N. Production of Malonic Acid through the Fermentation of Glucose. *Scholarly Commons.* 2018.
15. Ioannidou, V.; Misailidis, N.; Petrides, D.; Georgiadis, M. C. Process Synthesis, Design and Techno-Economic Assessment of Malonic Acid Production. *Processes (Basel)* 2024, 12 (11), 2559.
16. Gu, S.; Zhao, Z.; Yao, Y.; Li, J.; Tian, C. Designing and Constructing a Novel Artificial Pathway for Malonic Acid Production Biologically. *Front. Bioeng. Biotechnol.* 2021, 9, 820507.
17. Devi, A.; Bajar, S.; Sihag, P.; Sheikh, Z. U. D.; Singh, A.; Kaur, J.; Bishnoi, N. R.; Pant, D. A Panoramic View of Technological Landscape for Bioethanol Production from Various Generations of Feedstocks. *Bioengineered.* 2023, 14 (1), 81–112.
18. Rosegrant, M. W.; Msangi, S. Consensus and Contention in the Food-versus-Fuel Debate. *Annu. Rev. Environ. Resour.* 2014, 39 (1), 271–294.
19. Pattnaik, F.; Patra, B. R.; Nanda, S.; Mohanty, M. K.; Dalai, A. K.; Rawat, J. Drivers and Barriers in the Production and Utilization of Second-Generation Bioethanol in India. *Recycling* 2024, 9 (1), 19.
20. Jadidi, Y., Frost H., Das, S., Liao, W., Draths, K., Saffron, C. Life cycle and technoeconomic comparisons of shikimic acid sourced from corn grain or corn stover. 2025, *in preparation*
21. İlhami, Y.; MacEachern, C. 1.2 Historical Aspects of Energy. In *Comprehensive Energy Systems*; Elsevier, 2018; pp 24–48.
22. Igwebuike, C. M.; Awad, S.; Andrès, Y. Renewable Energy Potential: Second-Generation Biomass as Feedstock for Bioethanol Production. *Molecules* 2024, 29 (7), 1619.
23. 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.
24. Henard, C. A.; Akberdin, I. R.; Kalyuzhnaya, M. G.; Guarnieri, M. T. Muconic Acid Production from CH₄ Using Rationally-Engineered Methanotrophic Biocatalysts. *Green Chem.* 2019, 21 (24), 6731–6737.
25. Lai, N.; Luo, Y.; Fei, P.; Hu, P.; Wu, H. One Stone Two Birds: Biosynthesis of 3-Hydroxypropionic Acid from CO₂ and Syngas-Derived Acetic Acid in *Escherichia coli*. *Synth. Syst. Biotechnol.* 2021, 6 (3), 144–152.
26. Stefan. *Ullmanns Encyclopedia of Industrial Chemistry Malonic Acid and Derivatives*. DOKUMEN.TIPS. <https://dokumen.tips/documents/ullmanns-encyclopedia-of-industrial-chemistry-malonic-acid-and-derivatives.html> (accessed 2022-03-18).

27. 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>.
28. Bozell, J. J.; Petersen, G. R. 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.
29. Dietrich, J.; Lygos, C. T. O. *Malonic Acid: A Bioadvantaged Chemical*. <https://lygos.com/app/uploads/2015/03/2014-12-Website-Flyer.pdf> (accessed 2021-12-09).
30. Xu, L.; Xiu, Y.; Liu, F.; Liang, Y.; Wang, S. Research Progress in Conversion of CO₂ to Valuable Fuels. *Molecules* 2020, 25 (16), 3653.
31. 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 CO₂. *European J. Org. Chem.* 2013, 2013 (10), 1867–1871.
32. Dinh, D. K.; Lee, D. H.; Song, Y.-H.; Jo, S.; Kim, K.-T.; Iqbal, M.; Kang, H. Efficient CH₄-to-Acetylene Conversion Using Low-Current Arcs. *RSC Adv.* 2019, 9 (56), 32403–32413.
33. Kasprzycka, A.; Lalak-Kańczugowska, J.; Walkiewicz, A.; Bulak, P.; Proc, K.; Stępień, Ł. Biocatalytic Conversion of CH₄ – Selected Aspects. *Curr. Opin. Chem. Eng.* 2019, 26, 28–32.
34. Jiang, C.; Hu, J. Conversion of CH₄ to Acetylene; 2023. <https://doi.org/10.1039/9781839160257-00093>.
35. Bedard, R. L.; Naunheimer, C.; Towler, G. P.; Leonard, L. E.; Dudebout, R.; Woodcock, G. O.; Mittendorf, D. L.; Morris, M. C.; Hausen, R. V.; MirzaMoghadam, A. CH₄ Conversion Apparatus and Process Using a Supersonic Flow Reactor. 2014. 20140058174:A1
36. Dinh, D. K.; Lee, D. H.; Song, Y.-H.; Jo, S.; Kim, K.-T.; Iqbal, M.; Kang, H. Efficient CH₄-to-Acetylene Conversion Using Low-Current Arcs. *RSC Adv.* 2019, 9 (56), 32403–32413.
37. Díaz Velázquez, H.; Wu, Z.-X.; Vandichel, M.; Verpoort, F. Inserting CO₂ into Terminal Alkynes via Bis-(NHC)-Metal Complexes. *Catal. Letters* 2017, 147 (2), 463–471.
38. Manjolinho, F.; Arndt, M.; Gooßen, K.; Gooßen, L. J. Catalytic C–H Carboxylation of Terminal Alkynes with CO₂. *ACS Catal.* 2012, 2 (9), 2014–2021.
39. Gooßen, L. J.; Rodríguez, N.; Manjolinho, F.; Lange, P. P. Synthesis of Propiolic Acids via Copper-Catalyzed Insertion of CO₂ into the C–H Bond of Terminal Alkynes. *Adv. Synth. Catal.* 2010, 352 (17), 2913–2917.
40. Steckhan, E. Electrochemistry, 3. Organic Electrochemistry. *Ullmann’s Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2011. https://doi.org/10.1002/14356007.o09_o04.

41. Jers, C.; Kalantari, A.; Garg, A.; Mijakovic, I. Production of 3-Hydroxypropanoic Acid From Glycerol by Metabolically Engineered Bacteria. *Front. Bioeng. Biotechnol.* 2019, 7. <https://doi.org/10.3389/fbioe.2019.00124>.
42. 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>.
43. Kumar, V.; Ashok, S.; Park, S. Recent Advances in Biological Production of 3-Hydroxypropionic Acid. *Biotechnol. Adv.* 2013, 31 (6), 945–961. <https://doi.org/10.1016/j.biotechadv.2013.02.008>.
44. Danchin, A. In Vivo, in Vitro and in Silico: An Open Space for the Development of Microbe-Based Applications of Synthetic Biology. *Microb. Biotechnol.* 2022, 15 (1), 42–64.
45. *Cargill cooperation*. <http://www.novozymes.com/en/news/news-archive/2008/01/44469> (accessed 2022-06-14).
46. Kissman, E. N.; Sosa, M. B.; Millar, D. C.; Koleski, E. J.; Thevasundaram, K.; Chang, M. C. Y. Expanding Chemistry through in Vitro and in Vivo Biocatalysis. *Nature* 2024, 631 (8019), 37–48.
47. Yang, F.; Hanna, M. A.; Sun, R. Value-Added Uses for Crude Glycerol--a Byproduct of Biodiesel Production. *Biotechnol. Biofuels* 2012, 5 (1), 13.
48. Wang, Z.; Zhuge, J.; Fang, H.; Prior, B. A. Glycerol Production by Microbial Fermentation: A Review. *Biotechnol. Adv.* 2001, 19 (3), 201–223.
49. 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.
50. 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.
51. 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.
52. 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.
53. Johnson, E. A.; Lin, E. C. *Klebsiella pneumoniae* 1,3-Propanediol: NAD⁺ Oxidoreductase. *J. Bacteriol.* 1987, 169 (5), 2050–2054.

54. 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.
55. 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.
56. 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.
57. 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.
58. 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.
59. 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.
60. Zhao, P.; Tian, P. Biosynthesis Pathways and Strategies for Improving 3-Hydroxypropionic Acid Production in Bacteria. *World J. Microbiol. Biotechnol.* 2021, 37 (7), 117.
61. 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.
62. 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.
63. 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.
64. Murzin, A. G. Structural Classification of Proteins: New Superfamilies. *Curr. Opin. Struct. Biol.* 1996, 6 (3), 386–394.
65. 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.
66. Poelarends, G. J.; Whitman, C. P. Evolution of Enzymatic Activity in the Tautomerase Superfamily: Mechanistic and Structural Studies of the 1,3-Dichloropropene Catabolic Enzymes. *Bioorg. Chem.* 2004, 32 (5), 376–392.

67. Huddleston, J. P.; Johnson, W. H.; Schroeder, G. K.; Whitman, C. P. The Accidental Assignment of Function in the Tautomerase Superfamily. *Perspectives in Science*. 2015, 4, 38–45.
68. Baas, B.-J.; Medellin, B. P.; LeVieux, J. A.; Erwin, K.; Lancaster, E. B.; Johnson, W. H., Jr; Kaoud, T. S.; Moreno, R. Y.; de Ruijter, M.; Babbitt, P. C.; Zhang, Y. J.; Whitman, C. P. Kinetic and Structural Analysis of Two Linkers in the Tautomerase Superfamily: Analysis and Implications. *Biochem*. 2021, 60 (22), 1776–1786.
69. Whitman, C. P. The 4-Oxalocrotonate Tautomerase Family of Enzymes: How Nature Makes New Enzymes Using a Beta-Alpha-Beta Structural Motif. *Arch. Biochem. Biophys*. 2002, 402 (1), 1–13.
70. Baas, B.-J.; Medellin, B. P.; LeVieux, J. A.; de Ruijter, M.; Zhang, Y. J.; Brown, S. D.; Akiva, E.; Babbitt, P. C.; Whitman, C. P. Structural, Kinetic, and Mechanistic Analysis of an Asymmetric 4-Oxalocrotonate Tautomerase Trimer. *Biochem*. 2019, 58 (22), 2617–2627.
71. Burks, E. A.; Fleming, C. D.; Mesecar, A. D.; Whitman, C. P.; Pegan, S. D. Kinetic and Structural Characterization of a Heterohexamer 4-Oxalocrotonate Tautomerase from *Chloroflexus Aurantiacus* J-10-FI: Implications for Functional and Structural Diversity in the Tautomerase Superfamily. *Biochem*. 2010, 49 (24), 5016–5027.
72. Harayama, S.; Rekik, M.; Ngai, K. L.; Ornston, L. N. Physically Associated Enzymes Produce and Metabolize 2-Hydroxy-2,4-Dienoate, a Chemically Unstable Intermediate Formed in Catechol Metabolism via Meta Cleavage in *Pseudomonas Putida*. *J. Bacteriol*. 1989, 171 (11), 6251–6258.
73. Whitman, C. P. The 4-Oxalocrotonate Tautomerase Family of Enzymes: How Nature Makes New Enzymes Using a Beta-Alpha-Beta Structural Motif. *Arch. Biochem. Biophys*. 2002, 402 (1), 1–13.
74. Subramanya, H. S.; Roper, D. I.; Dauter, Z.; Dodson, E. J.; Davies, G. J.; Wilson, K. S.; Wigley, D. B. Enzymatic Ketonization of 2-Hydroxymuconate: Specificity and Mechanism Investigated by the Crystal Structures of Two Isomerases. *Biochem*. 1996, 35 (3), 792–802.
75. Shao, D.; Zhong, X.; Zhou, Y.-F.; Han, Z.; Lin, Y.; Wang, Z.; Bu, L.; Zhang, L.; Su, X.-D.; Wang, H. Structural and Functional Comparison of MIF Ortholog from *Plasmodium Yoelii* with MIF from Its Rodent Host. *Mol. Immunol*. 2010, 47 (4), 726–737.
76. Pantouris, G.; Khurana, L.; Ma, A.; Skeens, E.; Reiss, K.; Batista, V. S.; Lisi, G. P.; Lolis, E. J. Regulation of MIF Enzymatic Activity by an Allosteric Site at the Central Solvent Channel. *Cell Chem Biol*. 2020, 27 (6), 740-750.e5.
77. Lubetsky, J. B.; Swope, M.; Dealwis, C.; Blake, P.; Lolis, E. Pro-1 of Macrophage Migration Inhibitory Factor Functions as a Catalytic Base in the Phenylpyruvate Tautomerase Activity. *Biochem*. 1999, 38 (22), 7346–7354.
78. Rosengren, E.; Aman, P.; Thelin, S.; Hansson, C.; Ahlfors, S.; Björk, P.; Jacobsson, L.; Rorsman, H. The Macrophage Migration Inhibitory Factor MIF Is a Phenylpyruvate Tautomerase. *FEBS Lett*. 1997, 417 (1), 85–88.

79. Kerschbaumer, R. J.; Rieger, M.; Völkel, D.; Le Roy, D.; Roger, T.; Garbaraviciene, J.; Boehncke, W.-H.; Müllberg, J.; Hoet, R. M.; Wood, C. R.; Antoine, G.; Thiele, M.; Savidis-Dacho, H.; Dockal, M.; Ehrlich, H.; Calandra, T.; Scheiflinger, F. Neutralization of Macrophage Migration Inhibitory Factor (MIF) by Fully Human Antibodies Correlates with Their Specificity for the β -Sheet Structure of MIF. *J. Biol. Chem.* 2012, 287 (10), 7446–7455.
80. Guo, Y.; Serrano, H.; Poelarends, G. J.; Johnson, W. H., Jr; Hackert, M. L.; Whitman, C. P. Kinetic, Mutational, and Structural Analysis of Malonate Semialdehyde Decarboxylase from Corynebacterium Bacterium Strain FG41: Mechanistic Implications for the Decarboxylase and Hydratase Activities. *Biochem.* 2013, 52 (28), 4830–4841.
81. Lancaster, E. B.; Hardtke, H. A.; Melkonian, T. R.; Venkat Ramani, M.; Johnson, W. H., Jr; Baas, B.-J.; Zhang, Y. J.; Whitman, C. P. Conversion of Inactive Non-Pro1 Tautomerase Superfamily Members into Active Tautomerases: Analysis of the Pro1 Mutants. *Biochemistry* 2025, 64 (4), 812–822.
82. 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. *Biochem.* 2004, 43 (3), 759–772.
83. van Hylckama Vlieg, J. E. T.; Janssen, D. B. Bacterial Degradation of 3-Chloroacrylic Acid and the Characterization of *cis*- and *trans*-Specific Dehalogenases. *Biodegradation.* 1991, 2 (3), 139–150.
84. Wang, S. C.; Person, M. D.; Johnson, W. H., Jr; Whitman, C. P. Reactions of *trans*-3-Chloroacrylic Acid Dehalogenase with Acetylene Substrates: Consequences of and Evidence for a Hydration Reaction. *Biochem.* 2003, 42 (29), 8762–8773.
85. Poelarends, G. J.; Almrud, J. J.; Serrano, H.; Darty, J. E.; Johnson, W. H., Jr; Hackert, M. L.; Whitman, C. P. Evolution of Enzymatic Activity in the Tautomerase Superfamily: Mechanistic and Structural Consequences of the L8R Mutation in 4-Oxalocrotonate Tautomerase. *Biochem.* 2006, 45 (25), 7700–7708.
86. Poelarends, G. J.; Saunier, R.; Janssen, D. B. *trans*-3-Chloroacrylic Acid Dehalogenase from Pseudomonas Pavonaceae 170 Shares Structural and Mechanistic Similarities with 4-Oxalocrotonate Tautomerase. *J. Bacteriol.* 2001, 183 (14), 4269–4277.
87. Pegan, S. D.; Serrano, H.; Whitman, C. P.; Mesecar, A. D. Structural and Mechanistic Analysis of *trans*-3-Chloroacrylic Acid Dehalogenase Activity. *Acta Crystallogr. D Biol. Crystallogr.* **2008**, 64 (Pt 12), 1277–1282.
88. de Jong, R. M.; Bazzacco, P.; Poelarends, G. J.; Johnson, W. H., Jr; Kim, Y. J.; Burks, E. A.; Serrano, H.; Thunnissen, A.-M. W. H.; Whitman, C. P.; Dijkstra, B. W. Crystal Structures of Native and Inactivated *cis*-3-Chloroacrylic Acid Dehalogenase. Structural Basis for Substrate Specificity and Inactivation by (R)-Oxirane-2-Carboxylate. *J. Biol. Chem.* 2007, 282 (4), 2440–2449.

89. Guo, Y.; Serrano, H.; Johnson, W. H., Jr; Ernst, S.; Hackert, M. L.; Whitman, C. P. Crystal Structures of Native and Inactivated *cis*-3-Chloroacrylic Acid Dehalogenase: Implications for the Catalytic and Inactivation Mechanisms. *Bioorg. Chem.* 2011, 39 (1), 1–9.
90. Schroeder, G. K.; Johnson, W. H., Jr; Huddleston, J. P.; Serrano, H.; Johnson, K. A.; Whitman, C. P. Reaction of *cis*-3-Chloroacrylic Acid Dehalogenase with an Allene Substrate, 2,3-Butadienoate: Hydration via an Enamine. *J. Am. Chem. Soc.* 2012, 134 (1), 293–304.
91. LeVieux, J. A.; Baas, B.-J.; Kaoud, T. S.; Davidson, R.; Babbitt, P. C.; Zhang, Y. J.; Whitman, C. P. Kinetic and Structural Characterization of a *cis*-3-Chloroacrylic Acid Dehalogenase Homologue in *Pseudomonas* sp. UW4: A Potential Step between Subgroups in the Tautomerase Superfamily. *Arch. Biochem. Biophys.* 2017, 636, 50–56
92. Robertson, B. A.; Schroeder, G. K.; Jin, Z.; Johnson, K. A.; Whitman, C. P. Pre-Steady-State Kinetic Analysis of *cis*-3-Chloroacrylic Acid Dehalogenase: Analysis and Implications. *Biochem.* 2009, 48 (49), 11737–11744.
93. Schroeder, G. K.; Huddleston, J. P.; Johnson, W. H., Jr; Whitman, C. P. A Mutational Analysis of the Active Site Loop Residues in *cis*-3-Chloroacrylic Acid Dehalogenase. *Biochem.* 2013, 52 (24), 4204–4216.
94. Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, W. 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. *Biochem.* 2004, 43 (3), 759–772.
95. Sevastik, R.; Whitman, C. P.; Himo, F. Reaction Mechanism of *cis*-3-Chloroacrylic Acid Dehalogenase: A Theoretical Study. *Biochem.* 2009, 48 (40), 9641–9649.
96. Huddleston, J. P.; Wang, S. C.; Johnson, K. A.; Whitman, C. P. Resolution of the Uncertainty in the Kinetic Mechanism for the *trans*-3-Chloroacrylic Acid Dehalogenase-Catalyzed Reaction. *Arch. Biochem. Biophys.* 2017, 623–624, 9–19.
97. Baas, B.-J.; Zandvoort, E.; Wasie, A. A.; Quax, W. J.; Poelarends, G. J. Characterization of a Newly Identified Mycobacterial Tautomerase with Promiscuous Dehalogenase and Hydratase Activities Reveals a Functional Link to a Recently Diverged *cis*-3-Chloroacrylic Acid Dehalogenase. *Biochem.* 2011, 50 (14), 2889–2899.
98. Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, W. H., Jr; Whitman, C. P. Characterization of Cg10062 from *Corynebacterium glutamicum*: Implications for the Evolution of *cis*-3-Chloroacrylic Acid Dehalogenase Activity in the Tautomerase Superfamily. *Biochem.* 2008, 47 (31), 8139–8147.
99. Huddleston, J. P.; Johnson, W. H., Jr; 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.

100. Mathes Hewage, A.; Nayeibi Gavgani, H.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochem.* 2021, 60 (51), 3879-3886. <https://doi.org/10.1021/acs.biochem.1c00524>.
101. Mathes Hewage, A.; Silva, K.; Kwiatkowski, K.; Lee, H.; Sreedhar, D.; Gavgani, H. Geiger, J., Draths, K. An Original Biosynthetic Route to 3-Hydroxypropionic Acid, **2025**, *in preparation*
102. Ikeda, M.; Nakagawa, S. The *Corynebacterium glutamicum* Genome: Features and Impacts on Biotechnological Processes. *Appl. Microbiol. Biotechnol.* 2003, 62 (2–3), 99–109.
103. Schwentner, A.; Neugebauer, H.; Weinmann, S.; Santos, H.; Eikmanns, B. J. Exploring the Potential of *Corynebacterium glutamicum* to Produce the Compatible Solute Mannosylglycerate. *Front. Bioeng. Biotechnol.* 2021, 9, 748155.)
104. Silva, K.; Al-Ahmad, N.; Geiger, J. H.; Draths, K. Elucidating Structural and Activity Differences Between Tautomerase *cis*-CaaD and Cg10062: Implications For 3-Hydroxypropionic Acid Production From Acetylenecarboxylic Acid *Biochem.* 2025. *in preparation*

**CHAPTER TWO: Comparison of *cis*-CaaD and Cg10062 Catalyzed Transformation
of Acetylenecarboxylic Acid**

2.1. *cis*-CaaD Engineering

Due to the tautomerases', *cis*-CaaD and Cg10062, β - α - β building block motif, small size, and ability to function without cofactor, they can be easily manipulated through protein engineering. Based on literature assessments and structural comparisons of *cis*-CaaD with Cg10062 and Cg10062 variants previously studied, several residues were identified for specific point mutations in *cis*-CaaD. The *cis*-CaaD mutations involved were E114N, E114D, E114Q, H28A, R73A, Y103F, Y103A, T32A, and T34A. Active site residues were selected to evaluate their role in catalysis, while similar mutations in Cg10062 were previously found to have an impact on activity and product production. Protein design was based on analysis of enzyme structure and these active site residues. These amino acids were targeted based on polarity, pK_a , and hydrogen bonding distance between active site residues and substrate. Previously, we recognized His-28, Glu-114, Tyr-103, Thr-32, and Thr-34 to be crucial for activity in *cis*-CaaD.^{1,2} Alanine variants were synthesized to further evaluate their influence on product formation and activity. Tyr-103 is understood to play a role in the hydrogen bonding network between the catalytic water molecule that is added to substrate in the active site. Substitution of Tyr-103 for phenylalanine was implemented to remove hydrogen bonding ability of the side chain. Similarly, Glu-114 was mutated to affect the hydrogen bonding distance between residue-114 and a catalytic water molecule involved in the hydration of substrate to malonic semialdehyde (MSA). Aspartate was substituted for Glu-114 to determine if shortening the side chain of the amino acid that coordinates the active water will affect activity by theoretically pulling the water molecule farther away from the active site. Glu-114 was also exchanged for glutamine to determine the effect of changing the charge of this residue from negative to neutral. Combining these ideas, asparagine was also substituted at Glu-114 to understand the mechanism when there is a short uncharged amino acid in the water coordination site.

cis-CaaD from *Corynebacterium* was codon optimized for expression in *E. coli* and cloned into the commercial pET-21a(+) vector containing a C-terminal His₆-tag. A modified TEV protease recognition site was inserted before the His₆-tag to enable removal of the His₆-tag if it were to interfere with enzyme activity or crystallization efforts. Removal was not necessary as the His₆-tag had no noticeable effect on activity or crystallization. The N-terminal was left unmodified to allow activity of the catalytic Pro-1 residue. The TEV recognition site was included so that the His₆ affinity tag may be removed by TEV protease cleavage if the affinity tag hindered catalysis or protein crystallization. The affinity tag did not disrupt catalysis or crystallization efforts. Therefore, the modified C-terminal was left intact. Q5 site-directed mutagenesis was used to implement individual point mutations. Further structural analysis of the *cis*-CaaD structure revealed a second loop of interest beginning at residue-120 in both *cis*-CaaD and Cg10062. Chimeras of these two enzymes as well as a truncated *cis*-CaaD were designed at residue-120 (*N-cis*-CaaD/*C*-Cg10062, *N*-Cg10062/*C-cis*-CaaD, and *tcis*-CaaD). These studies were performed to evaluate the effect of the C-terminal on active site movement and activity. A gene chimera consisting of amino acids 1-119 of *cis*-CaaD followed by amino acids 120-148 of Cg10062 was designed including NdeI and XhoI restriction sites, respectively, at the 5' and 3' positions. A gene chimera consisting of amino acids 1-119 of Cg10062 followed by amino acids 120-149 of *cis*-CaaD was designed similarly. Gene chimeras were cloned into pET-21a(+) between NdeI and XhoI restriction sites. A truncation of *cis*-CaaD consisting of amino acids 1-119 was also designed. HiFi Gibson assembly was used to construct chimeras and truncated *cis*-CaaD. Plasmid DNA was confirmed with Sanger sequencing and transformed into BL21(DE3) *E. coli* cells for protein expression and purification. Protein yield was typically 2 g L⁻¹.

2.1.1. *cis*-CaaD Variant Product Profile and Kinetic Activity

Kinetic studies were performed on these variants and wild-type *cis*-CaaD to obtain product profiles and Michaelis-Menten kinetics where possible. A coupled enzyme assay to determine product profiles for enzymes and their variants includes the use of MSAD and alcohol

dehydrogenase (ADH, Figure 2.1).^{3,4} MSAD from *Corynebacterium* bacterium designate FG41 was modified with a TEV recognition sequence and cloned into pET-21a(+) at NdeI and XhoI restriction sites. MSAD decarboxylates MSA to ACH.⁵ ACH can then be converted to ethanol by ADH at the expense of NADH. Performing the reaction in the presence of MSAD, all MSA produced by *cis*-CaaD or variant will first be converted to ACH. The ACH produced by the *cis*-CaaD or variant, as well as the MSA that has been converted to ACH will be converted to ethanol by ADH. Without MSAD, any ACH produced by the *cis*-CaaD or variant evaluated will be directly converted to ethanol by ADH, leaving MSA in solution undetected. Activity was measured on a UV-Vis spectrophotometer by monitoring the oxidation of NADH by ADH at 340 nm. By calculating the differences between activity with and without MSAD, we can determine product profiles for the variant being studied (Figure 2.2). Comparison of the ratios of product formed between the variants gives insight into which variants can contribute to an improved catalytic efficiency to produce MSA. All coupling enzymes were used in excess to ensure the rate of activity measured was that of *cis*-CaaD or variant. Substrates used in these assays are *cis*-3-chloroacrylic acid (CCA) and acetylenecarboxylic acid (ACA). Native substrate, CCA, was used in kinetic studies as a reference substrate for natural activity, and in intermediate trapping experiments to determine the mechanism of *cis*-CaaD.

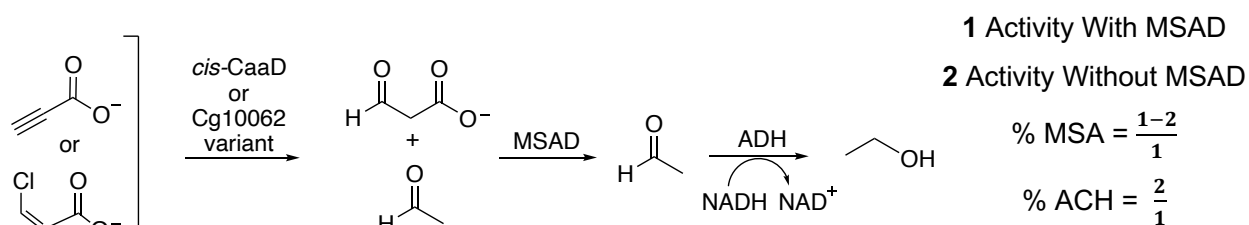


Figure 2.1. Coupled enzyme assay used in the evaluation of *cis*-CaaD and *cis*-CaaD variant activities.

Steady-state kinetics were carried out in triplicate at 25 °C in 100 mM sodium phosphate pH 8.0 with a final volume of 1 mL. Enzyme activity was measured using the same coupled reaction described previously. The reaction was monitored by following the oxidation of NADH at

340nm ($\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$) by the reduction of acetaldehyde by NADH-dependent alcohol dehydrogenase (ADH). Of the enzyme variants examined, steady state kinetics indicate wild-type *cis*-CaaD remains more active with ACA substrate than any of the *cis*-CaaD variants (Table 2.1). The initial rates of the *cis*-CaaD variants, relative to varied ACA concentrations were plotted to fit the Michaelis-Menten model and analyzed (Figure 2.3).

A Enzyme Variant	ACA		CCA	
	Product Ratio (%)		Product Ratio (%)	
	MSA	ACH	MSA	ACH
<i>cis</i> -CaaD (wild-type)	>99 \pm 1.0	<1 \pm 0.0	>99 \pm 1.0	<1 \pm 0.0
<i>cis</i> -CaaD (E114D)	>99 \pm 1.0	<1 \pm 0.0	>99 \pm 0.9	<1 \pm 0.1
<i>cis</i> -CaaD (E114N)	35 \pm 0.5	65 \pm 0.5	43 \pm 0.4	57 \pm 0.6
<i>cis</i> -CaaD (E114Q)	12 \pm 1.1	88 \pm 2.0	13 \pm 0.4	87 \pm 0.6
<i>cis</i> -CaaD (Y103F)	>99 \pm 0.4	<1 \pm 0.6	96 \pm 0.2	4 \pm 1.2
<i>cis</i> -CaaD (Y103A)	>99 \pm 1.0	<1 \pm 0.0	98 \pm 0.8	2 \pm 0.1
<i>cis</i> -CaaD (H28A)	-	-	-	-
<i>cis</i> -CaaD (R73A)	-	-	-	-
<i>cis</i> -CaaD (T32A)	>99 \pm 0.5	<1 \pm 0.5	>99 \pm 0.8	<1 \pm 0.2
<i>cis</i> -CaaD (T34A)	>99 \pm 0.9	<1 \pm 0.1	98 \pm 0.4	2 \pm 0.6
Cg10062 (wild-type)	15 \pm 0.6	85 \pm .04	14 \pm 0.8	86 \pm 0.2

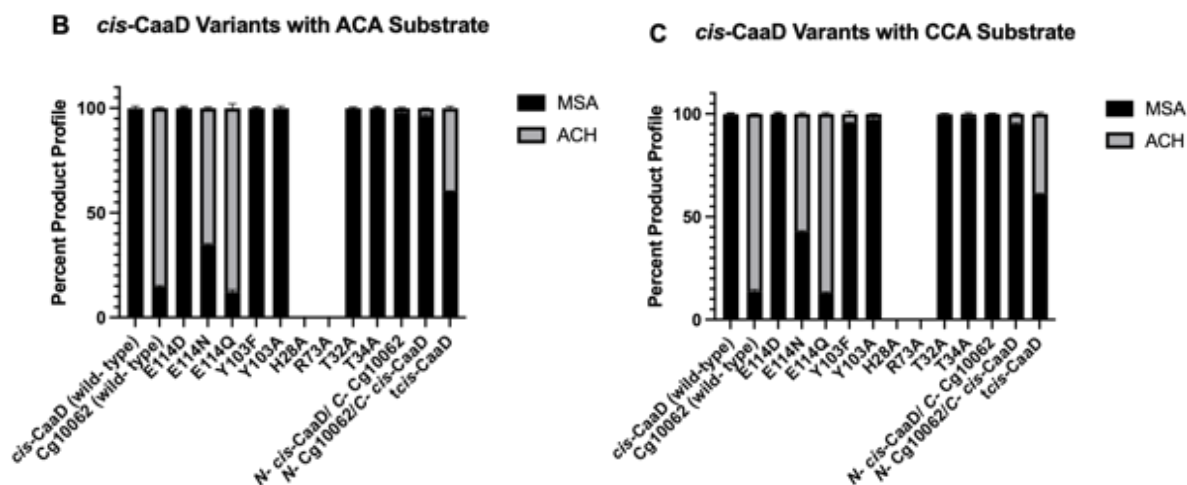


Figure 2.2. Product ratios of variants and *cis*-CaaD with (A) values shown in a table, (B) graphed product ratios for ACA substrate and (C) CCA substrate. (Those signified with – indicate too low activity to determine product ratios)

Table 2.1. Michaelis-Menten values for *cis*-CaaD (wild-type), *cis*-CaaD (T32A), and *cis*-CaaD (T34A) using ACA and CCA substrates. Activity for other variants had insufficient activity to accurately determine steady state kinetics.

Enzyme Variant	ACA			CCA		
	K_m (mM)	k_{cat} (s ⁻¹)	$k_{cat}/K_m \times 10^4$ (M ⁻¹ s ⁻¹)	K_m (mM)	k_{cat} (s ⁻¹)	$k_{cat}/K_m \times 10^4$ (M ⁻¹ s ⁻¹)
<i>cis</i> -CaaD (wild-type)	1050 ± 109	1.40 ± 0.05	0.13 ± 0.01	372 ± 41	2.16 ± 0.07	0.58 ± 0.05
<i>cis</i> -CaaD (T32A)	-	-	-	3051 ± 602	2.04 ± 0.16	0.07 ± 0.01
<i>cis</i> -CaaD (T34A)	1730 ± 199	0.59 ± 0.03	0.03 ± 0.00	990 ± 136	3.57 ± 0.18	0.36 ± 0.03
<i>N</i> -Cg10062/ <i>C</i> - <i>cis</i> -CaaD	1330 ± 363	1.02 ± 0.11	0.08 ± 0.01	-	-	-
Cg10062 (wild-type)	66 ± 17	4.2 ± 0.28	6.3 ± 1.2	15750 ± 2940	4.13 ± 0.29	0.03 ± 0.00

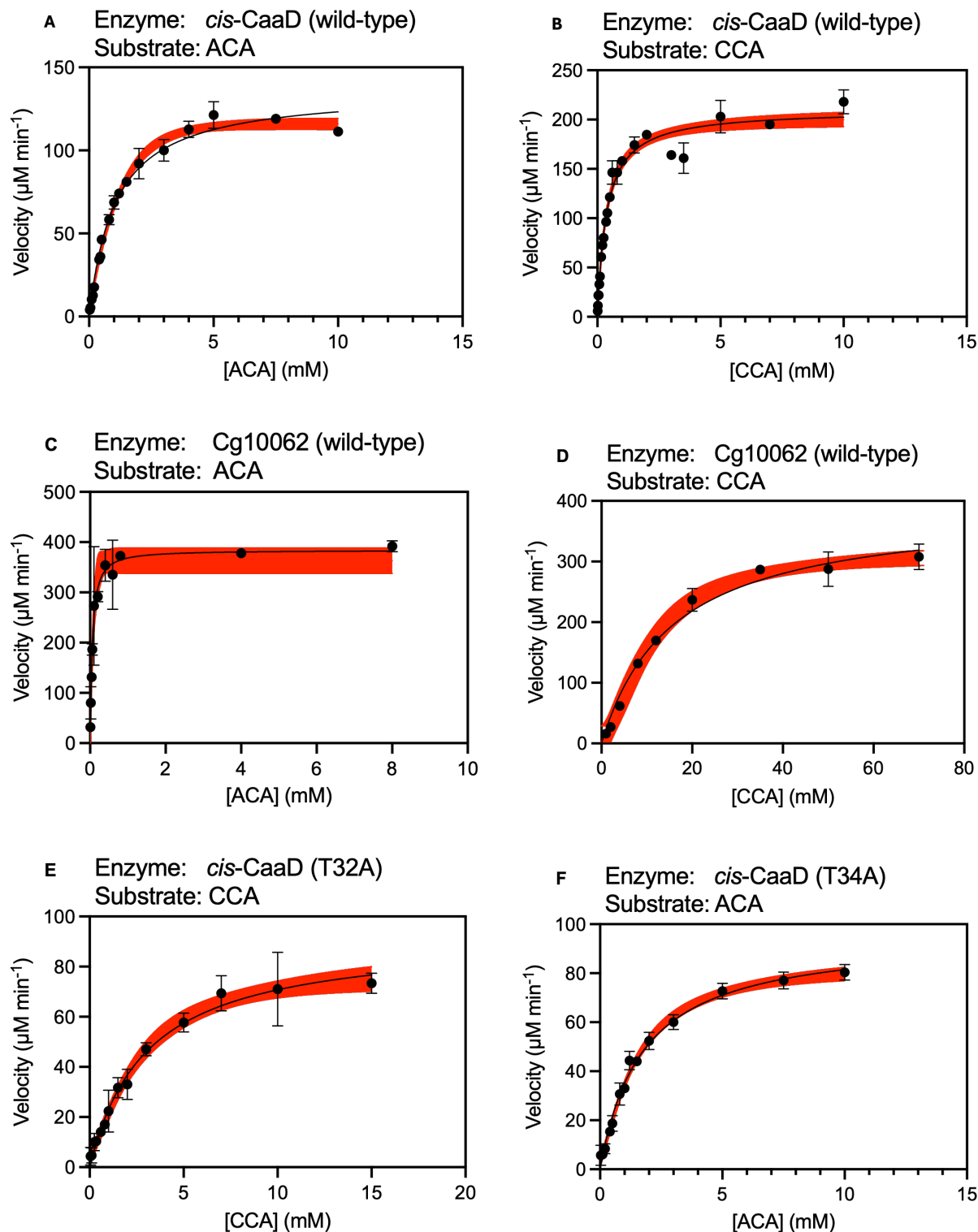
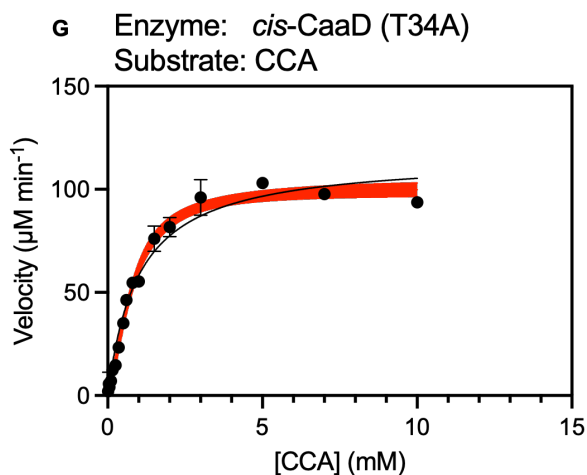
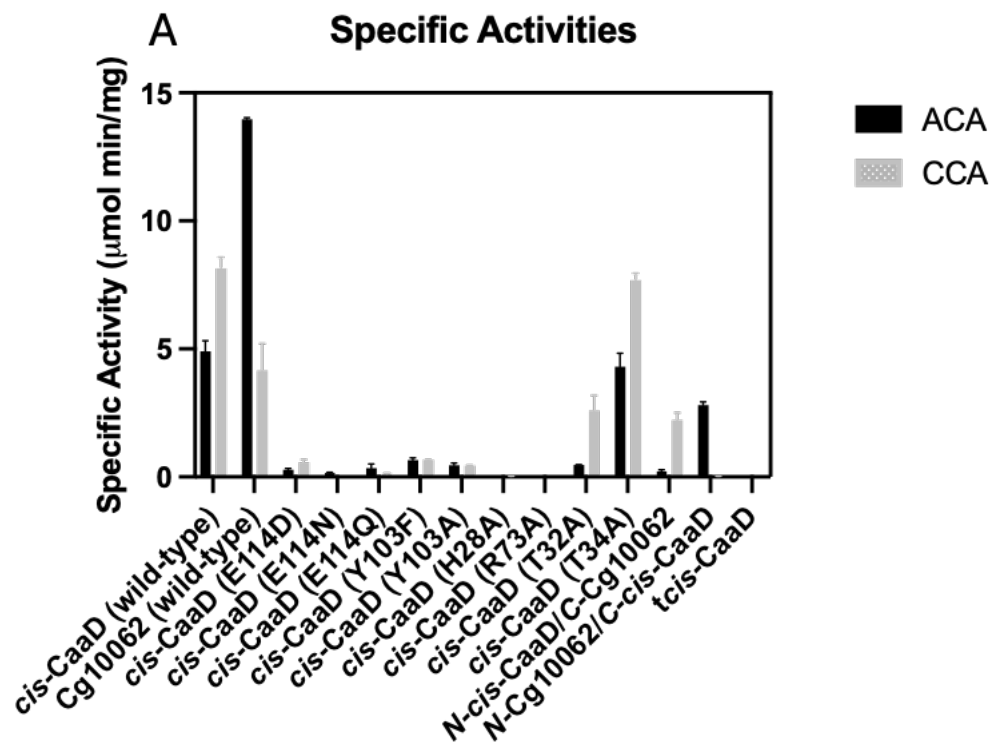


Figure 2.3. Steady state kinetics of (A) *cis*-CaaD (wild-type) with substrate ACA, (B) *cis*-CaaD (wild-type) with substrate CCA, (C) Cg10062 (wild-type) with ACA, (D) Cg10062 (wild-type) with CCA, (E) *cis*-CaaD (T32A) with CCA, (F) *cis*-CaaD (T34A) with ACA, and (G) *cis*-CaaD (T34A) with CCA.

Figure 2.3. (cont'd)



The specific activity of each variant was determined. Most variants of *cis*-CaaD exhibited lower specific activities than wild-type *cis*-CaaD, with *cis*-CaaD (T34A) exhibiting comparable activity to that of wild-type *cis*-CaaD. To maintain accuracy in measuring activity, varying concentrations of *cis*-CaaD variants were used (Figure 2.4). While wild-type *cis*-CaaD produces primarily MSA, in variants *cis*-CaaD (E114N) and *cis*-CaaD (E114Q), ACH is the favored product for both substrates ACA and CCA. These *cis*-CaaD variants exhibit similar product formation ratios to that of wild-type Cg10062. The appreciable loss in activity from variants of residues E114D, E114N, E114Q, Y103F, Y103A, H28A, and R73A confirm their critical role in catalysis.



B	Enzyme	[ACA] (mM)	[CCA] (mM)	[Enzyme] (mM)
	<i>cis</i> -CaaD (wild-type)	1.0	1.0	0.60
	Cg10062 (wild-type)	4.0	50	1.7
	<i>cis</i> -CaaD (E114D)	1.0	1.0	0.60
	<i>cis</i> -CaaD (E114N)	1.0	1.0	0.60
	<i>cis</i> -CaaD (E114Q)	1.0	1.0	0.60
	<i>cis</i> -CaaD (Y103F)	1.0	1.0	0.60
	<i>cis</i> -CaaD (Y103A)	2.0	2.0	2.1
	<i>cis</i> -CaaD (H28A)	2.0	2.0	15
	<i>cis</i> -CaaD (R73A)	2.0	2.0	22
	<i>cis</i> -CaaD (T32A)	2.0	2.0	0.84
	<i>cis</i> -CaaD (T34A)	2.0	2.0	0.60
	<i>N-cis</i> -CaaD/C-Cg10062	2.0	2.0	7.2
	<i>N-Cg10062/C-cis</i> -CaaD	2.0	2.0	1.2
	<i>tcis</i> -CaaD	2.0	2.0	290

Figure 2.4. Specific activities (A) of *cis*-CaaD (wild-type) and its variants with Cg10062 (wild-type) using substrates CCA and ACA. Differing substrate concentrations (B) were used to ensure saturation of substrate to enzyme.

Product formation was confirmed using ^1H NMR analysis (Figure 2.5). Reactions were prepared with substrate and initiated with the addition of *cis*-CaaD or variant and allowed to react. Aliquots were quenched with the addition of 5 M H_2SO_4 (2 μL) and timepoints were taken over 16 hours. Product formation was confirmed. 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. The resonance at δ 6.3 (d, 1H) and 6.4 (d, 1H) corresponds to CCA. Resonances at δ 3.20 (d, 2H), δ 9.50 (t, 1H) and δ 2.30 (d, 2H), 5.13 (t, 1H) correspond to malonate semialdehyde and its hydrate, respectively. Resonances at δ 2.03 (d, 3H), 9.47 (q, 1H) and δ 1.12 (d, 3H), 5.05 (q, 1H) correspond to acetaldehyde and its hydrate, respectively. The large peak at 4.9 ppm correspond to water in the reaction. The peak at δ 2.1 (s, 1H) comes from acetone contamination.

2.2. *cis*-CaaD Mechanism of Action

We looked to understand the mechanism of *cis*-CaaD in an attempt to better understand the kinetics of the product profile of the enzyme and improve activity while preventing decarboxylation. To determine the mechanism of *cis*-CaaD, we look to the structures of apo and substrate-soaked *cis*-CaaD. Both ACA and CCA substrates were used due to the possibility of native substrate CCA binding stronger to the active site and having superior trapping of intermediates than with ACA substrate. Proteins were concentrated to 18 mg mL⁻¹ in 10 mM Tris-SO₄ pH 8.0 and mounted as a hanging drop in 32 or 96-well plates with various crystallization solutions. Crystals were soaked for various lengths of time in 2 μ L of substrate solutions, which contained 1 μ L of 2 mM ACA or CCA and 1 μ L of the appropriate crystallization solution. Crystals were protected in Paratone or glycerol with crystallization solution and subsequently flash-frozen in liquid nitrogen.

Fortunately, intermediate trapping was successful when soaking wild-type and variant *cis*-CaaD crystals in both ACA and CCA substrate. Due to the homotrimeric nature of *cis*-CaaD, multiple chains were generally found in the asymmetric unit, which allowed trapping of multiple ligands in the active site of each crystal evaluated. In apo structures for the *cis*-CaaD variants E114D and E114N, we see sulfate buffer ions at a resolution of 1.6 Å and 2.5 Å respectively, and phosphate in E114Q at 1.8 Å resolution (Figures 2.6). Apo *cis*-CaaD (Y103F) at 2.2 Å resolution and apo *cis*-CaaD (T34A) at 2.5 Å showed a trapped sulfate in the active site similar to *cis*-CaaD E114D and E114N mutants (Figure 2.7).

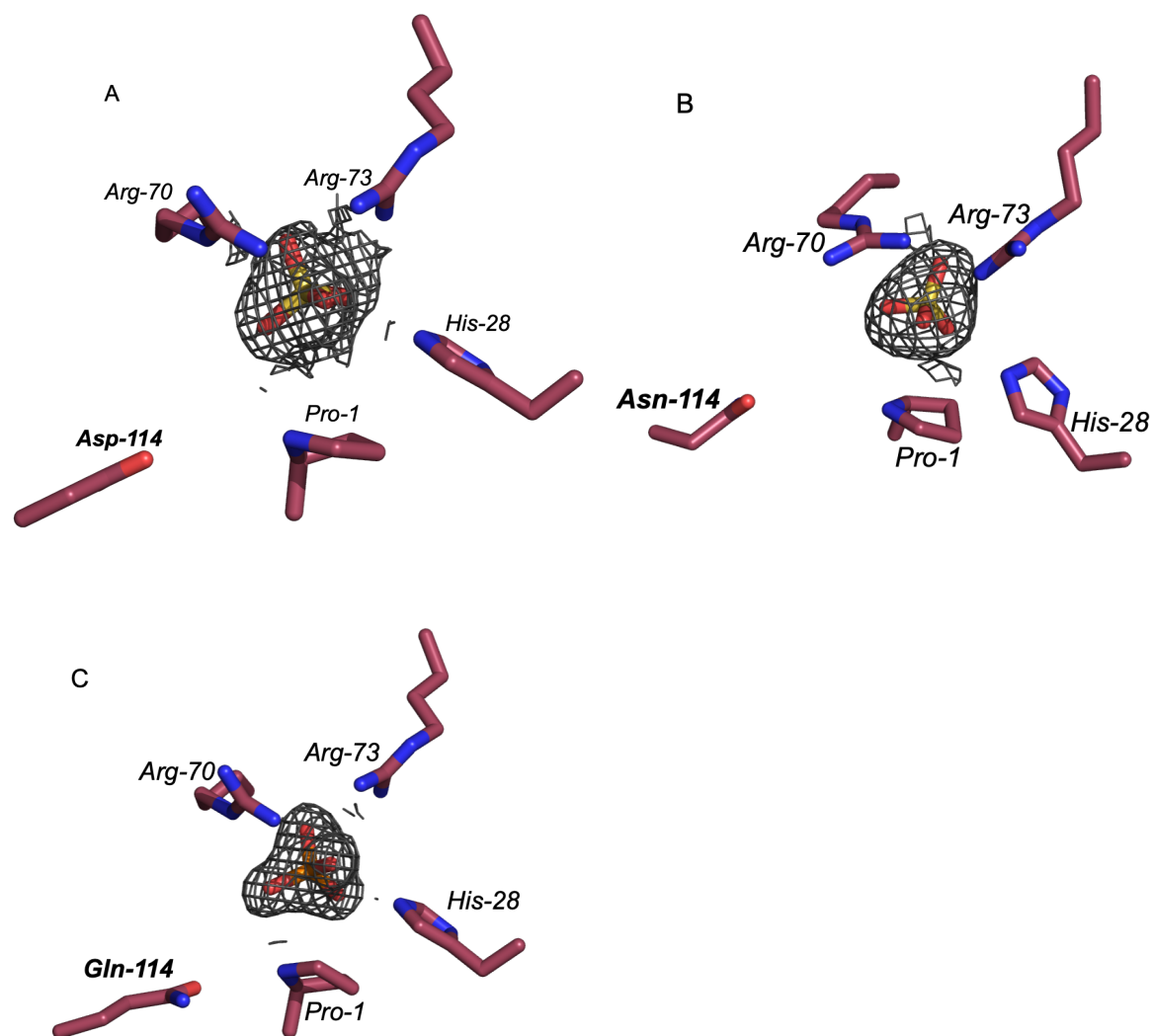


Figure 2.6. Crystal structure of (A) apo *cis*-CaaD (E114D) and (B) apo *cis*-CaaD (E114N) with sulfate in the active site and (C) apo *cis*-CaaD (E114Q) with phosphate in the active site.

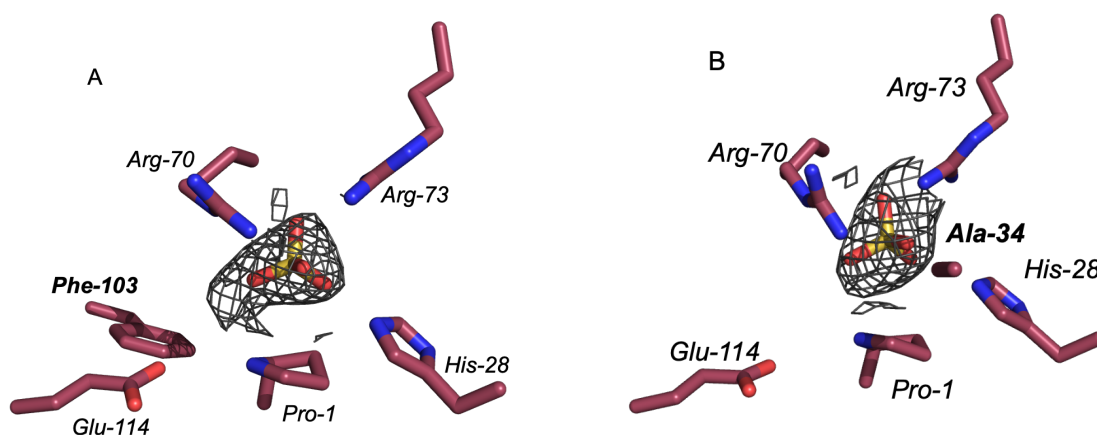


Figure 2.7. Crystal structure of apo *cis*-CaaD (Y103F) with sulfate in the active site (A). Crystal structure of apo *cis*-CaaD (T34A) with sulfate in the active site (B).

Evaluation of electron density of ACA-soaked structures of *cis*-CaaD variants showed unreacted ACA, MSA product, as well as trapped intermediates in the active site that gave insight into the mechanism of *cis*-CaaD. At 2.5 Å resolution, ACA-soaked *cis*-CaaD (E114D) shows the unreacted substrate, ACA with the catalytic water molecule and a covalent intermediate of 3-(*N*-prolyl)-acrylate bonded to Pro-1 (Figure 2.8). ACA substrate along with the catalytic water molecule used in the formation of MSA product can be seen in the active site of *cis*-CaaD (E114N) at 2.3 Å (Figure 2.9), however no covalent intermediate was trapped. Crystal structures of ACA-soaked *cis*-CaaD (E114Q) at 1.9 Å resolution show the covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate (Figure 2.10).

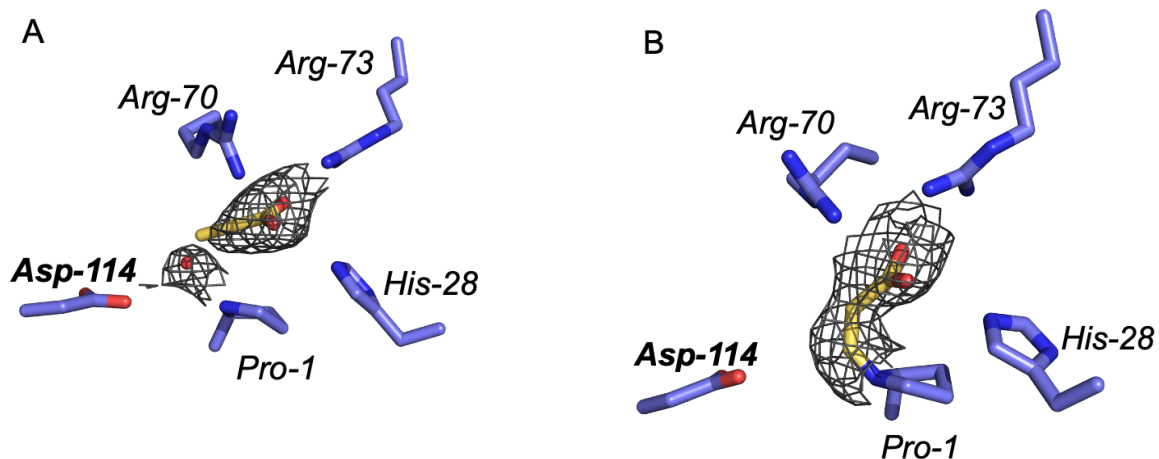


Figure 2.8. Crystal structures of ACA soaked *cis*-CaaD (E114D) at 2.5 Å showing (A) unreacted substrate and (B) covalent intermediate of 3-(N-prolyl)-acrylate.

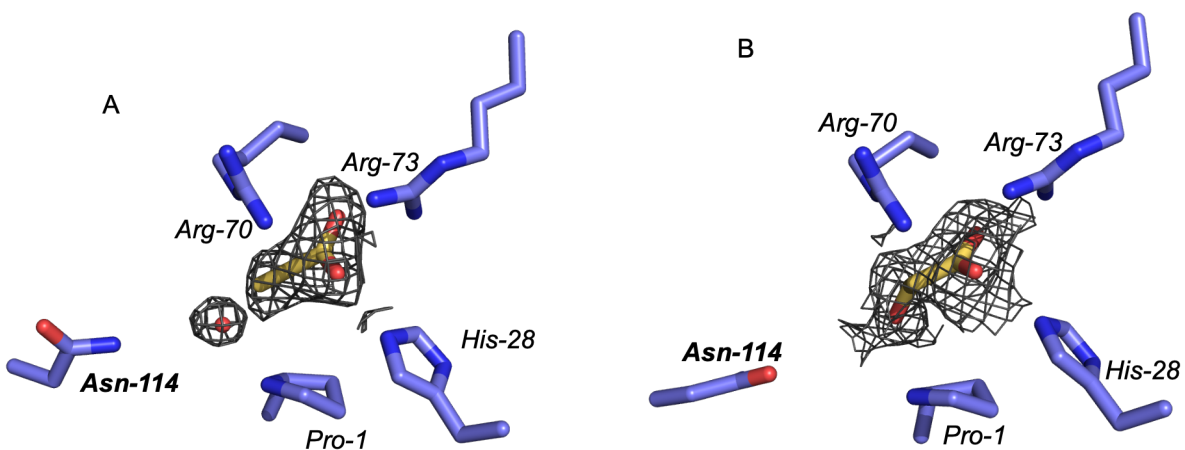


Figure 2.9. Crystal structure of shifted orientation of (A) ACA and (B) MSA product in the active site of *cis*-CaaD (E114N) is shown at 2.3 Å.

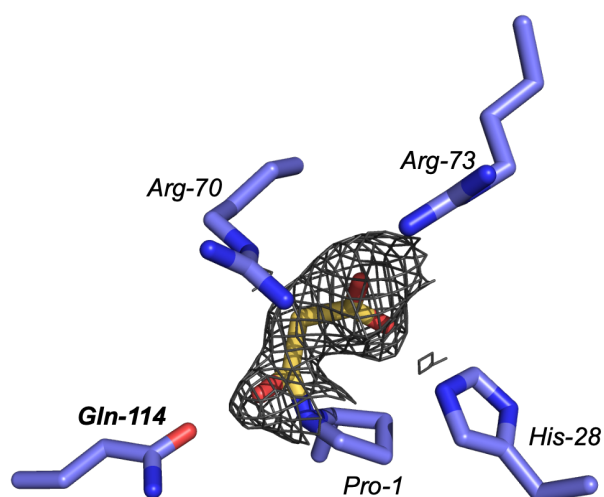


Figure 2.10. Crystal structures of ACA soaked *cis*-CaaD (E114Q) at 1.9 Å crystals showing covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1(A).

Although *cis*-CaaD (E114N) and *cis*-CaaD (E114Q) produce more ACH than MSA (65:35 and 88:12 respectively), we were able to trap the larger MSA molecule resulting from the hydration reaction without decarboxylation but did not trap the product of hydration and decarboxylation, ACH. This is likely because ACH is small and readily leaves the active site, while the carboxyl of MSA forms a salt bridge with Arg-70 and Arg-73 and is more stable in the active site. Acetate ion and water are seen at 2.3 Å in *cis*-CaaD (H28A) (Figure 2.11). Similarly, no substrate or product was trapped in the T34A or Y103F variants. Instead, a sulfate ion from the buffer solution was trapped in the Y103F crystal resolved at 2.3 Å and T34A crystal resolved at 2.5 Å (Figure 2.12). This is likely due to the higher activity of *cis*-CaaD (T34A) and (Y103F) than the other variants evaluated. Lower activity variants allowed for slower conversion of substrate to product and intermediate trapping was able to be obtained during crystal soaks.

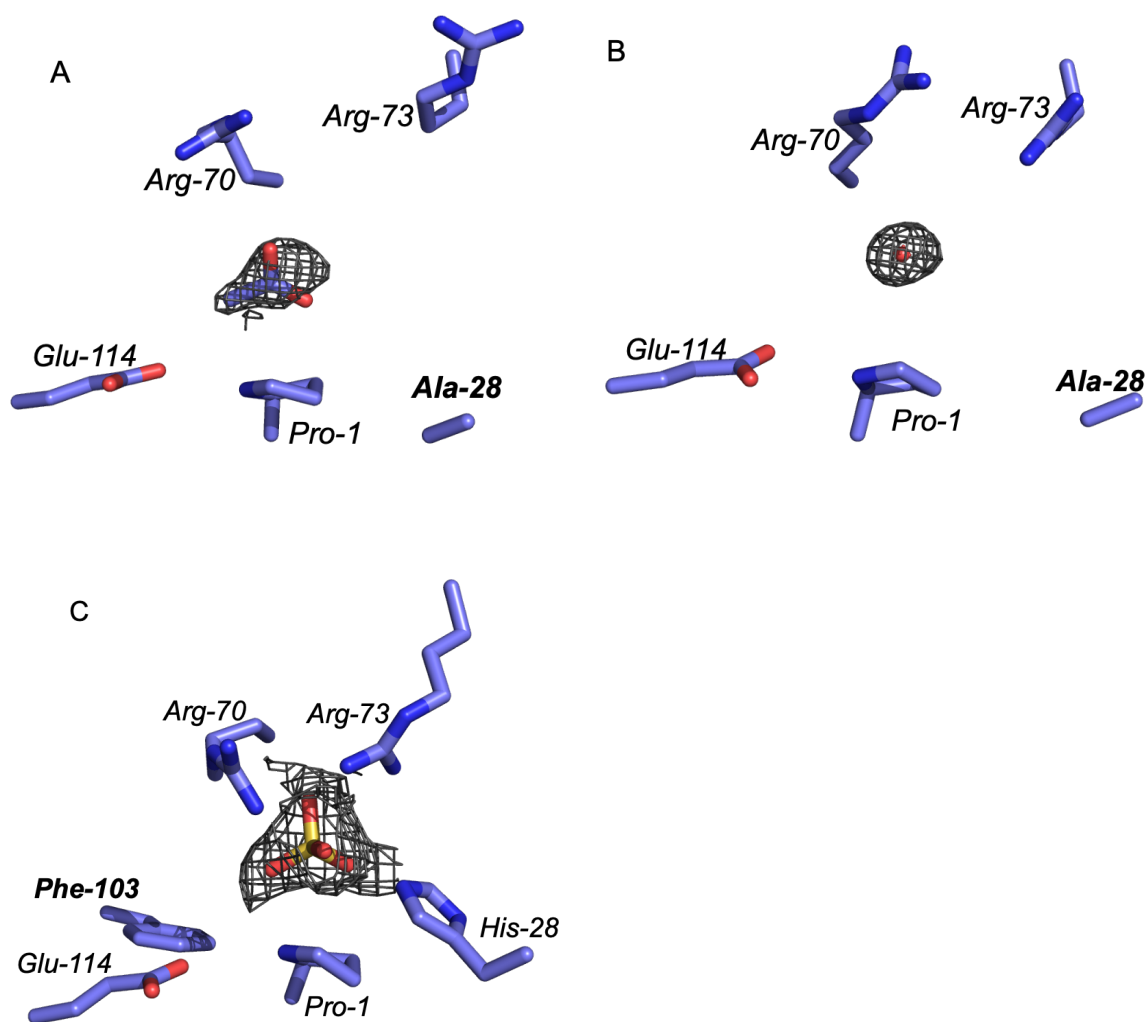


Figure 2.11. Crystal structures of ACA soaked *cis*-CaaD (H28A) with (A) acetate and (B) water in the active site. Crystal structures of ACA soaked *cis*-CaaD (Y103F) with (C) sulfate in the active site.

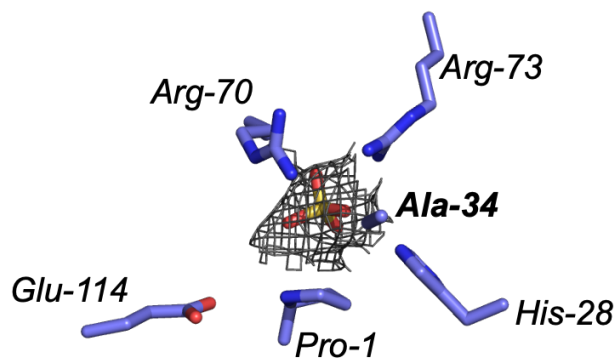


Figure 2.12. Crystal structures of ACA soaked *cis*-CaaD (T34A) with sulfate in the active site.

When using CCA as a substrate for crystal soaking, similar results are seen. When soaked with ACA, wild-type *cis*-CaaD presents unreacted ACA substrate at 2.0 Å resolution but when soaked with CCA, buffer ion acetate is seen in the active site at 1.3 Å (Figure 2.13).

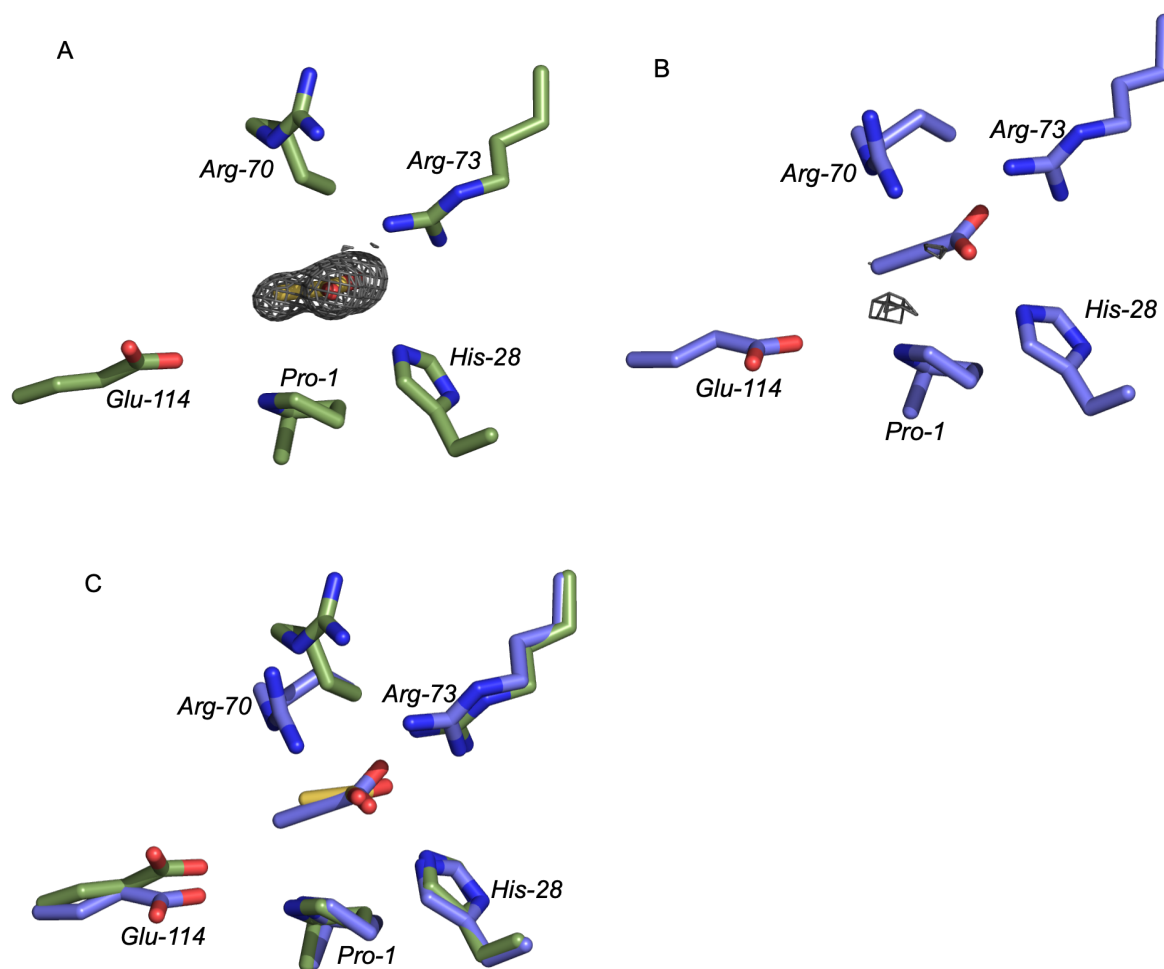


Figure 2.13. Crystal structure of wild-type *cis*-CaaD at 1.3 Å (A) with an acetate in the active site, (B) soaked with ACA substrate showing an unreacted ACA in the active site and 2.0 Å, and (C) the two aligned.

CCA substrate soaked *cis*-CaaD (E114D) show a covalently linked (*N*-prolyl)-ethene intermediate as well as sulfate at 1.8 Å resolution (Figure 2.14). The E114Q variant of *cis*-CaaD show covalently linked intermediate of 3-(*N*-prolyl)-3-hydroxypropionate to Pro-1, the covalent intermediate of 3-(*N*-prolyl)-acrylate linked to Pro-1, as well as product MSA at 2.2 Å resolution (Figure 2.15). For *cis*-CaaD (Y103F) we were able to trap three waters at 2.2 Å (Figure 2.16). Again, for *cis*-CaaD (T34A) we see screen buffer sulfate ion trapped in the active site at 2.4 Å,

likely due to the higher activity than other variants shown (Figure 2.17C). In the crystals obtained we did not see clear definition of an unreacted trapped CCA substrate.

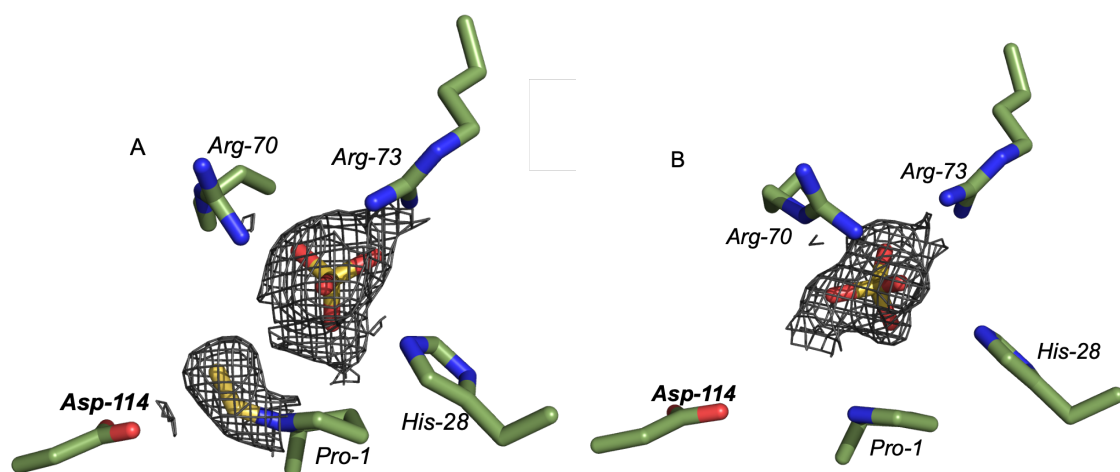


Figure 2.14. Crystal structures of CCA soaked *cis*-CaaD (E114D) at 1.8 Å showing (A) the covalent intermediate of (N-prolyl)-ethene linked to catalytic Pro-1 with sulfate and (B) sulfate alone.

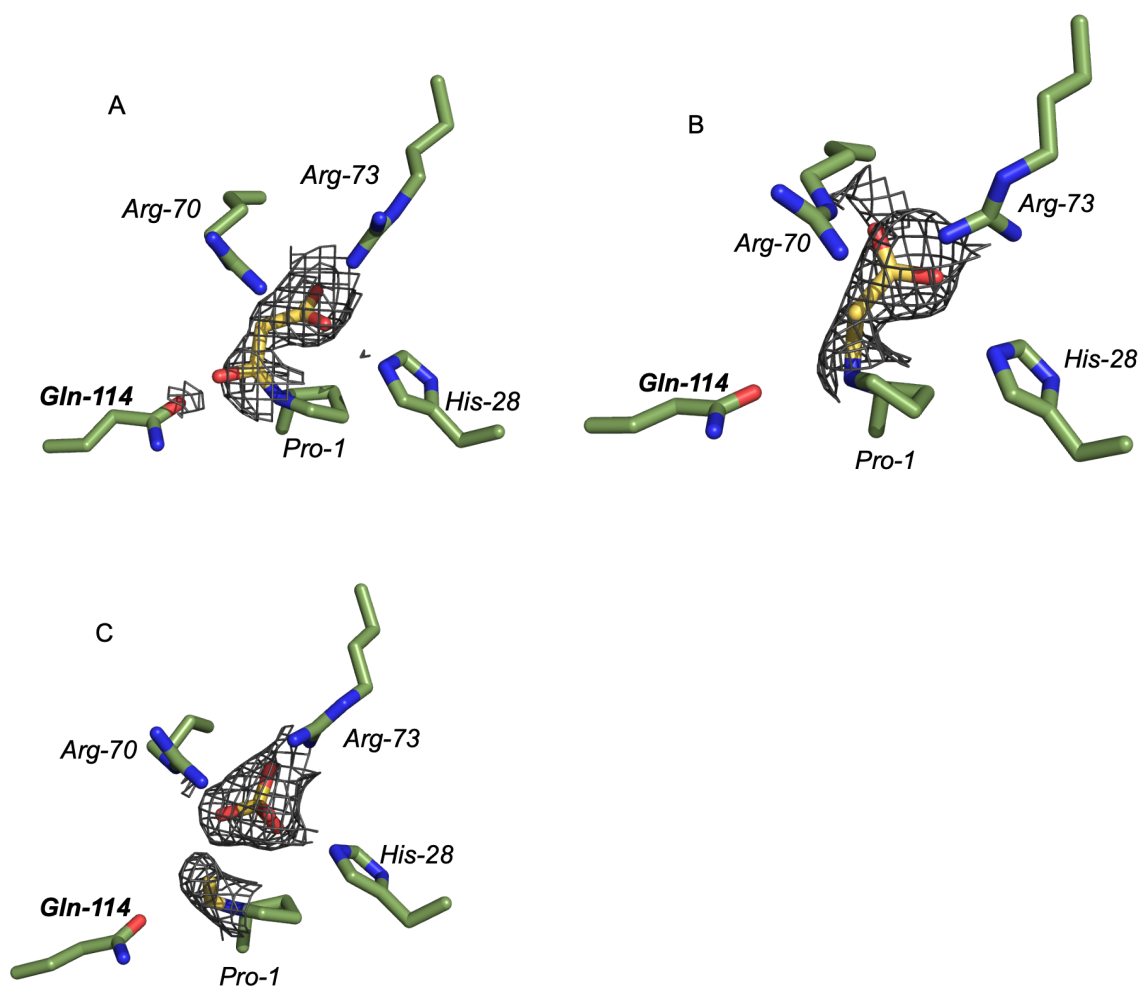


Figure 2.15 Crystal structures of CCA soaked *cis*-CaaD (E114Q) at 2.2 Å crystals showing (A) covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1, (B) covalent intermediate of 3-(*N*-prolyl)-acrylate linked to catalytic Pro-1 and (C) (*N*-prolyl)-ethene covalent intermediate with sulfate.

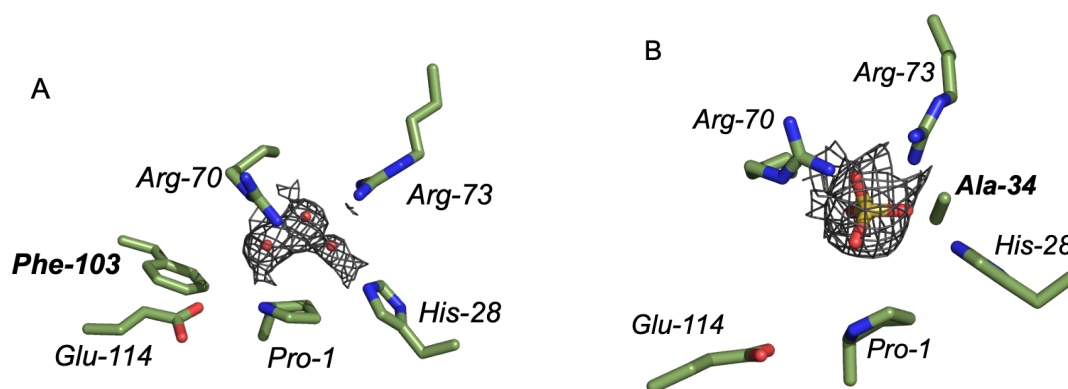


Figure 2.16 Crystal structures of CCA soaked *cis*-CaaD (Y103F) with (A) three waters in the active site. Crystal structures of CCA soaked *cis*-CaaD (T34A) with (B) sulfate in the active site.

In all *cis*-CaaD structures that show intermediates trapped, substrate, or product trapped in the active site, there is flexibility in Arg-70. This seems to play a role in the orientation of the substrate as it is converted to product. The product MSA seems to be released during the reaction and stabilized through hydrogen bonding by active site arginine residues Arg-70 and Arg-73.

Hypothesized mechanisms have previously been reported for *cis*-CaaD as well as Cg10062.^{1,2,5,6,7,8,9} However, by obtaining crystal structures of trapped intermediates of *cis*-CaaD when soaked with substrate we are able to provide strong supporting evidence for a similar mechanism for both (Figure 2.17).¹⁰ The reaction of substrate with *cis*-CaaD enzyme is initiated by a covalent interaction. Based on information gathered from structural data of substrate-soaked *cis*-CaaD variants, we hypothesize that *cis*-CaaD can proceed through a hydration mechanism (purple arrows) and a minor hydration/decarboxylation mechanism (orange arrows) to produce MSA and ACH. Both ACA (**1a**) and CCA (**1b**) situate in the active site poised for nucleophilic attack at C-3, resulting in formation of 3-(*N*-prolyl)-acrylate (ACR, **2**). Tautomerization of the enamine intermediate (**2**) affords a hypothesized iminium species (**3**). The electron withdrawing nature of **3** can result in decarboxylation (orange arrow) with formation of (*N*-prolyl)-ethene (ETH, **4**), which following hydration and bond breakage affords ACH. Alternatively, either intermediate **2** or **3** may

undergo hydration at C-1 to form the 3-(*N*-prolyl)-3-hydroxypropionate intermediate (HPA, **6**, purple arrow). Collapse of the covalent bond between Pro-1 nitrogen and the intermediate provides MSA (**7**). Given the propensity for *cis*-CaaD to form MSA over ACH, we hypothesize that a catalytic water molecule in the active site is well-situated for addition, minimizing subsequent decarboxylation. Therefore, the electron withdrawing nature of the iminium is required for decarboxylation to occur. This is similar to the formation of an iminium withdrawing group that is the key determinant in the enormous family of PLP-dependent enzymes, representing one of the largest family of enzyme decarboxylases. In this mechanism the amino group of a lysine residue and the aldehyde of PLP forms a Schiff-base structure, referred to as an aldimine, which then breaks to form a new Schiff base structure between the amino group of the substrate and the aldehyde of PLP through transaldimination.¹¹ Our mechanism of *cis*-CaaD can be seen as a derivative of this mechanism, where the Pro-1 is a secondary amine which does not require Schiff base protonation to provide a powerful electron sink. The *cis*-CaaD mechanism proceeds correspondingly with CCA as a substrate as ACA, except that the chloride ion of CCA is removed to form HCl by product with a neighboring water molecule. This mechanism is indistinguishable to that of Cg10062. Yet, the wild-type of *cis*-CaaD favors formation of MSA (>99%) over ACH (<1%) and wild-type Cg10062 favors formation of ACH (85%) over MSA (15%). Thus, *cis*-CaaD favors hydration only while Cg10062 favors hydration with decarboxylation. We therefore set out to understand the subtle differences between the two enzymes which afford such a dramatic disparity in product profile.

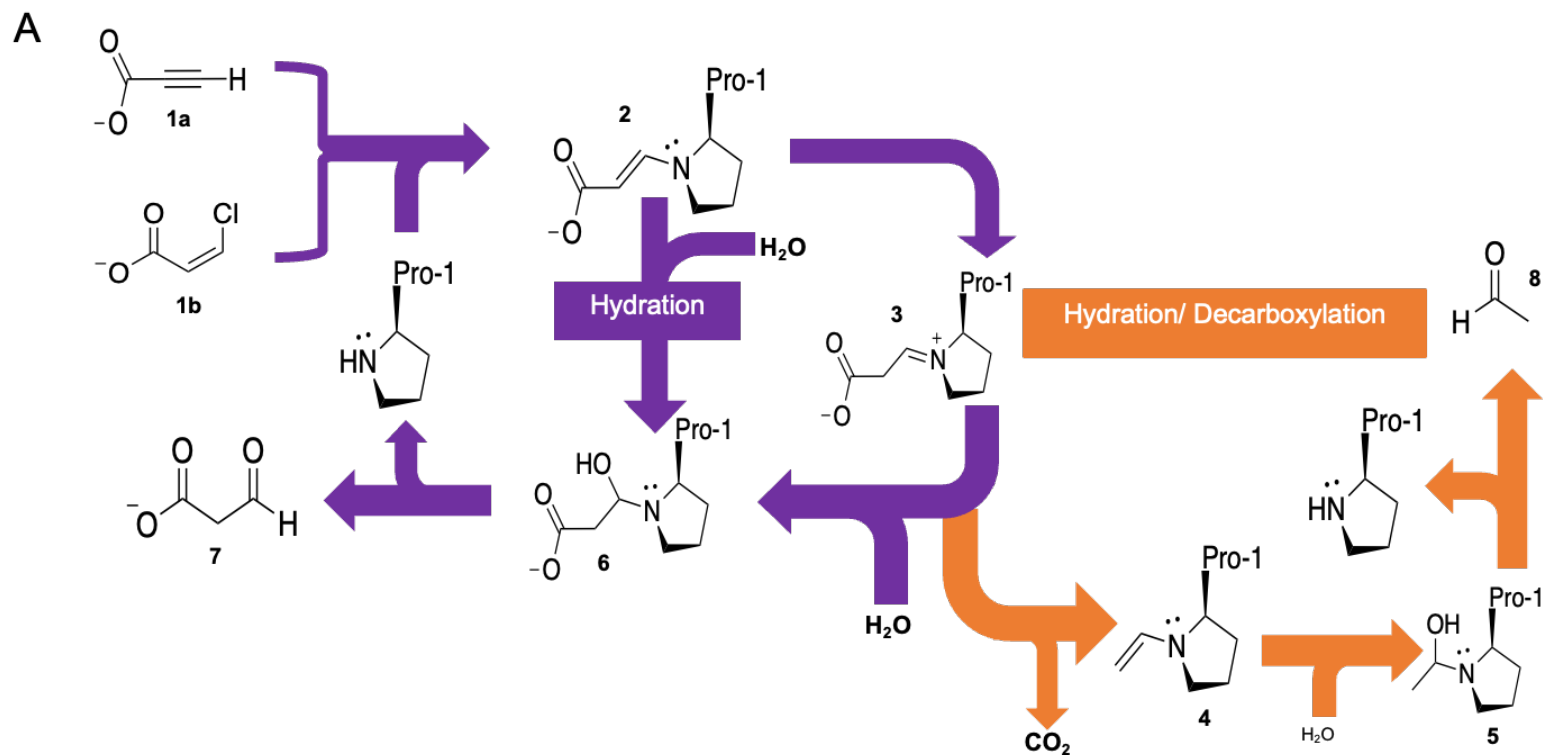


Figure 2.17 Mechanism of *cis*-CaaD (A) without and (B) with crystal structures showing intermediates trapped in the active site.

B

2.3 Cg10062 (E114N) Engineering

We next sought to discover an enzyme variant with improved kinetics for ACA transformation that selectively catalyzes hydration only. Although *cis*-CaaD and Cg10062 can both catalyze transformation of ACA, their native forms produce different product profiles of MSA and ACH. Wild-type Cg10062 leads to rapid decarboxylation to ACH while *cis*-CaaD forms exclusively MSA, although Cg10062 is more catalytically efficient. Here we look to variations of previously characterized Cg10062 variants to find an enzyme that converts ACA to exclusively MSA with high activity. The kinetic parameters of Cg10062 and its variants have been extensively studied.^{3,4,7} Additionally, we uncovered a novel variant Cg10062 (E114N) which converts ACA to exclusively MSA at high efficiency. Structural biochemistry was used to evaluate the structure and function of this variant. Kinetic constants of variants of Cg10062 were previously characterized.⁴ Determination of the more effective enzyme for catalytic conversion of ACA substrate to exclusively MSA was performed for Cg10062 as well as *cis*-CaaD. Previous reports described hydratase activity with variants Cg10062 (E114D) and Cg10062 (E114Q).³ These variants showed major hydration to MSA but with kinetics similar to that of wild-type Cg10062. Active site variants E114N, Y103F, Y103A, E114D-Y103F, H28A, R70A, R70K, R73K, E114A, and E114S were rationally designed using Q5 site-directed mutagenesis. These mutations were influenced by the active site residues' function within the active site. Arg-70, Arg-73, and His-28 mutations were performed to determine the effect on altering the presumed stability that these residues provide. Glu-114 and Tyr-103 modifications were performed to address their role in the activation of a catalytic water molecule during hydration of substrate. Cg10062 from *Corynebacterium glutamicum* was cloned similarly to *cis*-CaaD into the commercial pET-21a(+) vector containing a TEV protease recognition site and His₆-tag on the C-terminus.

2.3.1 Cg10062 Variant Product Profile and Kinetic Activity

An identical assay was used to characterize Cg10062 as *cis*-CaaD. The kinetic parameters of variants of Cg10062 have previously been characterized by others in Draths lab.^{2,4}

Just as with variants of *cis*-CaaD, MSAD was used to identify products formed with Cg10062 variants. The difference of the two rates in the absence and presence of MSAD measures the hydratase or hydratase with decarboxylase activity of each variant. ACA was the sole substrate used for the evaluation of Cg10062 variant activities. Wild-type Cg10062 is highly catalytic for substrate ACA, but MSA is the minor product (15%) with decarboxylation to ACH as the major product (85%). Alanine mutants of His-28, Arg-70, Arg-73, Tyr-103, and Glu-114 previously showed appreciable loss in activity with R70A and R73A reducing activity to the point of unattainable product profiles. Non-alanine Glu-114 variants designed to affect hydrogen bonding of w Table 2.2. (cont'd) led to a change in the ratio of products observed. Cg10062 E114D, E114N, and E114Q all yield exclusive MSA production (>99%, Table 2.2). Cg10062 E114A, Y103A and Y103F maintained decarboxylase activity, while H28A showed majority MSA production albeit not exclusively.

Table 2.2. Product profile of variants using ACA substrate. ^{4,6}

Enzyme Variant	Product Ratio (%)	
	MSA	ACH
Cg10062 (wild-type)	15 ± 0.6	85 ± .04
Cg10062 (E114N)	>99	<1
Cg10062 (E114D)	>99	<1
Cg10062 (E114Q)	>99	<1
Cg10062 (Y103F)	36	64
<i>cis</i> -CaaD (wild-type)	>99 ± 1.0	<1 ± 0.0
<i>cis</i> -CaaD (E114D)	>99 ± 1.0	<1 ± 0.0
<i>cis</i> -CaaD (E114N)	35 ± 0.5	65 ± 0.5
<i>cis</i> -CaaD (E114Q)	12 ± 1.1	88 ± 2.0
<i>cis</i> -CaaD (Y103F)	>99 ± 0.4	<1 ± 0.6
<i>cis</i> -CaaD (Y103A)	>99 ± 1.0	<1 ± 0.0
<i>cis</i> -CaaD (H28A)	-	-
<i>cis</i> -CaaD (R73A)	-	-
<i>cis</i> -CaaD (T32A)	>99 ± 0.5	<1 ± 0.5
<i>cis</i> -CaaD (T34A)	>99 ± 0.9	<1 ± 0.1
<i>N-cis</i> -CaaD/C-Cg10062	98 ± 0.6	2 ± 0.4
<i>N</i> -Cg10062/ <i>C-cis</i> -CaaD	96 ± 0.9	4 ± 0.1
<i>tcis</i> -CaaD	60 ± 0.2	40 ± 0.8

The mechanism of Cg10062 has previously been described.⁴ Investigation of steady-state kinetics of the hydratase-only variants will allow the careful selection of a highly catalytic sole MSA

producing variant. Although Cg10062 (E114D), (E114N), and (E114Q) variants of Cg10062 lead to solely hydratase activity and MSA production, Cg10062 (E114D) and (E114Q) have lower catalytic activity than that of wild-type (Table 2.3). Cg10062 (E114N) exhibited higher activity than the other hydratase only variants studied. Additionally, Cg10062 (E114N) has improved binding affinity (45 μ M) than wild-type *cis*-CaaD (1050 μ M), making it currently the most catalytic efficient enzyme with exclusive MSA. production. Cg10062 (E114N) may be utilized in subsequent studies for the production of building block chemicals such as 3HP. Additional studies by others demonstrate production of MSA by Cg10062 (E114N) *in vitro* and further formation of 3HP by NADPH dependent enzyme YdfG.⁶ YdfG is a 3-hydroxy dehydrogenase from *E. coli* that is able to convert MSA to 3HP.^{12,13}

Table 2.3. Michaelis-Menten kinetics for Cg10062 and variants.^{4,6}

Enzyme Variant	ACA		
	K_m (mM)	k_{cat} (s ⁻¹)	$k_{cat}/K_m \times 10^4$ (M ⁻¹ s ⁻¹)
Cg10062 (wild-type)	66 \pm 17	4.2 \pm 0.28	6.3 \pm 1.2
Cg10062 (E114N)	45 \pm 1	1.8 \pm 0.07	4.0 \pm 0.4
Cg10062 (E114D)	557 \pm 62	1.12 \pm 0.05	0.20 \pm 0.02
Cg10062 (E114Q)	64 \pm 12	0.66 \pm 0.04	1.08 \pm 0.21
Cg10062 (Y103F)	122 \pm 17	0.43 \pm 0.01	0.36 \pm 0.05
<i>cis</i> -CaaD (wild-type)	1050 \pm 109	1.40 \pm 0.05	0.13 \pm 0.01
<i>cis</i> -CaaD (T32A)	-	-	-
<i>cis</i> -CaaD (T34A)	1730 \pm 199	0.59 \pm 0.03	0.03 \pm 0.00
<i>N</i> -Cg10062/ <i>C-cis</i> -CaaD	1330 \pm 363	1.02 \pm 0.11	0.08 \pm 0.01

2.4 Cg10062 Mechanism

Crystal structures were obtained of variants of Cg10062. From the structural and kinetic analysis of Cg10062 and its variants, a detailed mechanism of the enzyme's decarboxylation and hydration pathway has been published (Figure 2.18).² The first step in the hydration and

hydration/decarboxylation of ACA is nucleophilic attack on the substrate, forming a covalent enamine 3-(*N*-prolyl)-acrylate intermediate. This can either go through direct hydration, as seen with Cg10062 (E114D), or the enamine may tautomerize to the iminium form. This highly reactive iminium can then either undergo hydration or decarboxylation. The strong electron withdrawing nature of iminium can lead to decarboxylation and ACH production. Alternately, water addition to the iminium may lead to hydration and the intermediate, 3-(*N*-prolyl)-3-hydroxypropionate. This intermediate has been trapped in many variants including Cg10062 (R73A) and Cg10062 (H28A). Finally, the covalent bond between the nitrogen of Pro-1 and the intermediate is lost and MSA is formed.

2.4.1. Cg10062 (E114N) Crystal Structure

Incubation of Cg10062 and variants' protein crystals with the substrate ACA afforded structures of Cg10062 with intermediates covalently bonded Pro-1, indicating covalent catalysis (Figure 2.18). In our previous study, the active site of the hydratase-only E114D variant contained a tetra-coordinated water molecule that was presumed to prevent decarboxylation.⁴ Relative to wild type Cg10062, the shorter side chain of Asp-114 creates space for hydrogen bond formation between the water molecule and the nitrogen of Pro-1, the hydroxyl group of Tyr-103, the carboxylate group of Asp-114 and the main chain carbonyl of Leu-38. It was presumed that the water interacts with the Pro-1 nitrogen, preventing formation of an iminium intermediate required for decarboxylation activity.

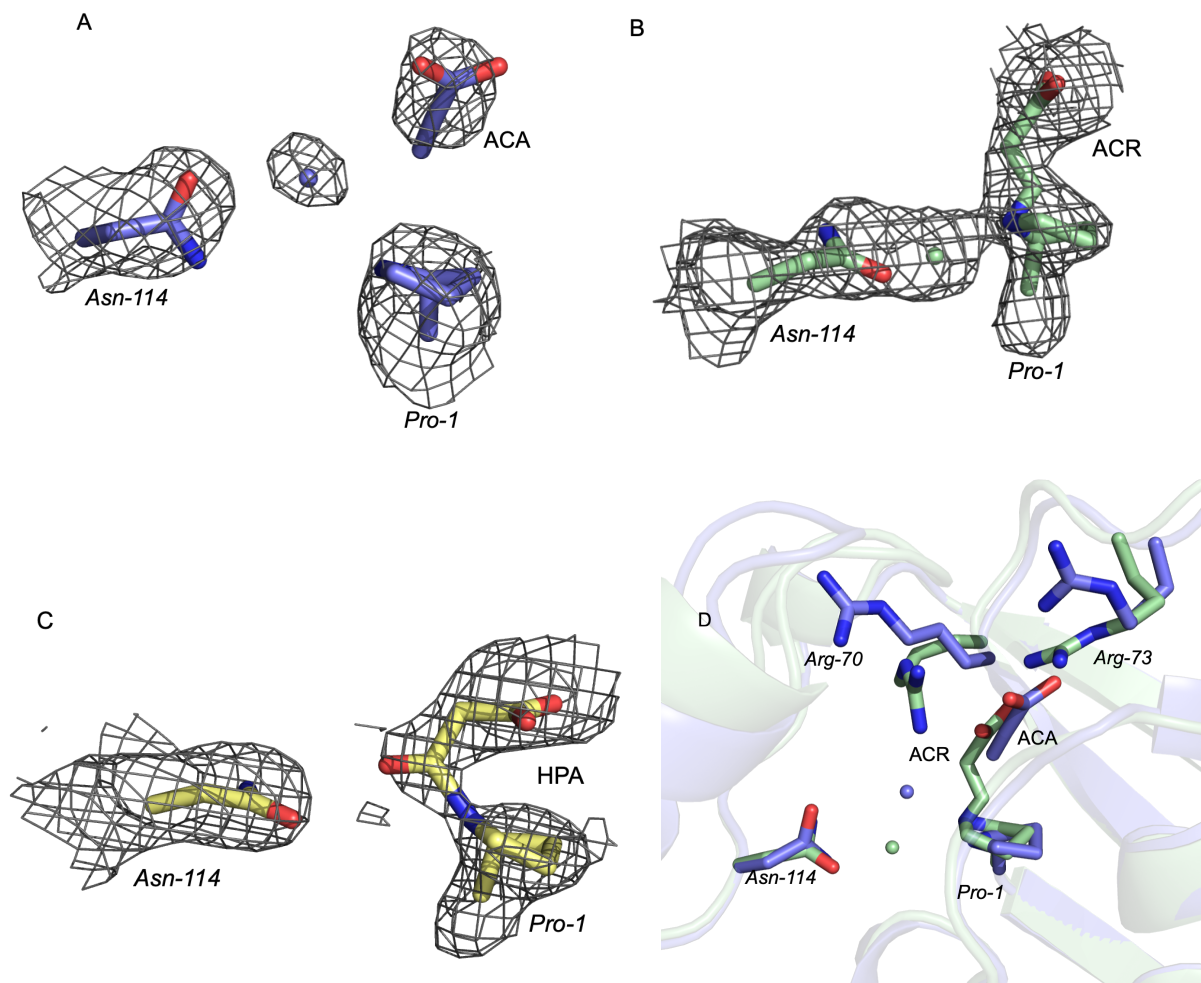
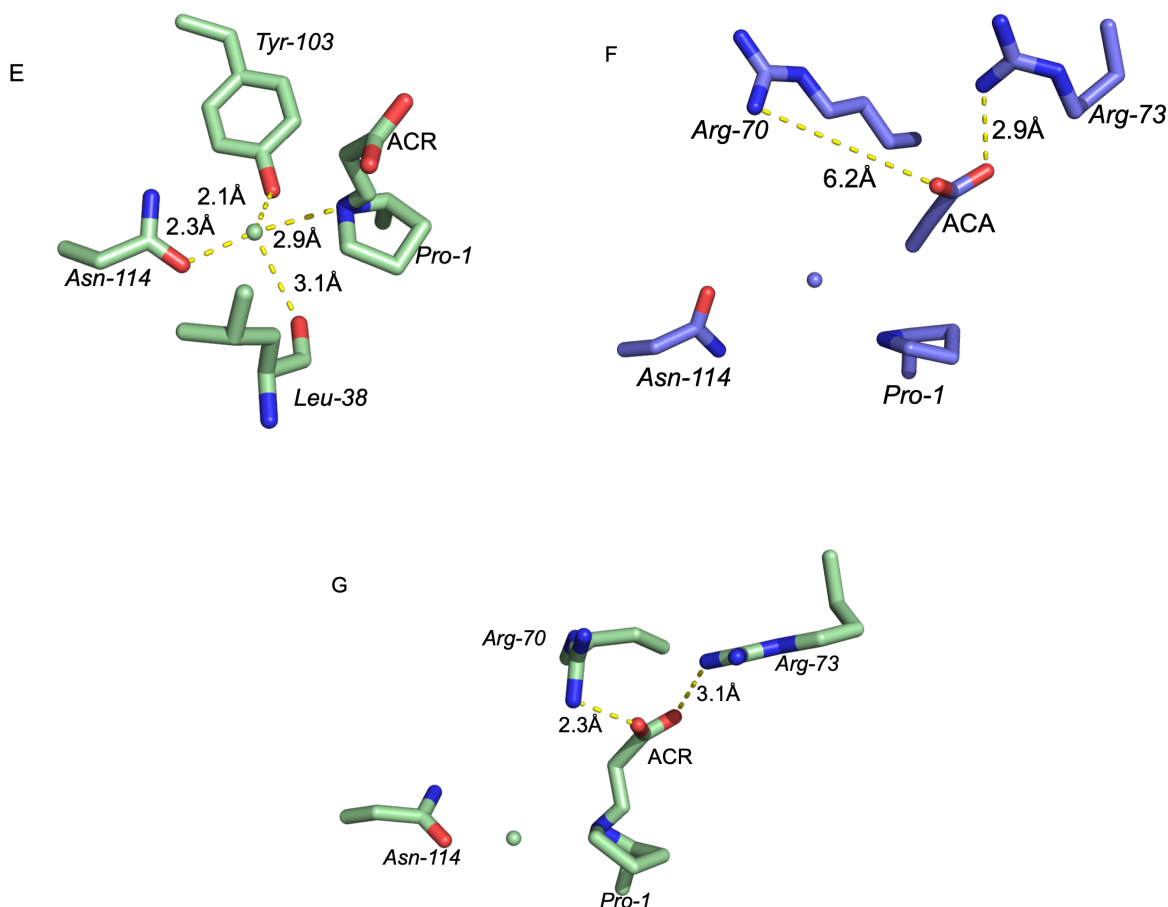


Figure 2.19. Cg10062 (E114N) is shown with (A) trapped ACA in the active site, (B) 3-(*N*-prolyl)-acrylate (ACR) intermediate, (C) 3-(*N*-prolyl)-3-hydroxypropionate (HPA) intermediate, and (D) 3-(*N*-prolyl)-acrylate and ACA aligned. The acrylate (E) is in hydrogen bonding distance to the side chain of the Tyr-103 (2.1 Å), Pro-1 (2.9 Å), and Asn-114 (2.3 Å) residues and Leu-38 carbonyl main chain (3.1 Å). Arg-70 and Arg-73 are (F) farther away in the unreacted substrate ACA and (G) move closer within hydrogen bonding distance to the reacted intermediate 3-(*N*-prolyl)-3-hydroxypropionate forming a salt bridge.

Figure 2.19. (cont'd)



The absence of the water molecule in wild-type Cg10062 led to the conclusion that the shorter side chain of Asp-114 creates space in the active site for the water-mediated hydrogen bonding network, leading to the design of E114N variant with a side chain of similar length. The crystal structure of ACA-soaked E114N showed a similar hydrogen bond network between an ordered water molecule and the residues Pro-1, Leu-38, Asn-114 and Tyr-103 residues, supporting the hypothesis that the water molecule plays a crucial role in preventing decarboxylation (Figure 2.19E). Electron withdrawing amino acids Arg-70 and Arg-73 lead to salt bridge formation with the carbonyl of the reacted intermediate leading to the protonation of the amine group of Pro-1 allowing the hydration to occur (Figure 2.20G). Biochemical characterization of the novel E114N variant confirmed hydratase-only activity higher than the previously identified E114Q and E114D variants (Table 2.1).

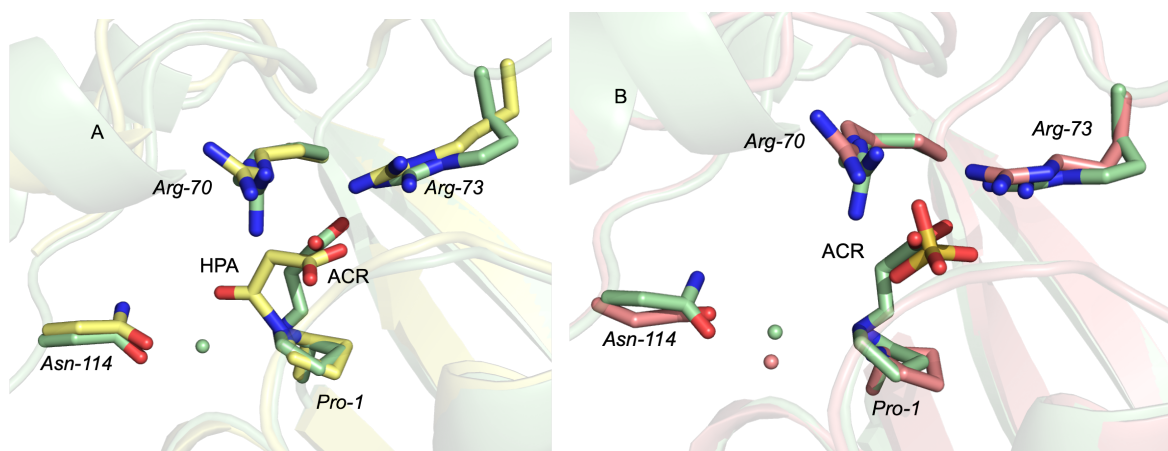


Figure 2.20. 3-(*N*-prolyl)-acrylate (ACR) intermediate aligned with (A) 3-(*N*-prolyl)-3-hydroxypropionate (HPA) intermediate or (B) sulfate within the active site, respectively. The ordered water resembles the stage prior to water's nucleophilic attack.

Cg10062 (E114N) producing only MSA would suggest that this variant follows the hydration pathway that is proposed in the mechanism of Cg10062 previously supported.⁴ Cg10062 (E114N) appears to follow the mechanism previously described, where MSA formation proceeds via covalent catalysis. Nucleophilic attack on C3 of ACA by the nitrogen of Pro-1 yields the tautomerization of enamine to iminium 3-(*N*-prolyl)-acrylate (Figure 2.20). Alignment of a water in the active site with the electron withdrawing iminium leads to subsequent water addition, forming a covalent 3-(*N*-prolyl)-3-hydroxypropionate intermediate (Figure 2.19C), prior to release of MSA from the active site.

2.4.2. Cg10062 Error Prone Mutagenesis for Improved Activity

Error-prone mutagenesis was performed to screen for a more active Cg10062 variant (Figure 2.21). Error-prone mutagenesis introduces random mutations to create a library of genetically diverse variants.¹⁴ This is a useful tool in protein engineering as it may enable optimization of enzyme activity and stability through screening assays. Mutants are selected based on desired characteristics which accelerates discovery of optimal variants.¹⁵ By utilizing error-prone mutagenesis and a coupled kinetic reaction we may screen multiple variants at once for a highly active tautomerase for the conversion of ACA to MSA.

Error-prone mutagenesis was carried out using wild-type Cg10062 as well as Cg10062 (E114N) templates using a GeneMorph II Random Mutagenesis kit according to manufacturer's instruction.¹⁶ PCR products were ligated using T4 ligase, transformed into BL21(DE3) cells, plated on LB with ampicillin, and grown overnight at 37 °C. From the viable colonies, 96 single colonies were selected and inoculated into a deep well 96 well plate, each well containing 1 mL of LB with antibiotic. Cultures were allowed to grow at 37 °C overnight with shaking. Cultures were then reinoculated to an OD₆₀₀ of 0.05 and induced with IPTG to a final concentration of 1 mM after one and a half hours. Cells were then pelleted and resuspended in 500 µL of cell lysis buffer containing 0.25 mg mL⁻¹ lysozyme and incubated for 10 minutes at 37 °C shaking. The plate was then centrifuged and 10 µL of supernatant was added to fresh 1 mL 96 well plates containing 10 µL NADH (10 mg/ml) and 0.5 mg mL⁻¹ MmsB. MmsB, a 3-hydroxyisobutyrate dehydrogenase known to convert MSA to 3HP, was obtained by the amplification of the *mmsB* gene *P. putida* KT2440 genomic DNA and cloning into the pET-21a(+) vector at the NdeI and XhoI sites. The kinetic reaction was initiated by the addition of 50 µL ACA (20 mM). Absorbance was taken at 340 nm following the oxidation of NADH. ACA that has been converted to MSA is converted to 3HP by the NADH dependent enzyme MmsB. Conversion of MSA to 3HP is monitored by NADH depletion. Increase in activity indicates high MSA production and a promising variant. Cg10062 (E114N) was also screened similarly without undergoing error-prone mutagenesis as a control to compare activities and screen for a more active mutant.

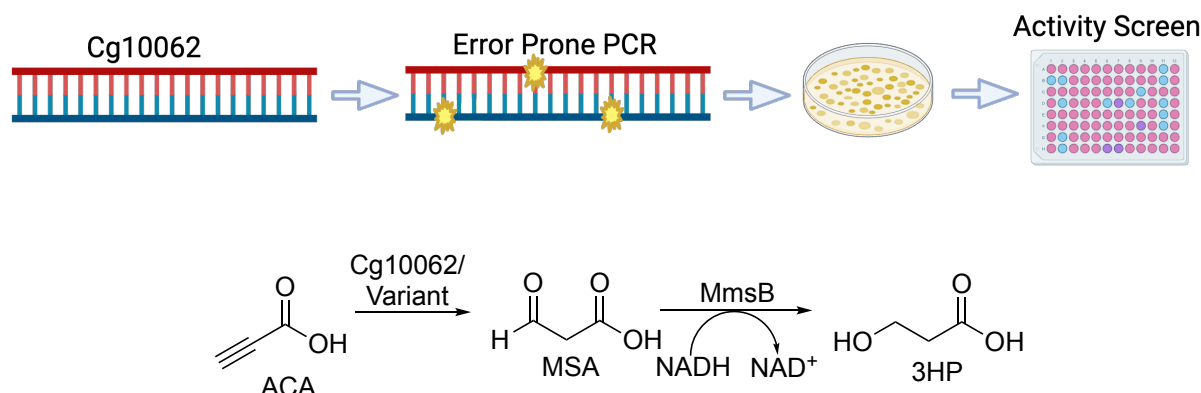


Figure 2.21. Error prone PCR and activity assay to screen for highly active variants.

Of the screened variants mutants were categorized by their activities. The top variants displayed activities too high to obtain a readable slope to determine reaction rates. These were diluted twofold, fourfold, and 24-fold and activity was remeasured. From this we obtained seven promising variants that were then sequenced. Three of these came from error-prone mutagenesis performed on Cg10062 (E114N) and four of these variants came from error-prone mutagenesis performed on wild-type Cg10062. From sequencing of the three Cg10062 (E114N) variants, one variant reverted to wild-type, one retained the E114N mutation with no other mutations, and one contained the mutations E114N, V46E, A58S, as well as two silent mutations. From sequencing the four wild-type Cg10062 variants, three contained E114N mutation and one contained a frameshift mutation that affected the affinity tag and purification. Due to the unknown effect further genetic manipulation would have on this variant to replace the affinity tag, this variant was removed as a potential target. The variants containing wild-type and E114N sequences were also removed. The repetition of sole E114N variants could be due to the use of Cg10062 (E114N) plasmid backbone vector during ligation. This left one of the seven original potential target variant E114N:V46E:A58S or “NL60”. A second round of mutagenesis and subsequent sequencing showed mutant “NL15b” (R11H:L31V:V102L:E114N:T141I:E142G) with promising activity. Large scale purification and kinetics using MSAD-coupled kinetics showed NL60 and NL15b activity

comparable for but not improved from Cg10062 (E114N). This process was repeated giving no variant to date with better enzyme kinetics than Cg10062 (E114N).

2.4.3 Cg10062 Biosensor for Improved Activity

Fast and efficient optimization of enzymes is of great importance. Directed evolution may be used to generate a library of enzyme variants. This library is then screened for improved activity and selectivity. Biosensors enable high-throughput screening of enzyme variants by providing real-time data of the assay being performed.¹⁷ It continuously monitors the product being formed and is designed to detect a reaction product or intermediate, for example by using fluorescence. This enables easy optimization of reaction conditions and quick readouts.

Previous studies have used biosensors to detect bioproduction of lactate and 3-hydroxypropionic acid by cells.^{18,19} Here, random mutagenesis can be used on Cg10062 (E114N) to screen for a higher activity variant. Literature has described methods of protecting certain amino acid residues during random mutagenesis, so that we may maintain the E114N mutation and screen for other additional variants that may improve activity.²⁰ Fluorescence is initiated through a two-plasmid component system. MSA converted from ACA is subsequently converted to 3HP by MmsB. A MmsR transcriptional regulator from *P. dentrificans* binds 3HP and form a 3HP-MmsR complex. This complex binds to the P_{mmsA} promoter to turn on transcription of the red fluorescence protein *rfp* gene. The increase in fluorescence would indicate higher activity and fluorescence activated cell sorting (FACs) could be used to screen for a more active variant for the conversion of ACA to MSA. We designed a biosensor system using Cg10062 (E114N) to optimize the conversion of ACA to MSA by Cg10062.²¹ However, limitations were placed due to the inability of ACA to enter the cell. Nevertheless, we were able to show the sensing of increased fluorescence by red fluorescence protein (RFP) based on increased IPTG concentration.

For the control of fluorescence readout, a single colony of MG1655 with a plasmid containing RFP (MG1655/pBbA1a-RFP) was grown overnight and reinoculated in M9/glu media to an initial OD₆₀₀ of 0.05. The cells were allowed to grow at 37 °C until the OD₆₀₀ reached 0.5–

0.7. Aliquots were transferred to a 96-well plate and IPTG was added to each well to a final concentration of 0, 0.01, 0.1 and 1 mM. OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 hours (Figure 2.22). An increase in fluorescence was seen indicating the effectiveness of the assay. However, no fluorescence was seen when using ACA in media due to the inability for ACA to enter the cell.

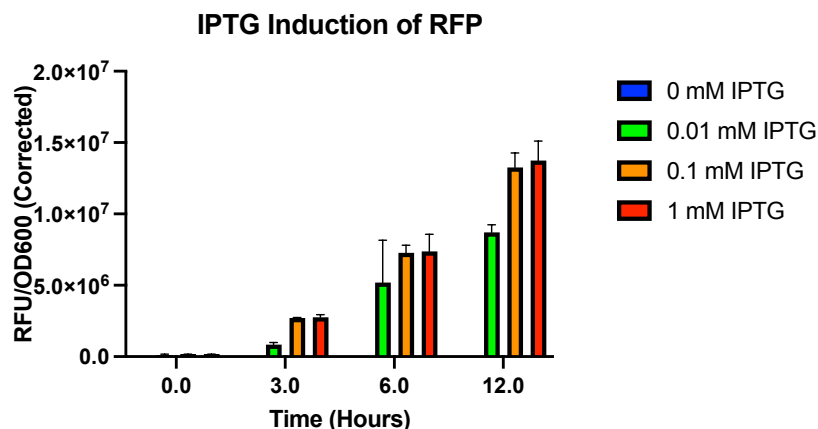


Figure 2.22. Induction of RFP by biosensor.

Although the original Cg10062 biosensor was not sensitive due to the lack of ACA cell permeation, we can apply the same method using Cg10062 externally. Here, MSA produced from Cg10062 (E114N) lysate may be added to media. Using Cg10062 (E114N) lysate supplemented media would allow us to screen for a more active variant than Cg10062 (E114N) using a similar fluorescent readout biosensor system as described previously (Figure 2.23). The addition of Cg10062 (E114N) crude lysate to media converts ACA in the media to MSA. MSA can then be transported into the cell. MSA is then converted to 3HP by MmsB within the cell. A plasmid designed with MmsR transcriptional regulator from *P. dentrificans* binds 3HP and form a 3HP-MmsR complex, which in turn binds to the P_{mmsA} promoter to turn on the transcription of red fluorescence protein.

Error prone mutagenesis may be implemented on Cg10062(E114N). Multiple colonies maybe selected for small scale 96-well plate expression and lysed. These lysates may be added

to *E. coli* cells containing the plasmids expressing *mmsB* and *mmsR* with *rfp*. Samples displaying an increase in fluorescence would indicate a more active variant for the conversion of ACA to MSA.

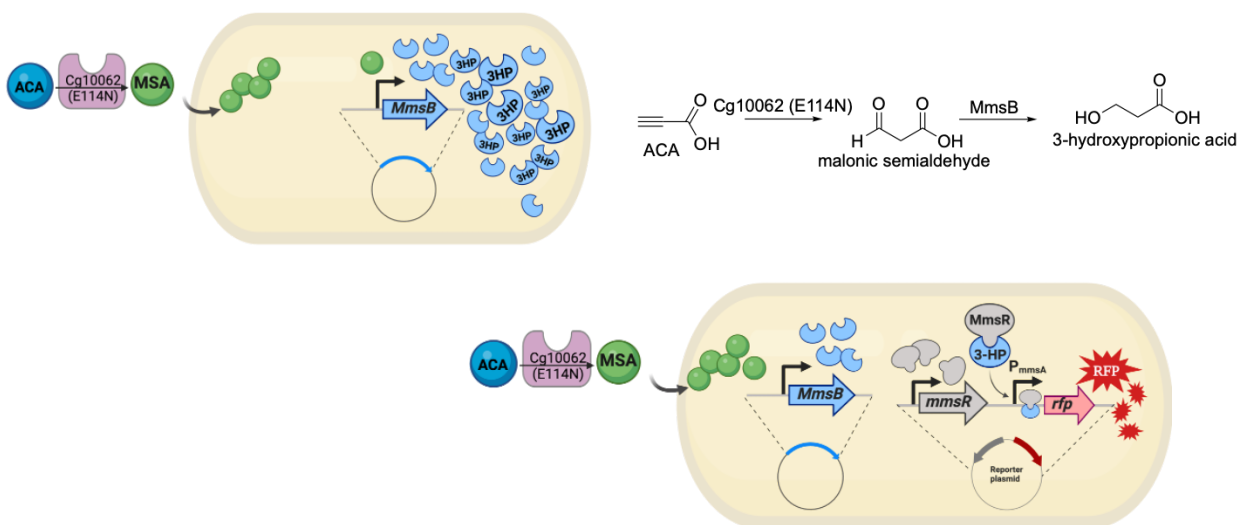


Figure 2.23. Lysate supplemented biosensor for Cg10062 (E114N).

2.5 α -Carbon Shift Leads to Product Change

To better understand the differing product profiles for wild-type *cis*-CaaD and Cg10062, we look to Glu-114. In *cis*-CaaD, the carboxyl of the side chain of residue Glu-114 is within hydrogen bonding distance of the Pro-1 nitrogen, which prevents the formation of the iminium tautomer of the acrylate intermediate. The acrylate intermediate is trapped in the enamine state (Figure 2.18) and can then form MSA as a major product after the addition of water. In Cg10062, the α -carbon of residue Glu-114 is shifted farther back and is unable to interact with the nitrogen of Pro-1 in the active site. This allows for iminium formation of the acrylate intermediate, and subsequent decarboxylation, which yields ACH as the major product. The C-terminal shifts with respect to the α -carbon of Glu-114, leading to the hypothesis that the C-terminal may be responsible for product specificity.

2.5.1 α -Carbon Shift Alters Product Profiles

We examined the structures of *cis*-CaaD and Cg10062 to understand why two similar enzymes with the same mechanism follow two different native pathways. When aligning variants, specifically the Y103F variant, of Cg10062 and *cis*-CaaD, we noticed a shift in the α -carbon position of residue 114, as well as a shift in a nearby C-terminal loop starting at residue Pro-120 (Figure 2.24). The Y103F variants of *cis*-CaaD and Cg10062 follow wild-type in major product formation. Cg10062 (Y103F) forms predominantly ACH, while *cis*-CaaD (Y103F) forms exclusively MSA. However, the Y103F variants have slower activities and are therefore easier to obtain substrate-soaked crystal structures than wild-type.

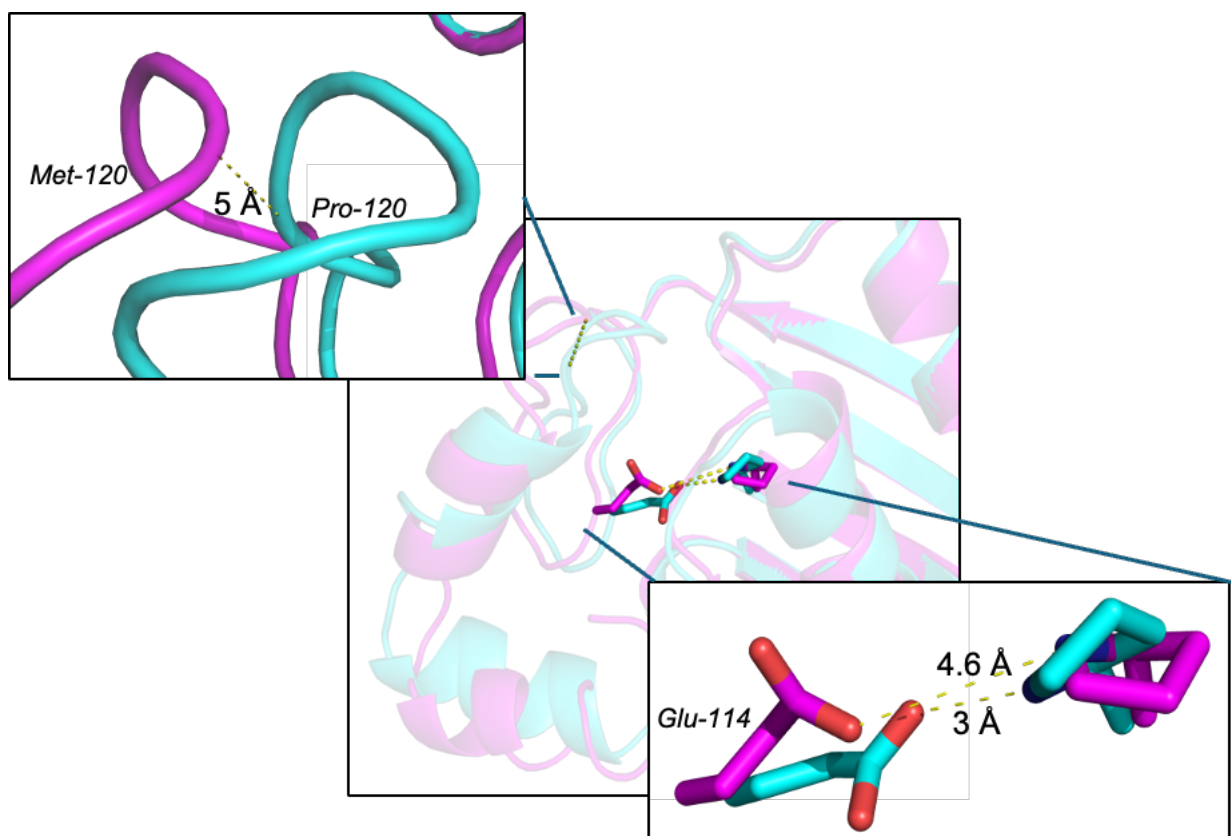


Figure 2.24. Alignment of Y103F variants of Cg10062 (pink) and *cis*-CaaD (blue). There is a 1.4 Å shift in the α -carbon position of residue Glu-114 and around 5 Å shift in the loop at residue 120.

In *cis*-CaaD (Y103F), the α -carbon of the Glu-114 is about 1.6 Å closer to Pro-1 and the carboxyl of Glu-114 is within hydrogen bonding distance of the nitrogen of Pro-1, which prevents the iminium form of the acrylate intermediate (Figure 2.18). Since a protonated nitrogen of a secondary amine cannot form an iminium, hydrogen bonding between the carboxyl of Glu-114 and the nitrogen of Pro-1 would be expected to suppress iminium formation and therefore inhibit the decarboxylation reaction. The acrylate intermediate is trapped in the enamine state leading to hydration and formation of MSA after the addition of water. In Cg10062 (Y103F), the α -carbon of residue Glu-114 is shifted farther back and is unable to interact with the nitrogen of Pro-1 in the active site allowing tautomerization to the iminium and decarboxylation to yield ACH. We can also see this shift between the carboxyl group of Glu-114 in Cg10062 (wild-type) as opposed to *cis*-CaaD (wild-type) (Figure 2.25). The carboxyl of Glu-114 of *cis*-CaaD (wild-type) is within hydrogen bonding distance of Pro-1 which prevents iminium formation and therefore prevents iminium formation, decarboxylation, and ACH formation. The acrylate remains in the enamine form leading to hydration and MSA formation. However, in Cg10062 (wild-type), the carboxyl of Glu-114 is farther away which allows for tautomerization to the iminium form and therefore decarboxylation to ACH ensues. However, when Glu-114 is mutated to the shorter asparagine residue in Cg10062 (E114N), there is room for a catalytic water molecule (Figure 2.25). Water mediated hydrogen bonding occurs, and iminium formation is prevented. Therefore, no decarboxylation occurs and the hydration pathway is followed by direct water addition to form MSA.

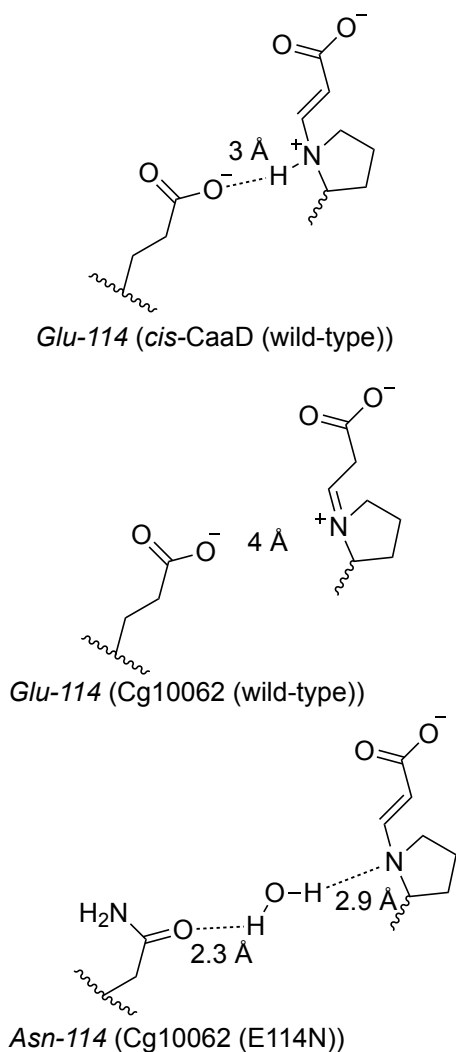


Figure 2.25. Hydrogen bonding for different enzymes.

We can see this mechanism followed through hydrogen bond distance correlation with other variants crystalized (Table 2.4). A distance between residue 114 and Pro-1 within hydrogen bonding distance indicated hydration product, MSA. A distance between residue 114 and Pro-1 outside hydrogen bonding distance, and without a water molecule between the two residues, indicated decarboxylation. If there was a water molecule within hydrogen bonding distance between residue 114 and Pro-1, water mediated hydrogen bonding was considered for hydration product. In structures such as *cis*-CaaD (E114N) and *cis*-CaaD (E114Q) we see a water molecule within the active site. However, it is too far for water coordinated hydrogen bonding. This allowed

for decarboxylation and a majority of ACH production. In addition, we see a 5 Å shift in the C-terminal at residue 120 (proline in *cis*-CaaD and methionine in Cg10062) with respect to the shift seen at the α -carbon of residue 114, leading to the hypothesis that the C-terminal may be responsible for product specificity. In our efforts to discover a more active *cis*-CaaD that solely produces MSA, we were able to discover why Cg10062 and *cis*-CaaD have different product profiles.

Table 2.4. Distances between residue 114 and Pro-1 for crystal structures evaluated.

Mutant	Activity	Soak	Ligand			Catalytic Water			Distance between Residue-114 and Pro-1 (Å)		
			A	B	C	A	B	C	A	B	C
<i>cis</i> -CaaD (wild-type)	Hydration	ACA	ACA	ACA	ACA	No	No	No	3.34	3.09	3.10
		CCA	Acetate	-	-	No	-	-	3.17	-	-
<i>cis</i> -CaaD (E114D)	Hydration	APO	SO ₄	SO ₄	SO ₄	Yes	Yes	Yes	4.46	4.35	4.41
		ACA	ACR	ACR	ACA	Yes	No	Yes	4.13	4.5	3.94
		CCA	SO ₄	SO ₄	SO ₄ / ETH	Yes	Yes	No	4.39	4.31	4.31
<i>cis</i> -CaaD (E114N)	Hydration/ Decarboxylation	APO	SO ₄	SO ₄	SO ₄	No	No	Yes	4.76	4.19	6.10
		ACA	ACA	MSA	MSA	Yes	Yes	No	5.02	4.44	4.02
<i>cis</i> -CaaD (E114Q)	Hydration/ Decarboxylation	APO	SO ₄	SO ₄	SO ₄	No	No	No	3.14	3.18	3.27
		ACA	HPA	HPA	HPA	No	No	No	3.40	3.42	3.44
		CCA	ACR	HPA	SO ₄ / ETH	No	No	No	2.99	3.63	3.47
<i>cis</i> -CaaD (Y103F)	Hydration	APO	SO ₄	SO ₄	SO ₄	No	No	No	2.98	2.85	3.29
		ACA	SO ₄	SO ₄	SO ₄	No	No	No	3.53	2.82	3.02
		CCA	H ₂ O	H ₂ O	H ₂ O	No	No	No	2.92	2.95	3.32
<i>cis</i> -CaaD (H28A)	-	ACA	H ₂ O	Acetate	Acetate/ H ₂ O	No	No	No	3.01	3.19	2.97
<i>cis</i> -CaaD (T34A)	Hydration	APO	SO ₄	SO ₄	SO ₄	No	No	No	3.38	3.34	3.08
		ACA	SO ₄	SO ₄	SO ₄	No	No	No	3.21	3.00	3.08
		CCA	SO ₄	SO ₄	SO ₄	No	No	No	3.30	2.89	2.86
<i>tcis</i> -CaaD	Hydration	CCA	SO ₄	-	-	No	-	-	3.38	-	-
<i>N-cis</i> -CaaD/C-Cg10062	Hydration	ACA	Cl ⁻	-	-	No	-	-	3.01	-	-
Cg10062 (wild-type)	Hydration/ Decarboxylation	ACA	H ₂ O	-	-	No	-	-	4.37	-	-
Cg10062 (E114N)	Hydration	APO	SO ₄	SO ₄	SO ₄	No	No	No	5.78	5.51	5.05
		ACA	SO ₄ *	ACR*	HPA*	Yes	Yes	No	5.35	5.06	4.89
Cg10062 (E114D)	Hydration	ACA	ACA*	MSA*	MSA*	Yes	Yes	Yes	5.01	5.14	4.81
Cg10062 (Y103F)	Hydration/ Decarboxylation	ACA	Empty	-	-	No	-	-	4.63	-	-

2.6 C-Terminal is Necessary for Activity

To determine the role of the C-terminal on activity, chimera swaps between the C-termini of *cis*-CaaD and Cg10062 at amino acid residue 120, as well as a truncation of *cis*-CaaD was performed. A gene chimera consisting of amino acids 1-119 of *cis*-CaaD followed by amino acids 120-148 of Cg10062 included NdeI and XhoI restriction sites, respectively, at the 5' and 3' positions. A gene chimera consisting of amino acids 1-119 of Cg10062 followed by amino acids 120-149 of *cis*-CaaD was designed similarly. A truncated gene consisting of amino acids 1-119 was designed and called *tcis*-CaaD. Each gene was cloned into pET-21a(+) between the NdeI and XhoI restriction sites. Crystal structures of the *N-cis*-CaaD/C-Cg10062 (*cis*-CaaD with Cg10062 C-terminal) as well as truncated *cis*-CaaD were obtained, although density of the C-terminal was unclear for *N-cis*-CaaD/C-Cg10062 (Figure 2.26). A swap between Cg10062 and *cis*-CaaD, (*N*-Cg10062/C-*cis*-CaaD), was also performed but we were unable to obtain a crystal structure of this chimera however kinetic results were obtained (Table 2.5). From these results, we can determine that the C-terminal directs product formation by shifting the α -carbon of Glu-114 closer to or farther from the active site. This allows hydrogen bonding between the carboxyl of Glu-114 and the amine of Pro-1.

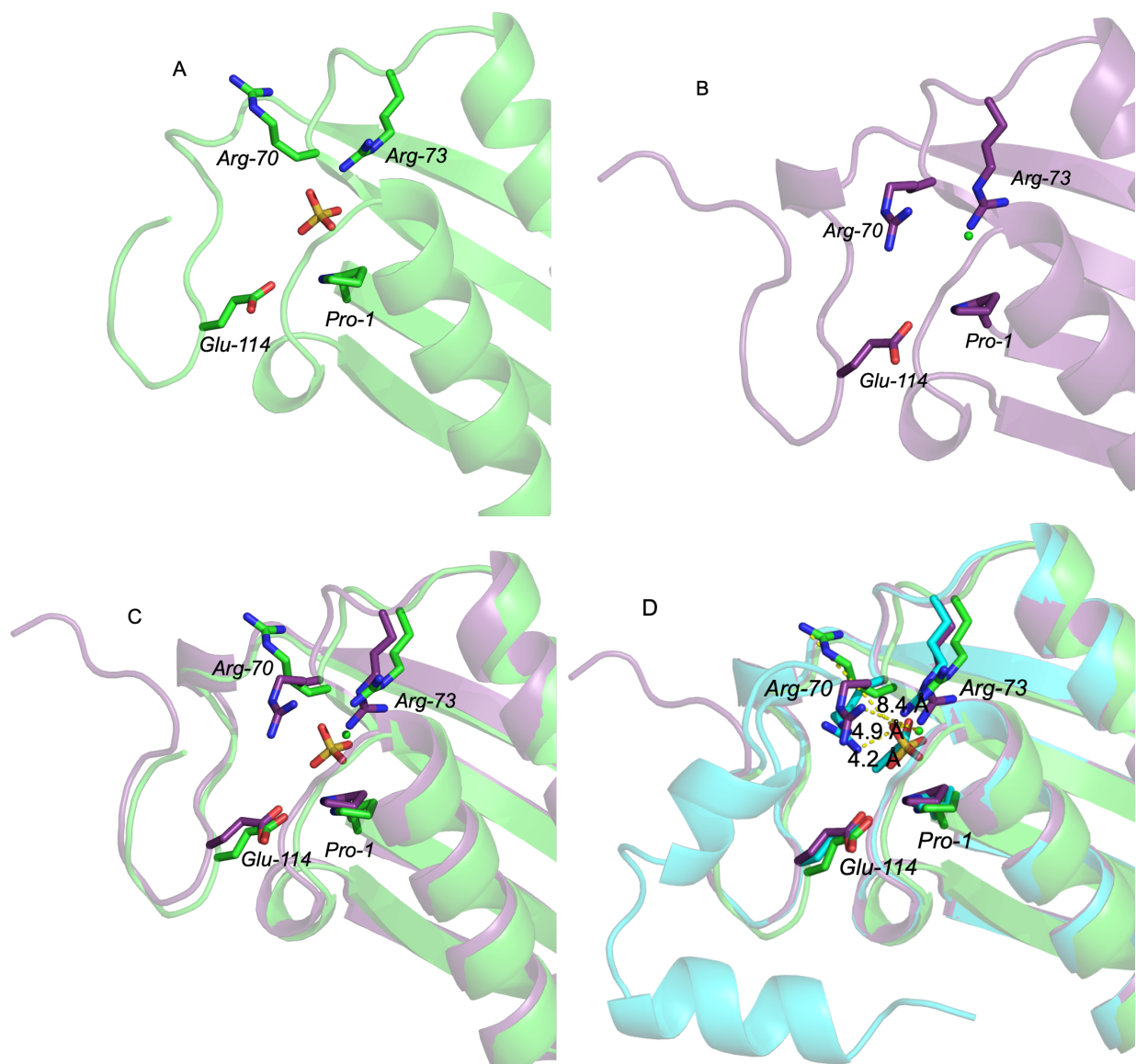


Figure 2.26. The crystal structure of (A) truncated *cis*-CaaD that has a sulfate in the active site. *N-cis*-CaaD/C-Cg10062 shows (B) a chloride ion trapped in the active site. The C-terminal is disordered (C), only the first four amino acids of the Cg10062 C-terminal sequence provide defined electron density. An overlay of the two structures with wild-type *cis*-CaaD (blue), reveals (D) Arg-70 of truncated *cis*-CaaD flipped out of the active site upon disorder of the C-terminal.

2.6.1 C-Terminal Chimeras Kinetics

When the C-terminal is removed, there is no activity establishing it as essential for activity, even though active site amino acids were unchanged from native. Although the active site residues in *N-Cg10062/C-cis-CaaD* are that of wild-type Cg10062, we see almost exclusively

MSA production. This may indicate the greater importance of the C-terminal in directing the α -carbon closer for hydrogen bonding than the E114 variant within the active site. *N-Cg10062/C-cis-CaaD* also shows higher catalytic activity for ACA than *N-cis-CaaD/C-Cg10062* however lower than that of each individual wild type. Inactivity and low activity in variants with manipulations of the C-terminal endorse the importance of the C-terminal starting at residue 120 on the activity of *cis-CaaD* and *Cg10062*. Furthermore, the findings here reveal the essential role of the C-terminal in both maintaining activity and directing product formation.

Table 2.5. Michaelis-Menten kinetics of chimera and truncated variants compared to wild-type *cis*-CaaD and Cg10062.

Enzyme Variant	ACA					CCA				
	K_m (μ M)	K_{cat} (s ⁻¹)	$K_{cat}/K_m \times 10^4$ (M ⁻¹ s ⁻¹)	Product Ratio (%)		K_m (μ M)	K_{cat} (s ⁻¹)	$K_{cat}/K_m \times 10^4$ (M ⁻¹ s ⁻¹)	Product Ratio (%)	
				MSA	ACH				MSA	ACH
<i>cis</i> -CaaD (wild-type)	1050 ± 109	1.40 ± 0.05	0.13 ± 0.01	>99	<1	372 ± 41	2.16 ± 0.07	0.58 ± 0.05	>99	<1
Cg10062 (wild-type)	66 ± 17	4.2 ± 0.28	6.3 ± 1.2	15	85	15800 ± 2940	4.13 ± 0.29	0.03 ± 0.00	14	86
N-Cg10062/ C- <i>cis</i> -CaaD	1330 ± 363	1.02 ± 0.11	0.08 ± 0.01	96	4	-	-	-	95	5
N- <i>cis</i> -CaaD/ C-Cg10062	-	-	-	96	4	-	-	-	>99	<1
<i>tcis</i> -CaaD	-	-	-	60	40	-	-	-	61	39

2.6.2 Arg-70 Movement

Within the chimeras there is a distinct repositioning likely due to the C-terminal's involvement of orientating Arg-70 into placement in the active site. However, we see the flexibility of Arg-70 in all structures of *cis*-CaaD wild-type and variants studied. In an overlay of all *cis*-CaaD structures, we can see the movement and flexibility of residue Arg-70 in the presence of bound substrate, whether ACA or intermediate (Figure 2.27). Arg-70 is positively charged and necessary to facilitate the binding of the negatively charged substrate. The highly negative charge on the substrate may facilitate the movement of Arg-70 towards the substrate to position the substrate into the active site. Structural movement and positions play an important role in the functioning of *cis*-CaaD and Cg10062. The α -carbon directs product formation, the C-terminal is required for activity and directing the α -carbon into position, and Arg-70 is necessary for positioning the substrate into the active site. Overall, these precise structural components are essential for ensuring effective substrate binding and product formation.

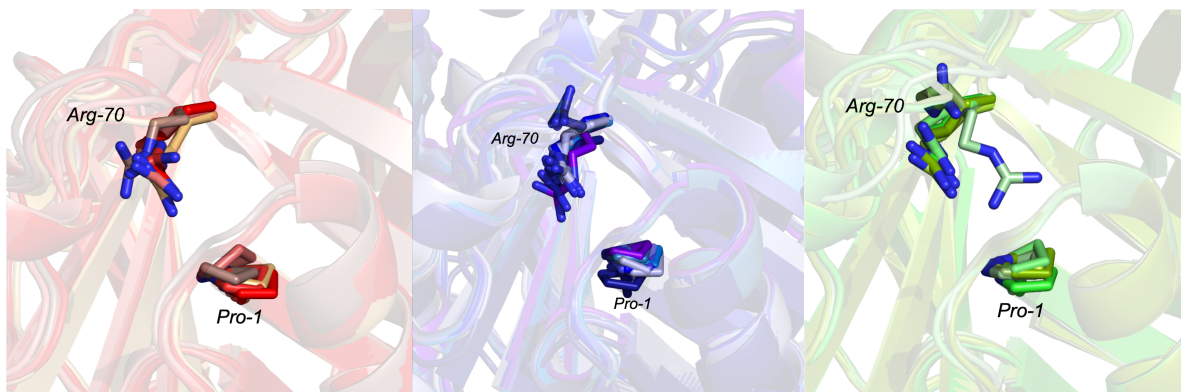


Figure 2.27. Arg-70 residue movement for all *cis*-CaaD crystal structure when in apo form (red), ACA soaked (blue), and CCA soaked (green).

REFERENCES

- Schroeder, G. K.; Huddleston, J. P.; Johnson, W. H., Jr; Whitman, C. P. A Mutational Analysis of the Active Site Loop Residues in *cis*-3-Chloroacrylic Acid Dehalogenase. *Biochem.* 2013, 52 (24), 4204–4216.
- de Jong, R. M.; Bazzacco, P.; Poelarends, G. J.; Johnson, W. H., Jr; Kim, Y. J.; Burks, E. A.; Serrano, H.; Thunnissen, A.-M. W. H.; Whitman, C. P.; Dijkstra, B. W. Crystal Structures of Native and Inactivated *cis*-3-Chloroacrylic Acid Dehalogenase. Structural Basis for Substrate Specificity and Inactivation by (R)-Oxirane-2-Carboxylate. *J. Biol. Chem.* 2007, 282 (4), 2440–2449.
- Huddleston, J. P.; Johnson, W. H., Jr; 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. *Biochem.* 2015, 54 (19), 3009–3023.
- Mathes Hewage, A.; Nayeibi Gavvani, H.; Chi, D.; Qiu, B.; Geiger, J. H.; Draths, K. Cg10062 Catalysis Forges a Link between Acetylenecarboxylic Acid and Bacterial Metabolism. *Biochem.* 2021. <https://doi.org/10.1021/acs.biochem.1c00524>.
- Guo, Y.; Serrano, H.; Poelarends, G. J.; Johnson, W. H., Jr; Hackert, M. L.; Whitman, C. P. Kinetic, Mutational, and Structural Analysis of Malonate Semialdehyde Decarboxylase from *Corynebacterium* Strain FG41: Mechanistic Implications for the Decarboxylase and Hydratase Activities. *Biochem.* 2013, 52 (28), 4830–4841.
- Mathes Hewage, A., Silva, K., Kwiatkowski, K., Lee, H., Sreedhar, D., Gavvani, H. Geiger, J., Draths, K. An Original Biosynthetic Route to 3-Hydroxypropionic Acid, 2025, *in preparation*
- Sevastik, R.; Whitman, C. P.; Himo, F. Reaction Mechanism of *cis*-3-Chloroacrylic Acid Dehalogenase: A Theoretical Study. *Biochem.* 2009, 48 (40), 9641–9649.
- Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, W. 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. *Biochem.* 2004, 43 (3), 759–772.
- Guo, Y.; Serrano, H.; Johnson, W. H., Jr; Ernst, S.; Hackert, M. L.; Whitman, C. P. Crystal Structures of Native and Inactivated *cis*-3-Chloroacrylic Acid Dehalogenase: Implications for the Catalytic and Inactivation Mechanisms. *Bioorg. Chem.* 2011, 39 (1), 1–9.
- Silva, K.; Al-Ahmad, N.; Geiger, J. H.; Draths, K. Elucidating Structural and Activity Differences Between Tautomerases *cis*-CaaD And Cg10062: Implications For 3-Hydroxypropionic Acid Production From Acetylenecarboxylic Acid *Biochem.* 2025. *in preparation*
- Liang, J.; Han, Q.; Tan, Y.; Ding, H.; Li, J. Current Advances on Structure-Function Relationships of Pyridoxal 5'-Phosphate-Dependent Enzymes. *Front. Mol. Biosci.* 2019, 6, 4.

12. 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. [https://doi.org/10.1016/S1570-9639\(02\)00533-2](https://doi.org/10.1016/S1570-9639(02)00533-2).
13. 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. <https://doi.org/10.1021/acssynbio.6b00007>.
14. Shafikhani, S.; Siegel, R. A.; Ferrari, E.; Schellenberger, V. Generation of Large Libraries of Random Mutants in *Bacillus Subtilis* by PCR-Based Plasmid Multimerization. *Biotechniques* 1997, 23 (2), 304–310.
15. Innis, M. A.; Gelfand, D. H.; Sninsky, J. J.; White, T. J. PCR Protocols: A Guide to Methods and Applications. *Academic Press*. 2014.
16. Manual, I. GeneMorph II Random Mutagenesis Kit. <https://www.agilent.com/cs/library/usermanuals/public/200550.pdf>.
17. Mor, R. A Review on Development of Enzymatic Biosensors for Industrial Applications. *World J. Adv. Res. Rev.* 2024, 23 (1), 2582–2590.
18. Nasu, Y.; Murphy-Royal, C.; Wen, Y.; Haidey, J. N.; Molina, R. S.; Aggarwal, A.; Zhang, S.; Kamijo, Y.; Paquet, M.-E.; Podgorski, K.; Drobizhev, M.; Bains, J. S.; Lemieux, M. J.; Gordon, G. R.; Campbell, R. E. A Genetically Encoded Fluorescent Biosensor for Extracellular L-Lactate. *Nat. Commun.* 2021, 12 (1), 7058.
19. Nguyen, N. H.; Kim, J.-R.; Park, S. Development of Biosensor for 3-Hydroxypropionic Acid. *Biotechnol. Bioprocess Eng.* 2019, 24 (1), 109–118.
20. Umeno, D.; Hiraga, K.; Arnold, F. H. Method to Protect a Targeted Amino Acid Residue during Random Mutagenesis. *Nucleic Acids Res.* 2003, 31 (16), e91.
21. Mathes Hewage, A. Synthesis of 3-Hydroxypropionic Acid from Acetylenecarboxylic Acid. *Michigan State University ProQuest Dissertations & Theses*. 2022. 29390957.

**CHAPTER THREE: Application of Cg10062 (E114N) to 3-Hydroxypropionic Acid
Synthesis**

3.1. Cg10062 (E114N) Use in the Production of Commodity Chemicals

Previous research has been performed using the modified enzyme Cg10062 (E114N) in the production of 3-hydroxypropionic acid (3HP).¹ We have targeted 3HP as a building block chemical due to its versatility in the further production of acrylamide, acrylic acid, and 1,3-propanediol.^{2,3} All of which are important chemicals in the production of textiles, adhesives, fragrances, and more. Previous *in vitro* studies couple Cg10062 (E114N) with YdfG from *E. coli* for the conversion of acetylenecarboxylic acid (ACA) to 3HP.⁵ This reaction is dependent on NADPH which was regenerated using an engineered NADP⁺ and phosphite dependent phosphite dehydrogenase, PTDH.⁴ In a second synthesis, MmsB from *P. putida* was used in place of YdfG. MmsB is dependent on NADH and NAD⁺ reducing soluble hydrogenase was used to couple the reaction and regenerate NADH cofactor.^{5,6} Additionally, a second enzyme Ynel from *E. coli* partakes in the direct conversion of MSA to malonic acid, another high value chemical of interest.⁷ Although previous efforts have shown Ynel to have low activity when purified in *in vitro* studies, Ynel may be utilized *in vivo* for the bioconversion of MSA to malonic acid.

3.1.1 Attempts at Bacterial Growth on ACA

In many cases, *in vitro* synthesis using isolated enzymes or cell-free systems leads to low efficiency and poor product yield due to manipulations of reaction conditions inhibit optimal enzyme function. By utilizing the full metabolic and regulatory network of the microbe, *in vivo* production may be able to increase production of chemicals by the microorganism. However, small genetic manipulations may affect yields as well. In addition, cost of production decreases with *in vivo* systems as a need to purify individual protein enzymes, an often costly and time-consuming process, is not required. *In vivo* processes also benefit from the ability to integrate multiple enzymatic steps into a single organism. This reduces the need for complex reaction cascades and optimizing reaction efficiency. Furthermore, microorganisms can often be grown on substrate feedstocks or evolved to grow on feedstocks enhancing the viability of this approach. Nevertheless, *in vivo* processes require further refinement involving metabolic engineering and

bioprocess optimization to enhance product formation, scalability and economic feasibility of using microbial production systems.

Previous *in vivo* attempts at the production of 3HP from ACA using Cg10062 (E114N) were unsuccessful. This is due to the decreased viability of *E. coli* using ACA as a carbon source. ACA is unable to be transported into the cell by any native system, therefore the *in vivo* production of 3HP is restricted. Extensive evaluation of growth of other mostly soil bacterial species produced similar results. Originally, soil samples from four separate locations (an apple orchard, corn field, pond, and pond drain) were obtained and evaluated for growth on ACA as a carbon source. When this proved unsuccessful, seventeen isolated bacterial strains were screened for growth on ACA. Glycerol stocks of each bacterial species were streaked out onto LB, M9, M9/Glu, or M9/ACA plates and evaluated for growth according to their required incubation conditions (Table 3.1). All bacterial species grew on rich LB carbon source with tryptone and yeast extract within 36 hours while twelve showed promise for growth on ACA as shown by growth on M9/ACA media plates. The positive growth on M9 with no carbon source is likely due to growth on thiamine that is added to M9 salts. To confirm growth on ACA, three colonies from each of the twelve organisms were inoculated into LB overnight. Cultures were then washed in M9 salts and inoculated to an OD₆₀₀ of 0.05 in either M9/Glu or M9/ACA both lacking thiamine. While growth was seen on culture plates, no growth was seen in liquid media for M9/ACA showing that no bacterial species evaluated were suitable for *in vivo* shake flask culture studies using ACA as a carbon source. Therefore, we maintain our focus on *E. coli* for its versatility and ease in use for bioengineering and look to other substrates as a carbon source for entry into the cell. This directs our attention to malonic semialdehyde (MSA); the first intermediate in the pathway from ACA to high value chemicals such as 3HP.

Table 3.1. Bacterial species growth on different solid medias. X indicates growth on the respective media.

Species	LB	M9/Glu	M9	M9/ACA
<i>Pseudomonas fluorescens Pf-5</i>	X	X	X	X

Table 3.1. (cont'd)

<i>Pseudomonas fluorescens</i> BL915	X	X		
<i>Rhodopseudomonas palustris</i>	X	X	X	X
<i>Pseudomonas putida</i>	X	X	X	X
<i>Pseudomonas denitrificans</i>	X	X	X	
<i>Zymomonas mobilis</i>	X			
<i>Enterobacter cloacae</i>	X			
<i>Pseudomonas chlororaphis</i> subsp. <i>aureofaciens</i>	X	X	X	X
<i>Bacillus subtilis</i>	X	X	X	X
<i>Asperigillus flavus</i>	X	X		X
<i>Bacillus pumilus</i>	X	X	X	X
<i>Corynebacterium glutamicum</i>	X			
<i>Bacillus cereus</i> UW85	X	X	X	X
<i>Candida parapsilosis</i>	X	X	X	X
<i>Klebsilla oxytoca</i>	X	X	X	X
<i>Bacillus circulans</i>	X	X	X	X
<i>Citrobacter braakii</i> ,	X	X	X	X

3.2 Short Term Adaptive Laboratory Evolution (STALE): *E. coli* Growth on Medias

Due to *E. coli*'s nonviability on ACA as a carbon source, we looked to MSA as a carbon feedstock. MSA is not easily obtained commercially, therefore it's production and utilization *in situ* was favored. ACA is a stable and commercially available chemical that can easily be added to media. Hypothetically, ACA can then be converted in media to MSA and the MSA produced can be utilized by cells. To determine how high of an ACA concentration and low of a concentration of minimal media at which cells were viable at, a series of short-term adaptive growth studies were carried out. Varying concentrations of minimal media containing either glucose or glycerol with ACA were inoculated with MG1655 *E. coli* cell and allowed to grow over 48 hours (Figure 3.1). At 16 mM ACA (4 mM Glucose or Glycerol) growth starts to decline. Therefore, a concentration of 4 mM glucose or glycerol is the lowest concentration that can be added to maintain cell growth. ACA of 16 mM added is the maximum amount of ACA that can be added to media to maintain cell growth. We can assume that cells will grow on preferred media, glucose or glycerol, first then utilize the secondary carbon source in media. This can be used for the secretion of recombinant proteins, where glucose is used to express a protein of interest. Cell can use the glucose in media

first to express Cg10062 (E11N) then this enzyme can use ACA in media to make MSA which can be taken into cell (Figure 3.2). However, Cg10062 (E114N) must be expressed outside of the cell for this process to occur. The addition of a secondary enzyme such as MmsB may be included to further the production of 3-hydroxypropionic acid from MSA that has been brought into the cell.

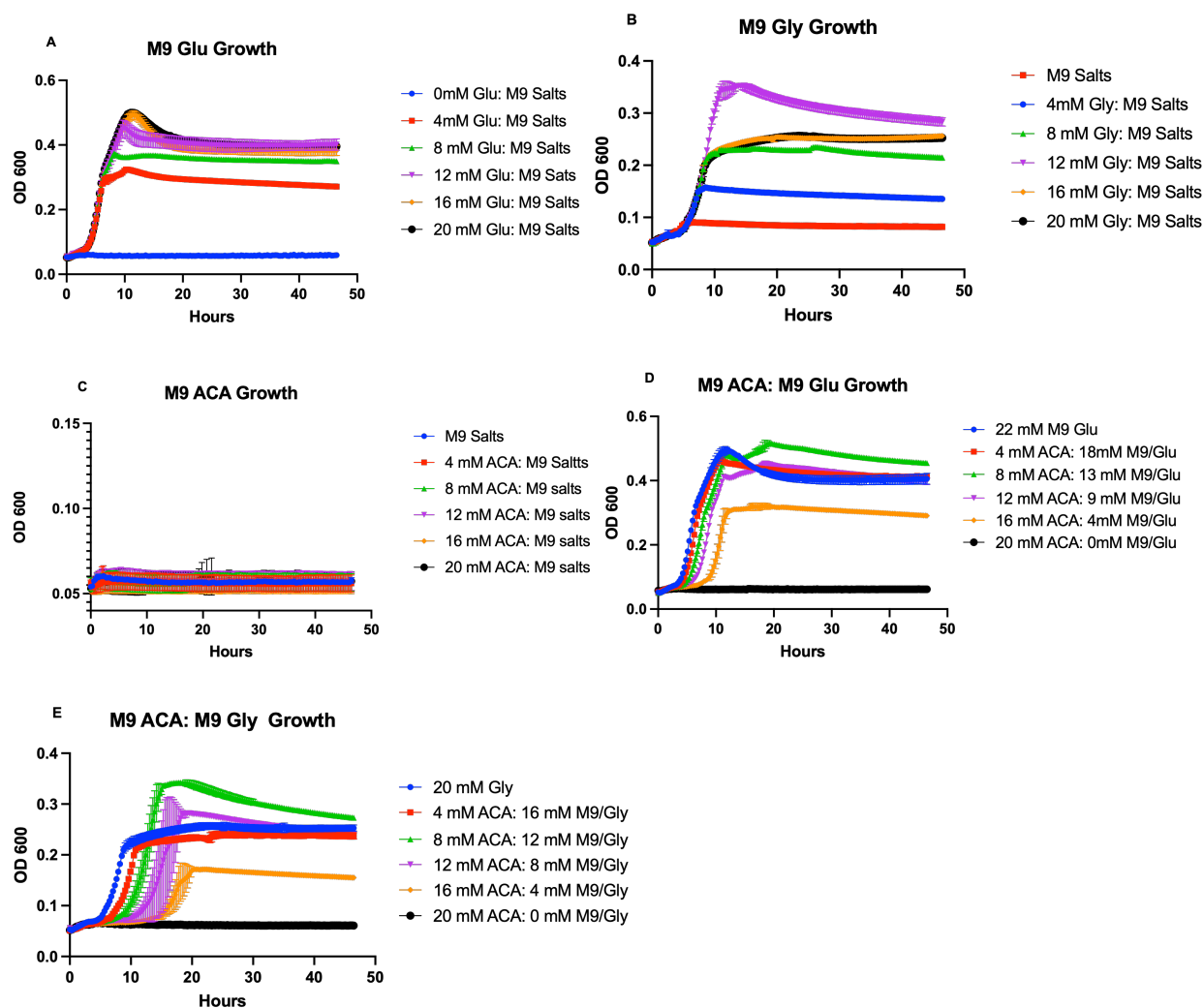


Figure 3.1. *E. coli* growth on different medias (A) M9/Glu, (B) M9/Gly, (C) M9/ACA, (D) M9/ACA with M9/Glu, (E) and M9/ACA with M9/Gly.

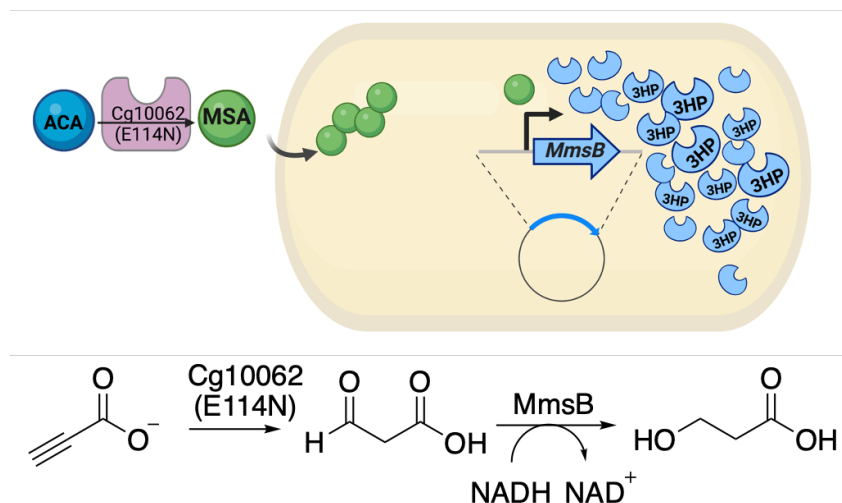


Figure 3.2. Malonic semialdehyde transport into the cell and further production of 3-hydroxypropionic acid.

3.2.1 Lysate Supplemented Growth on MSA

In order to ensure Cg10062 (E114N) was expressed outside of the cell media was supplemented with Cg10062 (E114N) lysate. A plasmid containing Cg10062 (E114N, pKS4.1114) was transformed into BL21(DE3) cells, expressed and lysed. An overnight culture of MG1655 *E. coli* was reinoculated to an OD_{600} of 0.05. The Cg10062 (E114N) lysate was added to M9/ACA media and MG1655 *E. coli* overnight culture. OD_{600} and ^1H NMR were taken at multiple timepoints (Figure 3.3). ^1H NMR spectra confirm the turnover of ACA to MSA. The immediate increase in OD_{600} signifies that cells were viable with MSA as a carbon source. There was also a color change in the media to a bright yellow signifying MSA production. To confirm MSA entrance into the cell and the ability to produce further chemicals in vivo, BL21(DE3) cells were transformed with a plasmid (pKK1.1025) encoding MmsB gene for a 3-hydroxyisobutyrate dehydrogenase which is known to convert MSA to 3HP. Similarly, lysate of enzyme Cg10062 (E114N) was added to M9/ACA media and used to inoculate an overnight culture of BL21(DE3)/pKK1.1025. OD_{600} and ^1H NMR were measured. Ethanol production is seen via ^1H NMR analysis (Figure 3.4). This is due to decarboxylation of MSA in cell and subsequent conversion to ethanol by anaerobic fermentation by *E. coli*.

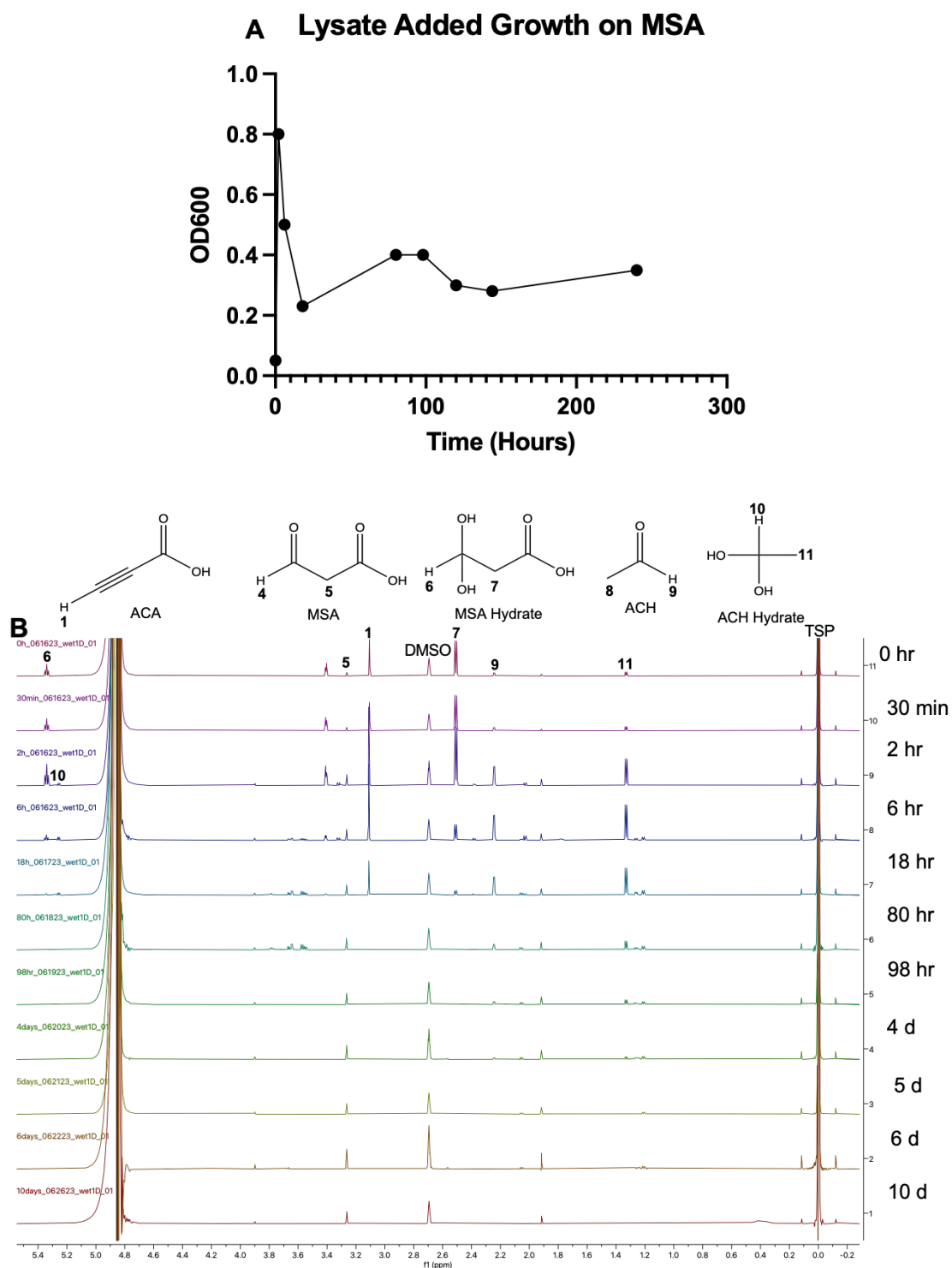


Figure 3.3. (A) Growth curve and (B) ^1H NMR analysis of products of MG1655 *E. coli* cells when grown on M9/ACA media containing Cg10062 (E114N) lysate in media.

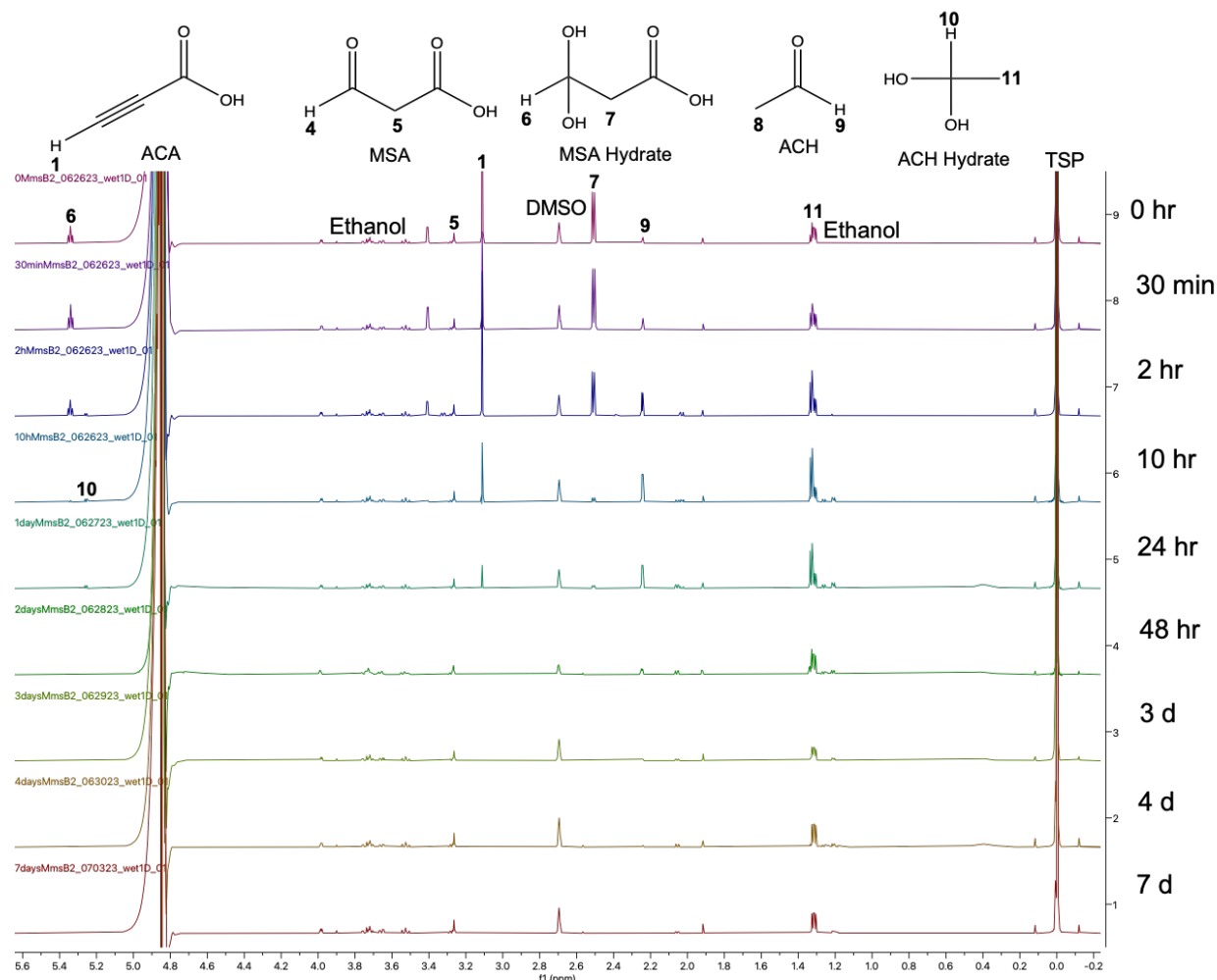


Figure 3.4. ^1H NMR analysis of products of BL21(DE3) *E. coli* cells encoding MmsB for the production of ethanol when grown on M9/ACA media containing Cg10062 (E114N) lysate in media.

3.3 A Novel Cg10062 (E114N) Fusion Protein

To increase efficiency and minimize the need for added lysate to produce MSA in media, a Cg10062 (E114N) fusion protein was designed. This fusion protein comprises Cg10062 (E114N) tethered to the surface of the *E. coli* membrane so that conversion of ACA to MSA occurs near the cell and fast entrance of MSA into the cell occurs (Figure 3.5). With the transformation of a plasmid containing MmsB, any MSA that has entered the cell can be converted to 3HP. 3HP can then further be used as a carbon source within the cell or can be harvested. This allows a one cell system for the conversion of ACA to high value chemicals *in vivo* (Figure 3.6).

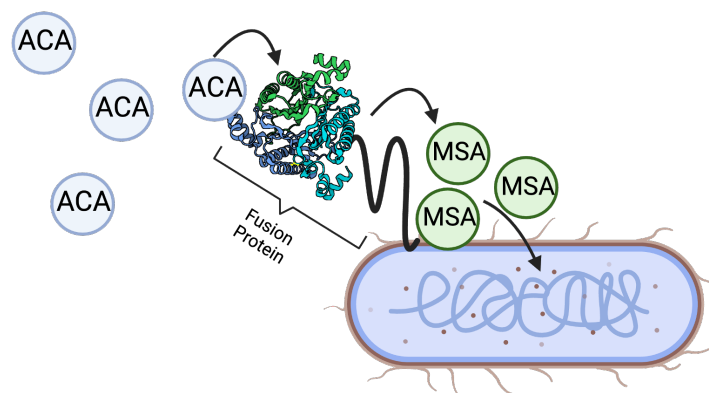


Figure 3.5. Cg10062 (E114N) fusion protein conversion of ACA to MSA outside the cell and subsequent transfer of MSA into the *E. coli* cell.

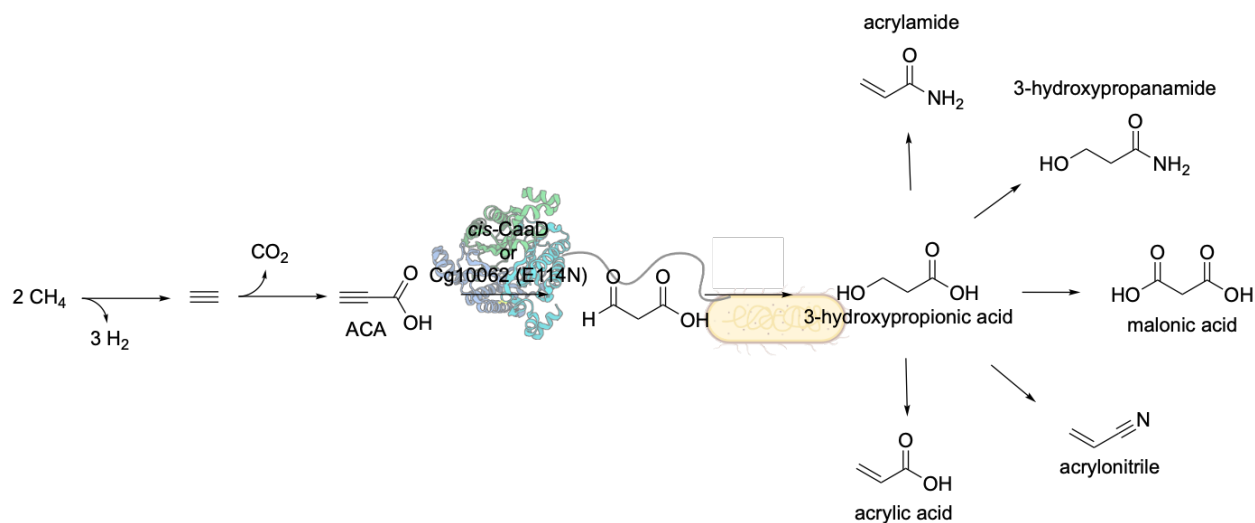


Figure 3.6. Cg10062 (E114N) fusion protein in the large-scale production of high value chemicals from CH₄ and CO₂ starting products.

3.3.1 pAIDA System to Produce Cg10062 (E114N) Fusion *In Vivo*

In the design of this Cg10062 (E114N) fusion protein, we look to the AIDA auto transport system. The glycosylated autotransporter adhesin involved in diffuse adherence (AIDA)-I system consists of two genes that encode specific autotransporter.⁸ The surface translocation system of AIDA-I is commonly used in the surface expression of recombinant proteins. It is characterized

by a conserved sequence with a *N*-terminal signal peptide, passenger protein (Cg10062 (E114N)), linker region, and C-terminal translocation unit (Figure 3.7).⁹ It is accepted that the conserved sequence uses the Sec-mediated transport system is native to *E. coli* to pass through the inner membrane.¹⁰ A 147 bp signal peptide is cleaved by signal peptidase 1 at AXA site.^{11,12} This leaves the catalytic N-terminal Pro-1 of Cg10062 (E114N) free for catalysis. This makes the pAIDA system optimal for use with Cg10062 as opposed to other transport systems when the *N*-terminal is not left free. The C-terminal translocation unit then inserts itself into the outer membrane, forming a β -barrel pore and pulls the linker and Cg10062 (E114N) through the pore to the surface to be exposed (Figure 3.8). The optimization of this system and its use in the fusion of many protein fusions have deemed this a successful option in the secretion of recombinant surface proteins.^{13,14} A plasmid (pKS8.0410) was designed and Cg10062 (E114N) was inserted using Gibson assembly as the passenger protein. Using the media concentration determined above, M9 salts containing 4mM Glu and 16 mM ACA and antibiotic were used to express the fusion protein using glucose as a food source for expression and ACA for the conversion to MSA by the AIDA/Cg10062 (E114N) fusion protein.

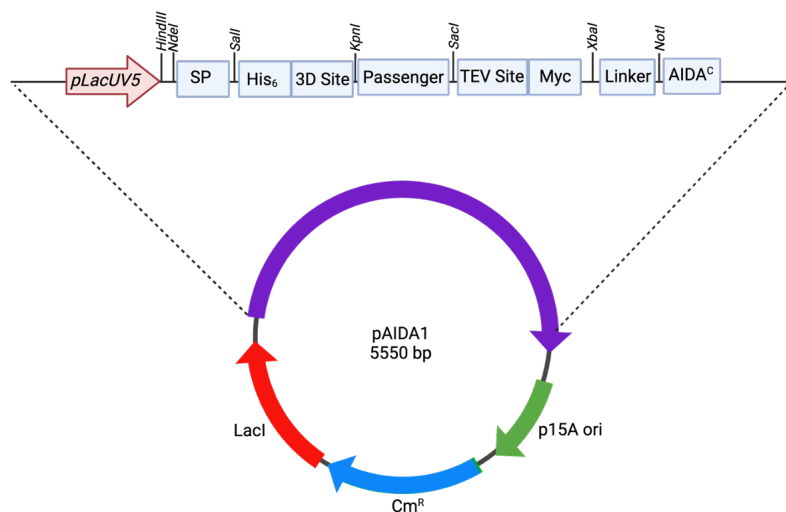


Figure 3.7. pAIDA1 expression vector.¹³

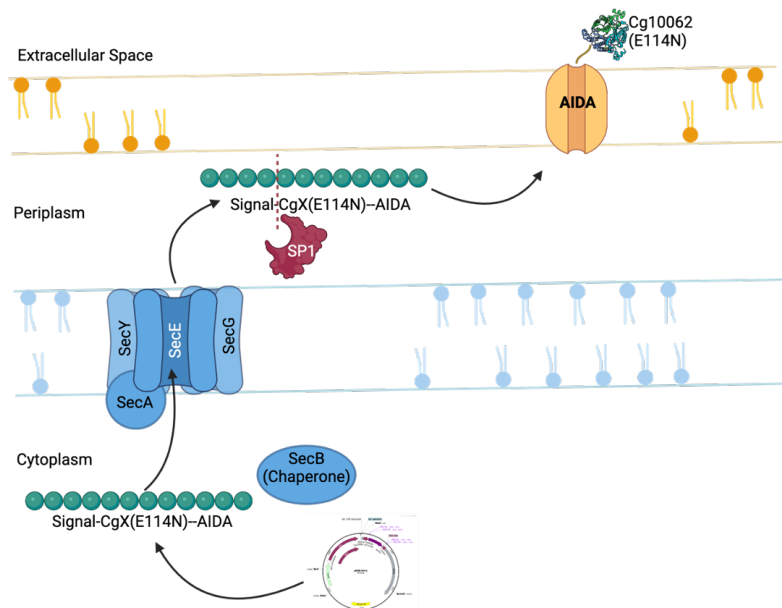


Figure 3.8. AIDA/Cg10062 (E114N) fusion expression and transport through the cell membrane of *E. coli*.

3.3.2 MSA Uptake by *E. coli*

Overnight cultures of fusion protein were inoculated to an OD₆₀₀ of 0.1 in M9/Glu 4mM/ACA16 mM/Cm and induced with 0.2 mM IPTG when OD₆₀₀ reached 0.2 (approximately 2 hours). OD₆₀₀ and ¹H NMR spectra were taken at 0 minute, 30 minute, 2 hour, 6 hour, and 24-hour time points (Figure 3.9). We can see that ACA concentration decreases with time (Figure 3.10). Yet, no MSA is seen in the NMR spectra. This is due to the fact that MSA is immediately taken up by cells. However, natural ethanol production is seen. This is due to the spontaneous enzymatic decarboxylation of MSA to acetaldehyde. Acetaldehyde is then converted to ethanol by anaerobic fermentation native to *E. coli*. This further confirms the entrance of MSA into cells. Additionally, the noticeable increase in cell growth indicates that cells are able to use MSA formed by ACA for survival (Figure 3.11).

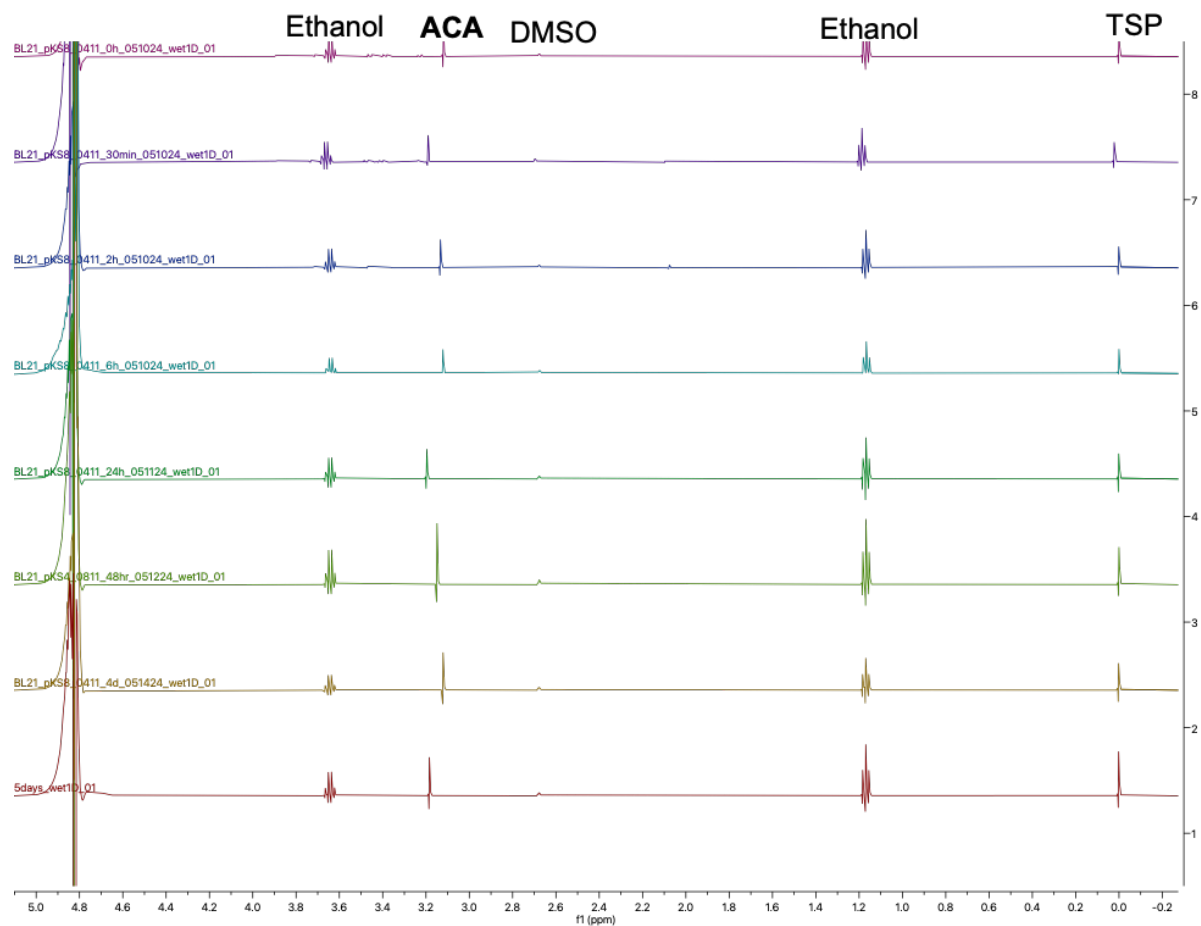


Figure 3.9. ^1H NMR spectra of ACA conversion in media. No MSA is shown implicating that MSA is readily taken up by the cell. Ethanol indicates acetaldehyde conversion via anaerobic fermentation.

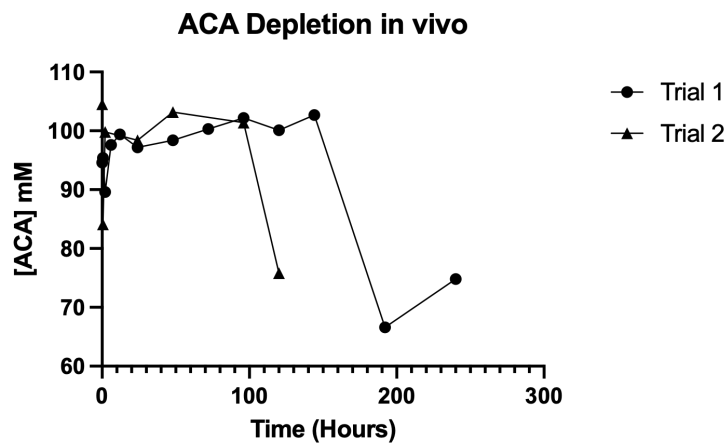


Figure 3.10. ACA depletion by cells with time.

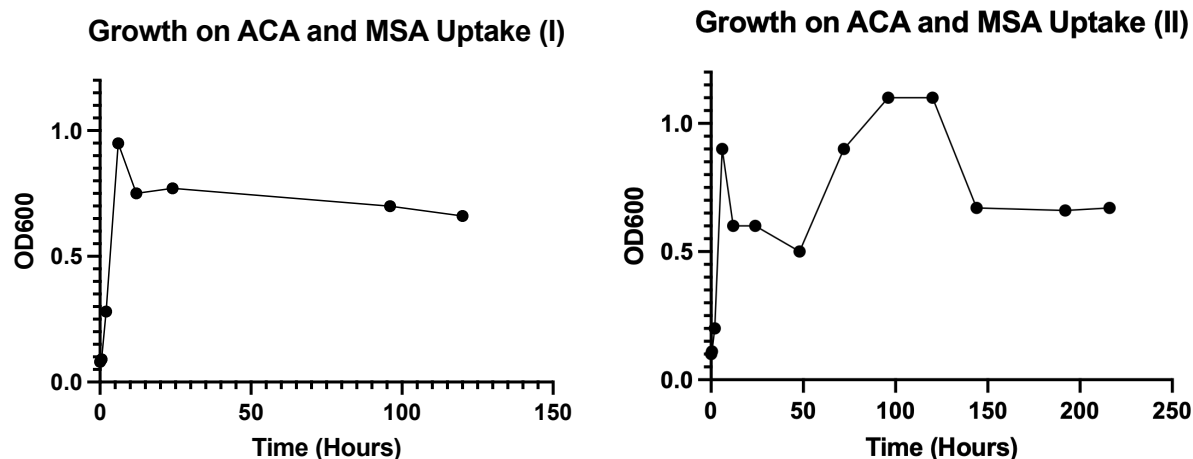


Figure 3.11. Increase in cell growth with time relative of ACA decline. (I) indicates trial 1 and (II) indicates trial 2.

3.3.3 Continual Production of High Priority Chemicals *In Vivo*

Here we see that Cg10062 (E114N) can be used as a fusion protein for *E. coli* growth on MSA. The transformation of a plasmid containing the gene for MmsB will allow the subsequent conversion of MSA to 3HP *in vivo*. MmsB may be transformed into *E. coli* cells containing the AIDA/Cg10062 (E114N) fusion protein to convert acquired MSA to 3HP. Additionally, genes related to anerobic fermentation should be knocked out to ensure direct transformation of MSA to 3HP. Alternately, the gene encoding the enzyme YneI may be inserted into *E. coli* cells containing the AIDA/Cg10062 (E114N) fusion to convert acquired MSA to malonic acid. Here we have evaded the limitation of *E. coli* nonviability on ACA carbon source and projected a novel route for the production of high value chemicals *in vivo* through the novel Cg10062 (E114N) enzyme using a unique AIDA/Cg10062 (E114N) fusion protein.

REFERENCES

1. Mathes Hewage, S., Silva, K., Kwiatkowski, K., Lee, H., Sreedhar, D., Gavgani, H. Geiger, J., Draths, K. An Original Biosynthetic Route to 3-Hydroxypropionic Acid, 2025, *in preparation*
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; National Renewable Energy Lab., Golden, CO (US), 2004. <https://doi.org/10.2172/15008859>.
3. 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. <https://doi.org/10.1039/B922014C>.
4. Johannes, T. W.; Woodyer, R. D.; Zhao, H. Efficient Regeneration of NADPH Using an Engineered Phosphite Dehydrogenase. *Biotechnol. Bioeng.* 2007, 96 (1), 18–26. <https://doi.org/10.1002/bit.21168>.
5. Lenz, O.; Lauterbach, L.; Frielingsdorf, S. Chapter Five - O₂-Tolerant [NiFe]-Hydrogenases of *Ralstonia eutropha* H16: Physiology, Molecular Biology, Purification, and Biochemical Analysis. In *Methods in Enzymology*; Armstrong, F., Ed.; Enzymes of Energy Technology; Academic Press, 2018, 613, 117–151.
6. Chowdhury, E.; Akaishi, Y.; Nagata, S.; Misono, H. Cloning and Overexpression of the 3-Hydroxyisobutyrate Dehydrogenase Gene from *Pseudomonas putida* E23. *Biosci. Biotechnol. Biochem.* 2003.
7. 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.
8. Benz, I.; van Alen, T.; Bolte, J.; Wörmann, M. E.; Schmidt, M. A. Modulation of Transcription and Characterization of the Promoter Organization of the Autotransporter Adhesin Heptosyltransferase and the Autotransporter Adhesin AIDA-I. *Microbiology.* 2010, 156 (4), 1155–1166.
9. Jarmander, J.; Gustavsson, M.; Do, T.-H.; Samuelson, P.; Larsson, G. A Dual Tag System for Facilitated Detection of Surface Expressed Proteins in *Escherichia coli*. *Microb. Cell Fact.* 2012, 11, 118.
10. Crane, J. M.; Randall, L. L. The Sec System: Protein Export in *Escherichia coli*. *EcoSal Plus.* 2017, 7 (2). <https://doi.org/10.1128/ecosalplus.ESP-0002-2017>.
11. Paetzel, M. Structure and Mechanism of *Escherichia coli* Type I Signal Peptidase. *Biochim. Biophys. Acta.* 2014, 1843 (8), 1497–1508.
12. Auclair, S. M.; Bhanu, M. K.; Kendall, D. A. Signal Peptidase I: Cleaving the Way to Mature Proteins. *Protein Sci.* 2012, 21 (1), 13–25.

13. Gustavsson, M.; Bäcklund, E.; Larsson, G. Optimisation of Surface Expression Using the AIDA Autotransporter. *Microb. Cell Fact.* 2011, *10*, 72.
14. Parks, L.; Ek, M.; Ståhl, S.; Löfblom, J. Investigation of an AIDA-I Based Expression System for Display of Various Affinity Proteins on *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 2024, *696*, 149534.

CHAPTER FOUR: Experimental

4.1 Materials

Chemicals, biochemicals, Luria-Bertani (LB) media components and buffer salts were purchased from Sigma Aldrich (St. Louis, MO), Becton, Dickinson and Company (Sparks, MD), Fisher Scientific (Pittsburgh, PA) and Gold Biotechnology (St. Louis, MO). Alcohol dehydrogenase from *Saccharomyces cerevisiae* was purchased from Sigma Aldrich. Protein Assay Dye Reagent, Precision Plus Protein standards and Mini-PROTEAN TGX Precast 4-20% polyacrylamide gels were purchased from Bio-Rad (Hercules, CA). Q5 Site-Directed Mutagenesis Kits, Monarch PCR and DNA Cleanup Kit and restriction enzymes were purchased from New England Biolabs (Ipswich, MA). QIAprep Spin Miniprep and Maxiprep Kits were purchased from Qiagen (Venlo, Netherlands). HisTrap FF 1 mL and 5 mL pre-packed columns were purchased from Cytiva (Marlborough, MA). Amicon Ultra-15 10K centrifugal filter units and 0.22 μ m 0.45 μ m syringe filters were purchased from MilliporeSigma (Burlington, MA). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Commercially synthesized plasmids were obtained from Genscript (Piscataway, NJ). PEG/Ion and PEG/Ion 2 crystallization screens were purchased from Hampton Research (Aliso Viejo, CA).

Culture media were prepared in distilled, deionized water. Culture media including LB, SOB and SOC were prepared according to standard protocols.¹ Ampicillin and isopropyl β -D-1-thiogalactopyranoside (IPTG) stock solutions were prepared using sterile deionized water and filtered through 0.22 μ m syringe filters. Restriction digests were carried out to confirm the insertion of mutations of the plasmids after site-directed mutagenesis, prior to sequencing (Table 4.1). The general components used for a double digest are shown below in Table 6.1. Reactions were prepared in 0.2 mL microfuge tubes and incubated at 37 °C for 1 hour prior to separation on a 0.7 % agarose gel.

4.2 Culture Media and Stock Solutions

All culture media were prepared according to standard protocols in double deionized water and sterilized by autoclaving (liquid cycle, 121 °C, 25 min).¹ LB medium (1 L) was prepared with Bacto™ tryptone (10 g), Difco™ yeast extract (5 g) and NaCl (10 g) and was sterilized by autoclaving.² Nutrient Broth (NB) medium (1 L) contained beef extract (3 g) and peptone (5 g) and was sterilized by autoclaving. 2XYeast-tryptone (2XYT) media (1 L) contained Bacto™ tryptone (16 g), Difco™ yeast extract (10 g) and NaCl (5 g) and sterilized. SOB medium (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. SOC medium (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).

Stock solutions of 2 M MgCl₂, 1 M MgSO₄, 1 M glucose, and 20% (w/v) glucose were prepared separately using deionized water and sterilized by autoclaving. Stock solutions of 1 mg mL⁻¹ thiamine hydrochloride and 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) stock solutions were prepared using deionized water and sterilized using 0.22 μm syringe filters. Antibiotics were added to all media where appropriate to the following final concentrations: ampicillin (Ap, 50 μg mL⁻¹), chloramphenicol (Cm, and 20 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).

4.3 Preparation and Transformation of Electro- and Chemically Competent *E. coli*

Electrocompetent cells were prepared using standard operating procedures.¹ A desired strain was streaked out on an LB agar plate. A single bacterial colony was selected to inoculate a 5 mL LB culture and incubated overnight with shaking (37 °C, 200 rpm). The overnight culture (2 mL) was used to inoculate 100 mL of sterile 2xYT media in a 500 mL baffled shake flask and

incubated at 37 °C until an OD₆₀₀ of 0.5-0.7 was reached. The cells were harvested by centrifugation using a Fiberlite™ F12-6x500LEX fixed angle rotor (4,500 x g, 5 min, 4 °C) and gently resuspended in 100 mL cold, sterile deionized water to remove residual salts from the culture media. The cells were pelleted by centrifugation (4,500 x g, 5 min, 4 °C) and carefully decanted to remove the supernatant. The cells were again washed with another 100 mL cold, sterile deionized water. The cell pellet was resuspended in 100 mL cold, sterile 10% aqueous (v/v) glycerol 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 10% (v/v) glycerol. 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.

The transformation of electrocompetent *E. coli* cells is as follows. Plasmid DNA (2 µL of 1-5 ng µL⁻¹ in sterile deionized water), purified, de-salted PCR products, or ligation products were transferred to pre-aliquoted 50 µL of electrocompetent cells thawed on ice. 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 to the cuvette. The resulting recovery culture was transferred to a sterile 5 mL culture tube and 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 60 µL aliquots were plated onto LB solid media containing the appropriate antibiotic.

A standard protocol adapted from Sambrook and Russell was followed to prepare chemical competent cells.¹ Using a freshly streaked plate of less than 5 days old, a single bacterial colony was inoculated into 5 mL LB and incubated overnight with shaking (37 °C, 200 rpm). A 1 mL aliquot of the overnight culture was used to inoculate 100 mL of sterile LB media in a 500 mL baffled shake flask and incubated at 37 °C until an OD₆₀₀ of 0.5-0.7 was reached. The cells were

transferred to a sterile 500 mL centrifuge bottle, incubated on ice for 5 minutes, collected by centrifugation (2,700 x g, 5 min, 4 °C), and gently resuspended were resuspended by gently rocking the centrifuge bottle on ice in 100 mL 0.9% NaCl. The cells were pelleted by centrifugation (2,700 x g, 5 min, 4 °C) and the sample was carefully decanted to remove the supernatant. While on ice, the cells were resuspended by gently rocking the centrifuge bottle in 50 mL of 100 mM CaCl₂, left on ice for 30 minutes to 1 hour, and centrifuged (2,700 x g, 5 min, 4 °C). The supernatant was discarded, and the cell pellet was resuspended in 4 mL of 100 mM CaCl₂ in 15% (w/v) glycerol solution. Aliquots (100 µL) were prepared in sterile microcentrifuge tubes on ice and flash-frozen in liquid nitrogen for storage at -80 °C.

For the transformation of chemical competent *E. coli* cells, plasmid DNA (1 µL of 10-100 ng µL⁻¹ in sterile deionized water), purified, de-salted ligation product, or PCR products were combined with 100 µL of chemical competent cells thawed on ice. The tube was gently tapped to mix, and the sample was placed on ice for 30 minutes. The cells were heat-shocked using a 42 °C (± 2 °C) water bath for exactly 60 seconds then immediately placed on ice for 2 minutes. Following, 0.5 mL pre-warmed LB media was added. The recovery culture was incubated at 37 °C for 1 hour without shaking. The culture was centrifuged at 5,000 rpm in a microcentrifuge to pellet cells, and 500 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. The plates were incubated at 37 °C overnight.

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 5 mL LB media containing the appropriate antibiotic and incubated at 37 °C overnight with shaking (200 rpm). Cells were harvested for 1 min at 13,000 rpm using a table top microcentrifuge. For large scale purification, a single colony from a freshly transformed plate of *E. coli* was inoculated into 100 mL LB media containing the appropriate antibiotic and incubated at 37 °C overnight with shaking (200 rpm). Cells were harvested by

centrifugation (6,100 x g, 10 min, 4 °C). Isolation and purification of the plasmid DNA was carried out using the Qiagen Plasmid Miniprep and Midiprep Kit's following the manufacturer's instructions. For the small scale Miniprep purification, DNA was eluted from the column with 50 μ L sterile, deionized water. For the large scale Midiprep purification of DNA, the air-dried plasmid pellet was dissolved in 100–200 μ L sterile, deionized water. All purified plasmids were stored at 4 °C.

4.5 Restriction Digestion of DNA

Restriction digestion was used to confirm the identity of purified plasmid DNA and also to isolate DNA fragments for T4 ligation. A typical digest had a final volume of 20 μ L and contained the components listed in Table 4.1. For a single digest, only one restrict enzyme was used. For a double digest, 1 μ L of each required restriction enzyme was used. If a larger amount of digested plasmid DNA was required for cloning purposes or DNA assembly reaction, the volume was increased to 50 μ L and all components were scaled up accordingly. To achieve a final volume of 20 μ L , the volume of deionized water was adjusted accordingly. Digestion reactions were incubated for 1 h at 37 °C and subsequently quenched with 4 μ L 6X loading dye containing SDS (NEB) prior to visualization on a 0.7% agarose gel containing ethidium bromide (0.5 μ g 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 typical restriction digest.

Component	Volume	Final Quantity
Plasmid (100 ng μ L ⁻¹)	5 μ L	500 ng
10X Buffer	2 μ L	1X
Restriction enzyme 1	1 μ L	10 U
Restriction enzyme 2	1 μ L	10 U
Deionized water	11 μ L	-

DNA was visualized using agarose gel electrophoresis. TAE buffer (1X) (0.5% tris(hydroxymethyl)aminomethane, 0.1% acetic acid, and 0.03% ethylenediaminetetraacetic acid) in water (pH 8) was used as the running buffer. To make one 2.5" x 4" gel, 50 mL 1x TAE buffer

and 0.35 g UltraPure agarose were combined in an Erlenmeyer flask and microwaved until the mixture came to a boil (approximately 30 s). Once the agar had cooled slightly, ethidium bromide (2.5 μL) was added. After cooling to approximately 60 °C, the solution was poured into a gel mold with appropriate lane markers and left to solidify. The gel was then placed in an electrophoresis gel box and submerged in 1x TAE running buffer. A DNA size marker stock was prepared with 20 μL of 1 kb size marker (0.5 $\mu\text{g } \mu\text{L}^{-1}$), 20 μL purple loading dye with no SDS, and 80 μL sterile deionized H_2O . Samples (15-25 μL) and DNA size marker (1 kb, 8-10 μL , 0.7-0.8 μg) were loaded into the wells by carefully pipetting. When running a DNA sample from a PCR, 5 μL of PCR product was combined with 10 μL water and 3 μL NEB 6X SDS Purple Loading Dye. The gel was run at 98 V until the loading dye was 1-2 inches from the bottom of the gel (approximately 1 hr) and was imaged using an Axygen imaging system.

4.6 DNA Sequencing

All plasmids constructed and used in this study were submitted for Sanger sequencing for sequencing confirmation prior to their use for *in vitro* and *in vivo* studies. Isolated DNA was sequenced at the Michigan State University Research Technology Support Facility (MSU RTSF) Genomics Core using custom primers or commercial T7 forward and T7 reverse primers. For each plasmid sequenced, forward and reverse primers were designed with sufficient distance between each set to obtain full sequence data from both strands of DNA. Sequencing data received from MSU RTSF were analyzed using the “Align to Reference Sequence” tool on SnapGene 6.0.

4.7 PCR Amplification

Unless otherwise specified, all PCR amplifications were conducted in a Bio-Rad DNA Engine® Peltier Thermal Cycler with a final reaction volume of 50 μL using Q5® High-Fidelity DNA Polymerase with other components as listed in Table 4.2. Routine PCR was conducted using the thermocycling conditions in Table 4.3. The PCR samples were quenched with 10 μL of 6X loading dye containing SDS (NEB) prior to visualization on a 0.7% agarose gel.

Table 4.2. PCR protocol of parent template.

Component	Volume	Final Concentration
Q5 Hot State High-Fidelity Master Mix (2X)	12.5 μ L	1X
Forward Primer (10 μ M)	1.25 μ L	0.5 μ M
Reverse Primer (10 μ M)	1.25 μ L	0.5 μ M
Template DNA (5 ng μ L ⁻¹)	1 μ L	0.2 ng μ L ⁻¹
Deionized water	9 μ L	-

Table 4.3. PCR conditions for Q5 site-directed mutagenesis.^a

Step	Temperature	Time
Initial Denaturation	98 °C	30 sec
30 cycles	98 °C	10 sec
	50–72 °C ^a	30 sec
	72 °C	3 min (30 sec kb ⁻¹)
Final Extension	72 °C	2 min
Hold	4 °C	-

^aannealing temperatures for each set of primers were determined using NEBasechanger.

4.8 Protein Expression

Protein expression was carried out in *E. coli* BL21(DE3) relying on a T7 promoter system, using either pET-21a(+) or pET-15b vectors. A plasmid encoding a His₆-tagged protein was transformed into *E. coli* BL21(DE3) and a single colony was inoculated into 50 mL of LB media containing ampicillin. Cells were grown at 37 °C with shaking (200 rpm) overnight.³ The overnight culture was used to inoculate either 100 mL or 1 L LB media containing antibiotic depending on small scale or large scale expression, respectively. Cultures were incubated at 37 °C for approximately 1.5 h with shaking (200 rpm) until an OD₆₀₀ of 0.5-0.7 was reached at which point IPTG (isopropyl β -D-1-thiogalactopyranoside) to a final concentration of 1 mM was added. IPTG-induced overexpression was carried out for 8–10 h at 30 °C (200 rpm). Cells were harvested by centrifugation using a Fiberlite™ F12-6x500LEX fixed angle rotor (4,500 x g, 10 min, 4 °C) and stored at -20 °C until purification.

Uninduced and induced cell lysates were analyzed on 4–20% (w/v) acrylamide Mini-PROTEAN® TGX™ precast gels from Bio-Rad (Figure 4.1). A 500 μ L sample of uninduced and

induced cell culture was centrifuged in a tabletop microcentrifuge at 14000 rpm for 5 min and supernatant was discarded and cell pellets were weighed. Sodium phosphate buffer 100 mM pH8 was added to each cell pellet at a ratio of 3 mL of buffer per g cell paste. Lysozyme (MilliporeSigma, 5 mg mL⁻¹, 50 µL) was added to each 200 µL sample resuspension. Samples were incubated at room temperature for 20 min shaking (200 rpm) and then centrifuged in a tabletop microcentrifuge at 14,000 rpm for 15 min to remove lysate debris. Each protein sample was quantified as described (see 4.10 Protein Quantification) using Bradford reagent and stock solutions (300 ng µL⁻¹) were prepared in sodium phosphate buffer. To 2X Laemmli Buffer (800 µL) containing 4% (w/v) SDS, 20% (w/v) glycerol, 0.004% (w/v) bromophenol blue and 125 mM Tris-HCl (pH 6.8), 200 µL of 2 M DTT (dithiothreitol) was added to obtain a final concentration of 400 mM DTT. The 2X Laemmli Buffer (50 µL) containing DTT was combined with protein sample (50 µL) and incubated at 98 °C for 5 min. The samples were centrifuged to remove any precipitate (17,000 x g, 2 min). Protein samples (30 µL) and Precision Plus Protein™ All Blue Ladder (10 µL) (Bio-Rad) were loaded onto the precast gel. SDS-PAGE was carried out under denaturing conditions in running buffer (pH 8.3) containing 0.1% (w/v) SDS, 192 mM L-glycine and 25 mM Tris base, at a constant voltage of 200 V. After the dye reached approximately one inch from the bottom of the gel (35-40 min) the gel was removed from the cassette and stained for 1.5–2 h in a solution containing 0.1% (w/v) Coomassie Brilliant Blue R, 45% (v/v) methanol and 10% (v/v) acetic acid in water. The gel was promptly destained for overnight in a solution containing 45% (v/v) methanol and 10% (v/v) acetic acid in water.

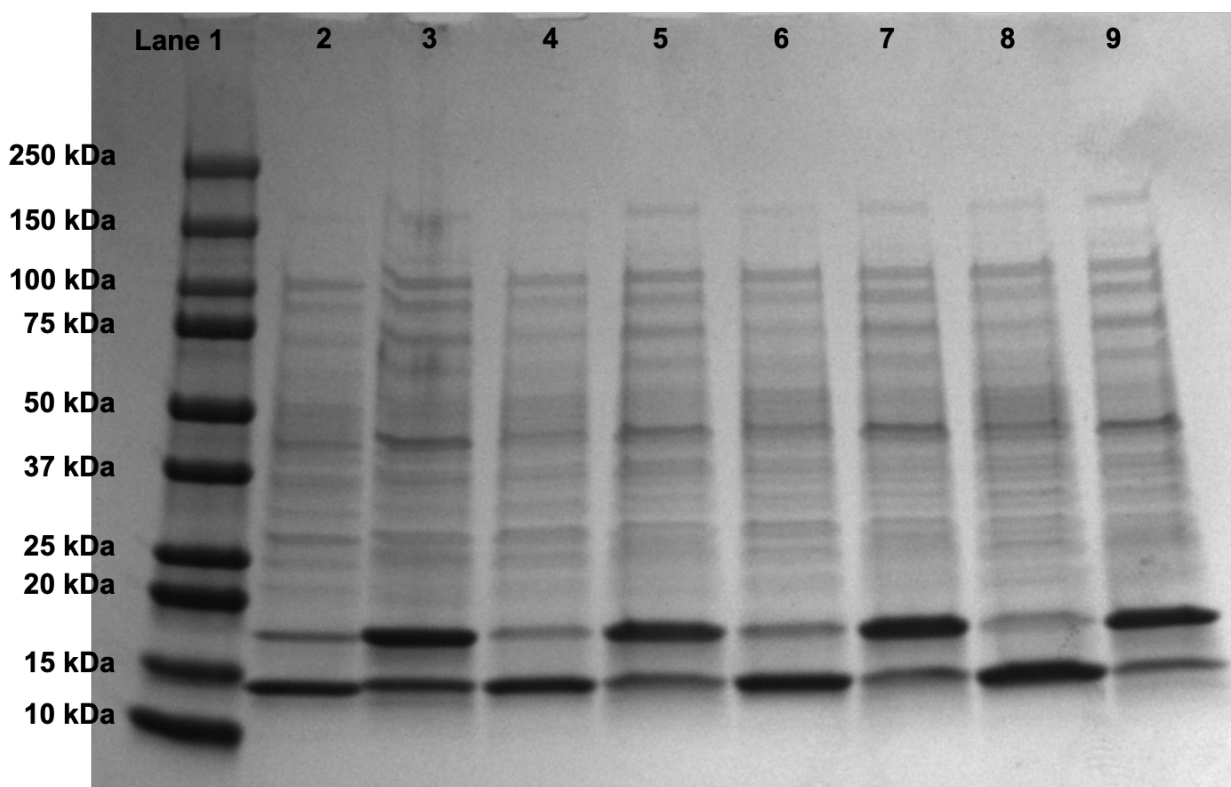


Figure 4.1. SDS-PAGE of *cis*-CaaD (18.5 kDa), *cis*-CaaD (E114D) (18.5 kDa), *cis*-CaaD (E114N) (18.5 kDa), and *cis*-CaaD (E114Q) (18.5 kDa). Lane 1: protein ladder, lane 2: uninduced culture of *cis*-CaaD (wild-type), lane 3: induced culture of *cis*-CaaD (wild-type), lane 4: uninduced culture of *cis*-CaaD (E114D), lane 5: induced culture of *cis*-CaaD (E114D), lane 6: uninduced culture of *cis*-CaaD (E114N), lane 7: induced culture of *cis*-CaaD (E114N), lane 8: uninduced culture of *cis*-CaaD (E114Q), and lane 9: induced culture of *cis*-CaaD (E114Q). (Other variants not shown.) The band at 14 kDa depicts lysozyme used in cell lysis during SDS-PAGE preparation.

4.9 Protein Purification

All proteins characterized and used in this dissertation were purified by means of a His₆-tag either on the C-terminus or N-terminus 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 previously stored at -20 °C were thawed at room temperature. The 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. 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 into a chilled centrifuge tube. Lysed cells were then centrifuged using a Fiberlite™

F21S-8x50y fixed angle rotor (47,500 x g) at 4 °C for 30 min. Clarified lysate was separated from 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 (Figure 4.2). Two buffers were used for the purification process. The binding buffer contained 20 mM sodium phosphate, pH 7.4 and 500 mM NaCl. The elution buffer contained 20 mM sodium phosphate pH 7.4, 500 mM sodium chloride and 500 mM imidazole. In each step (except elution and fractionation), buffer mixing was used (96% binding buffer and 4% elution buffer) to ensure that the imidazole concentration (20 mM) was identical to that of the sample that was resuspended in lysis buffer. The HisTrap FF column was equilibrated with 5 column volumes at a flow rate of 4.0 mL min⁻¹. Sample application was carried out using a flow rate of 1.0 mL min⁻¹ and unbound protein was washed out using 15 column volumes at a flow rate of 4.0 mL min⁻¹. An imidazole gradient from 20 mM (4% elution buffer) to 500 mM (100% elution buffer) over 20 column volumes was used to elute the His₆-tagged protein from the nickel column. Fractions containing the desired protein were pooled and divided evenly into two separate tubes. Each tube was individually concentrated and desalted using Amicon Ultra-15 10K filters. One sample were stored in 100 mM sodium phosphate pH 8.0 while the other was stored in 10 mM Tris-SO₄ pH 8.0. Protein purifications typically yielded between 10-150 mg of enzyme per liter of cell culture.

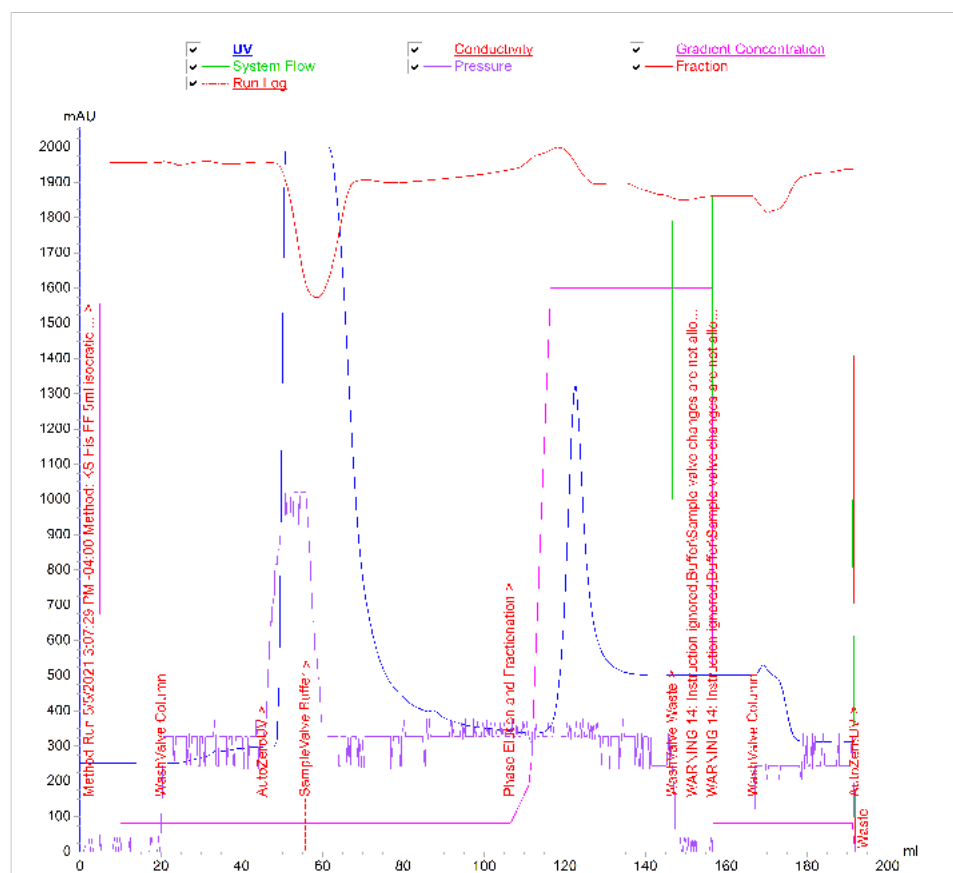


Figure 4.2. FPLC purification of wild-type *cis*-CaaD. Other variants (not shown) were purified similarly.

4.10 Protein Quantification

Protein concentrations of cell lysates were quantified using Bio-Rad Protein Assay reagent. For quantification of cell lysates, crude lysate was diluted 5-fold by addition of deionized water. Following addition of 1 mL of protein assay reagent to 4 μ L diluted protein with 16 μ L of deionized water, samples were incubated at room temperature for 5 min before measuring OD₅₉₅. The enzyme concentration was determined using a standard curve developed for the Bradford reagent using bovine serum albumin standard. Purified proteins were quantified using 6 M guanidinium chloride.³ Purified protein (10 μ L) was combined with 6 M guanidinium chloride (990 μ L) and the absorbance at 280 nm was measured. Purified protein solution was quantified using

the molar extinction coefficient of each protein at 280 nm and their respective molecular weights (Table 4.8).⁴

Table 4.4. Parameters used for enzyme quantification.⁵

Enzyme	Molecular Weight (Da)	ϵ (280 nm) ($M^{-1} cm^{-1}$)
<i>cis</i> -CaaD (wild-type)	18542	17780
<i>cis</i> -CaaD (E114D)	18528	17780
<i>cis</i> -CaaD (E114N)	18527	17780
<i>cis</i> -CaaD (E114Q)	18541	17780
<i>cis</i> -CaaD (Y103F)	18526	16500
<i>cis</i> -CaaD (Y103A)	18449	16500
<i>cis</i> -CaaD (H28A)	18475	17780
<i>cis</i> -CaaD (R73A)	18456	17780
<i>cis</i> -CaaD (T32A)	18512	17780
<i>cis</i> -CaaD (T34A)	18512	17780
Cg10062 (wild-type)	19013	30440
Cg10062 (E114N)	18998	30440
Cg10062 (E114Q)	19012	30440
Cg10062 (E114D)	18999	30440
Cg10062 (Y103F)	18997	29160
<i>C-cis</i> -CaaD/ <i>N</i> -Cg10062	18470	17780
<i>N</i> -Cg10062/ <i>C-cis</i> -CaaD	19085	30440
<i>tcis</i> -CaaD	15108	12090
MSAD	16464	8250

4.11 Short-Path Distillation of Acetylenecarboxylic Acid

cis-3-Chloroacrylic acid (CCA) purchased as a solid from MilliporeSigma was used directly. However, it was necessary to purify acetylenecarboxylic acid 95% (ACA) purchased from MilliporeSigma using short-path distillation. All glassware were rinsed with acetone and dried at 100 °C prior to use. ACA was transferred to a pre-weighed round-bottom flask. The short-path distillation apparatus was set up using a short-path distillation head connected to a cow-type distilling receiver which was connected to three round-bottom flasks for collection of distillates. Glass wool was used for insulation, all glass joints were greased to ensure a tight seal, a vacuum pump was connected, and a column containing DRIERITE[®] desiccant was inserted within the apparatus. The round-bottom flask containing the ACA was gently heated using a hot water bath with stirring. 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. Distillate temperature was monitored using an overhead thermometer inside the short-path condenser and purified ACA was collected in the second round-bottom flask connected to the cow-type distilling receiver at 60-65 °C and approximately 12 psi vacuum pressure. The distilled ACA was transferred to glass vials in 5 mL aliquots under nitrogen and stored at -20 °C. Exposure to air was minimized thereafter.

4.12 Chapter Two: Comparison of *cis*-CaaD and Cg10062 Transformation of Acetylenecarboxylic Acid

4.12.1 Bacterial Strains, Genes, and Plasmids

Escherichia coli strains BL21(DE3), DH5 α , and MG1655 were obtained from Invitrogen (Carlsbad, CA). Plasmids pET21-a(+) and pET-28a(+) were obtained from Invitrogen. The plasmid construct expressing a His₇-tagged TEV protease (pMHT Δ 238) and *E. coli* BL21-CodonPlus(DE3)-RP were kindly provided by Professor Heedeok Hong of Michigan State University.⁶ Cells were grown at 37 °C in LB media containing ampicillin (Ap, 50 μ g mL). The first amino acid in each sequence is methionine (Met-0) which is subsequently cleaved by native *E. coli* methionyl-aminopeptidase (MAP). Thus, numbering of amino acids residues begins with Pro-1. A list of strains used in chapter two is shown in table 4.5.

Table 4.5. Strains and plasmids used in Chapter Two of this study.

Strain or Plasmid	Description	Reference/ Source
<i>E. coli</i> DH5 α	K-12 B F ⁻ ϕ 80/ <i>lacZ</i> Δ M15 Δ (<i>lacZ</i> YA- <i>argF</i>)U169 <i>recA1 endA1 hsdR17</i> (r _K ⁻ , m _K ⁺) <i>phoA supE44</i> λ - <i>thi-1 gyrA96 relA1</i>	Invitrogen
<i>E. coli</i> BL21	K-12 B <i>fhuA2 [lon] ompT gal [dcm] ΔhsdS</i>	Invitrogen
<i>E. coli</i> BL21(DE3)	K-12 B F ⁻ <i>ompT hsdS_B</i> (r _B ⁻ , m _B ⁻) <i>gal dcm</i> (DE3)	Invitrogen
<i>E. coli</i> BL21-CodonPlus(DE3)-RP	F ⁻ <i>ompT hsdSx</i> (r _B ⁻ m _B ⁻) <i>dcm</i> ⁺ Tet ^R <i>gal</i> λ (DE3) <i>endA Hte [argU proL Cam^R]</i>	Agilent Technologies
<i>E. coli</i> MG1655	F ⁻ λ - <i>ilvG⁻ rfb-50 rph-1</i>	Invitrogen
pMHT Δ 238	Kan ^R , TEV protease expression vector	DNASU Plasmid Repository ⁷
pET-21a(+)	Ap ^R , <i>lacI</i> , P _{T7} pMB1 replicon	Invitrogen

Table 4.5. (cont'd)

pNJD1.002	<i>P_{TT}cis-CaaD</i> in pET-21a(+)	Devito
pKS1.001	<i>P_{TT}cis-CaaD</i> (E114D) in pNJD1.002	This Study
pKS1.002	<i>P_{TT}cis-CaaD</i> (E114N) in pNJD1.002	This Study
pKS1.003	<i>P_{TT}cis-CaaD</i> (E114Q) in pNJD1.002	This Study
pKS1.004	<i>P_{TT}cis-CaaD</i> (Y103A) in pNJD1.002	This Study
pKS1.005	<i>P_{TT}cis-CaaD</i> (Y103F) in pNJD1.002	This Study
pKS1.006	<i>P_{TT}cis-CaaD</i> (R73A) in pNJD1.002	This Study
pKS1.007	<i>P_{TT}cis-CaaD</i> (T32A) in pNJD1.002	This Study
pKS1.008	<i>P_{TT}cis-CaaD</i> (T34A) in pNJD1.002	This Study
pKS1.009	<i>P_{TT}cis-CaaD</i> (H28A) in pNJD1.002	This Study
pKS2.0503	<i>P_{TT}cis-CaaD/Cg10062</i> in pET-21a(+)	This Study
pKS2.0504	<i>P_{TT}Cg10062/cis-CaaD</i> in pET-21a(+)	This Study
pKS2.0615	<i>P_{TT}tcis-CaaD</i> in pET-21a(+)	This Study
pAS2.031	<i>P_{TT}msad</i> in pET-21a(+)	Sirinimal
pAS1.046	<i>P_{TT}cg10062</i> in pET-21a(+)	Sirinimal
pAS2.100	<i>P_{TT}Cg10062</i> (E114N) in pET-21a(+)	Sirinimal
pKK1.1025	<i>P_{TT}mmsB</i> in pET-21a(+)	Kwiatkowski
<i>Corynebacterium glutamicum</i>	NCBI 10025	ATCC 13032

Plasmid pNJD1.002

The gene expressing *cis*-CaaD (PDB ID: 3MF8) from *Corynebacterium* was codon-optimized for expression in *E. coli* and modified to replace the stop codon with a TEV protease recognition site (Figure 4.3 and Figure 4.4).⁶ The modified gene was cloned into the pET-21a(+) commercial vector at the NdeI and XhoI restriction sites, respectively, at the 5' and 3' positions to afford pNJD1.002. All variants of *cis*-CaaD designed using Q5 site-directed mutagenesis were designed using pNJD1.002 as a template.

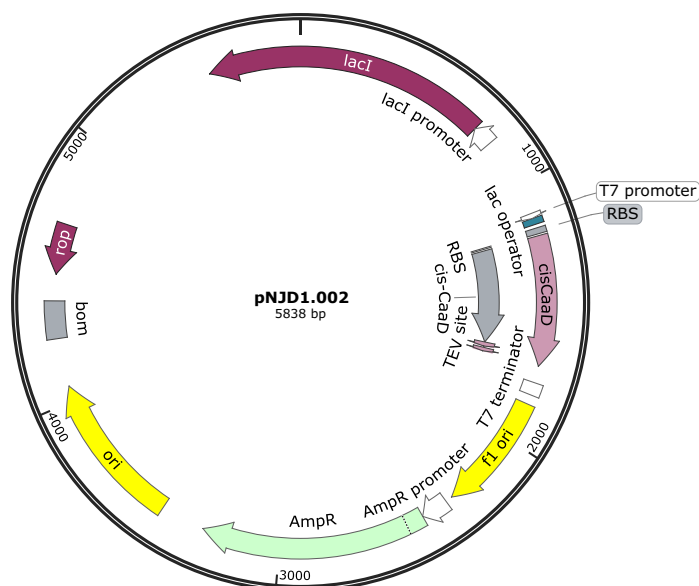


Figure 4.3. Plasmid map of pNJD1.002.

5' - ATGCCGGTTTACATGGTTTACGTTAGCCAGGACCGTCTGACCCCGAGCGCGAAGCACG
 CGGTTGCGAAGGCGATTACCGATGCGCACCGTGGTCTGACCGGCACCCAGCACTTCCTGG
 CGCAAGTGAACCTTCAGGAGCAACCGGCGGGTAACGTGTTCCCTGGGTGGCGTTCAGCAAG
 GTGGCGACACCATCTTTGTTTCATGGTCTGCACCGTGAGGGCCGTAGCGCGGATCTGAAGG
 GCCAGCTGGCGCAACGTATTGTTGACGATGTGAGCGTTGCGGCGGAAATCGACCGTAAAC
 ACATTTGGGTGTACTTCGGCGAGATGCCGGCGCAGCAAATGGTTGAATATGGCCGTTTCCT
 GCCGCAGCCGGGTCATGAGGGTGAATGGTTTGACAACCTGAGCAGCGATGAACGTGCGTT
 TATGGAGACCAATGTTGATGTGAGCCGTACC **GAAATCTGTACTTCCAAGGT** - 3'

Figure 4.4. Codon-optimized nucleotide sequence for expression of *cis*-CaaD in *E. coli*. Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.

Plasmid pKS2.0503

The plasmid pNJD1.002 was used as a template to construct the chimera and truncated variants of *cis*-CaaD (Figure 4.5). A gene chimera consisting of amino acids 1-119 of *cis*-CaaD followed by amino acids 120-148 of Cg10062 was obtained from Integrated DNA Technologies (Figure 4.6). The gene included NdeI and XhoI restriction sites, respectively, at the 5' and 3' positions. Each gene was cloned into pET-21a(+) between the NdeI and XhoI restriction sites. Restriction digestion followed by T4 ligation was used to clone the chimera genes into pET-21a(+).

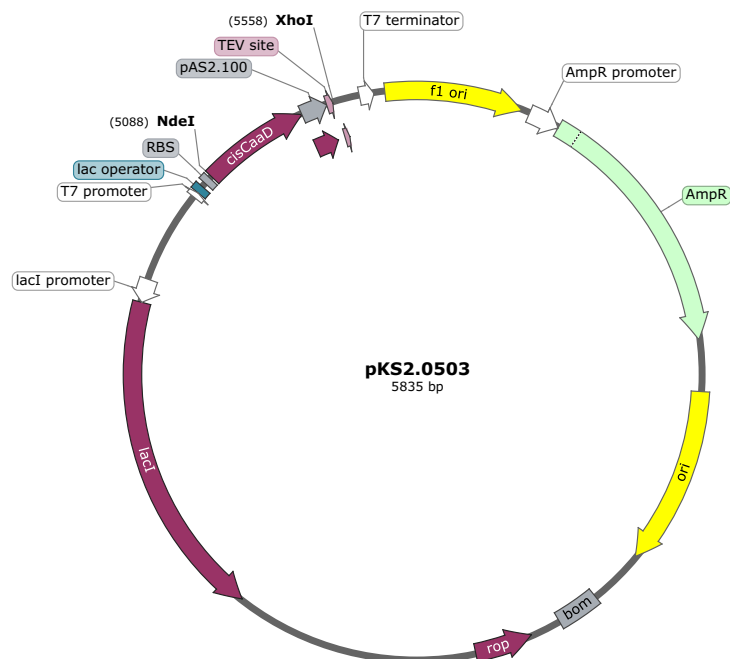


Figure 4.5. Plasmid map of pKS2.0503.

5' – ATGCCGGTTTACATGGTTTACGTTAGCCAGGACCGTCTGACCCCGAGCGCGAAGCACG
 CGGTTGCGAAGGCGATTACCGATGCGCACCGTGGTCTGACCGGCACCCAGCACTTCCTGG
 CGCAAGTGAAC TTTCAGGAGCAACCGGCGGGTAACGTGTTCTTGGGTGGCGTTT CAGCAAG
 GTGGCGACACCATCTTTGTTTCATGGTCTGCACCGTGAGGGCCGTAGCGCGGATCTGAAGG
 GCCAGCTGGCGCAACGTATTGTTGACGATGTGAGCGTTGCGGCGGAAATCGACCGTAAAC
 ACATTTGGGTGTACTTCGGCGAGATGCCGGCGCAGCAAATGGTTGAATATGGCCGTTTCCT
 GATGGAGCCGGGCGAGGAAGAGAAATGTTCAACAGCCTGCCGGAGGGCCTGCGTGAGC
 GTCTGACCGAACTGGAGGGTAGCAGCGAA **GAAAATCTGTACTTCCAAGGT** – 3'

Figure 4.6. Codon-optimized nucleotide sequence for expression of *N-cis-CaaD/C-Cg10062* in *E. coli*. Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.

Plasmid pKS2.0504

A gene chimera consisting of amino acids 1-119 of Cg10062 followed by amino acids 120-149 of *cis*-CaaD was obtained from Integrated DNA Technologies (Figure 4.7 and Figure 4.8). The gene included NdeI and XhoI restriction sites, respectively, at the 5' and 3' positions. Each gene was cloned into pET-21a(+) between the NdeI and XhoI restriction sites. Restriction digestion followed by T4 ligation was used to clone the chimera genes into pET-21a(+).

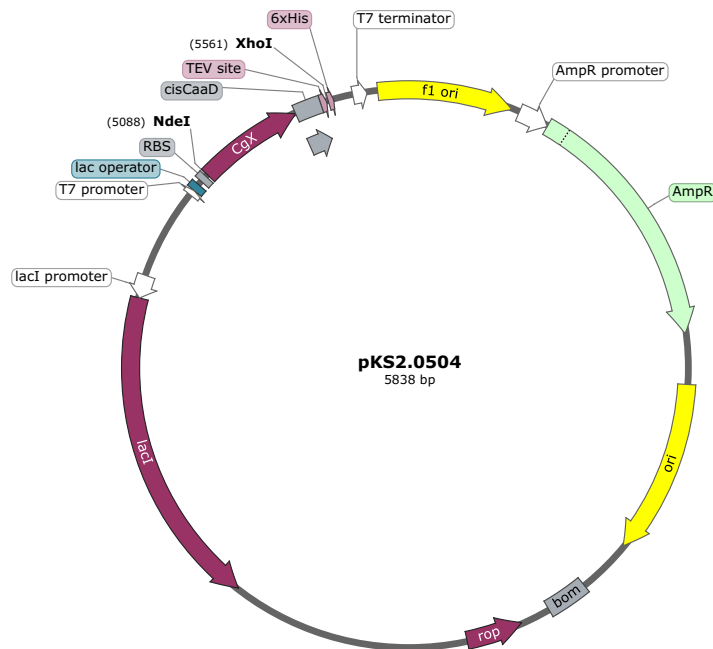


Figure 4.7. Plasmid map of pKS2.0504.

5' – ATGCCGACCTACACCTGCTGGAGCCAACGCATTCGTATTAGCCGTGAAGCGAAGCAAC
GCATCGCGGAAGCGATTACCGACGCGCACCATGAACTGGCGCACGCGCCGAAGTACCTGG
TGCAGGTTATTTTCAACGAAGTGGAGCCGGACAGCTATTTTATCGCGGCGCAGAGCGCGAG
CGAGAACCACATTTGGGTTCAAGCGACCATCCGTAGCGGCCGTACCGAAAAGCAGAAAGA
GGAAGTCTGCTGCTGCGTCTGACCCAAGAGATCGCGCTGATTCTGGGTATCCCGAACGAGGA
AGTGTGGGTTTACATTACCGAAATCCCGGGTAGCAACATGACCGAATATGGCCGTCTGCTG
CCGCAGCCGGGTCATGAGGGTGAATGGTTTGACAACCTGAGCAGCGATGAACGTGCGTTT
ATGGAGACCAATGTTGATGTGAGCCGTACC **GAGAACCTGTATTTTCAAGGC** – 3'

Figure 4.8. Codon-optimized nucleotide sequence for expression of *N-Cg10062/C-cis-CaaD* in *E. coli*. Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.

Plasmid pKS2.0615

A truncated gene consisting of amino acids 1-119 was designed and called *tcis-CaaD* (Figure 4.9 and Figure 4.10). HiFi Gibson assembly was used to truncate *cis-CaaD* at residue 119.

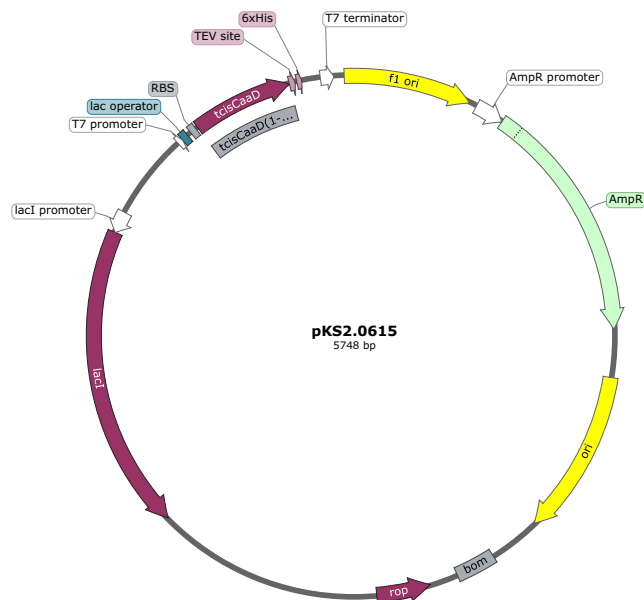


Figure 4.9. Plasmid map of pKS2.0615.

5' – ATGCCGGTTTACATGGTTTACGTTAGCCAGGACCGTCTGACCCCGAGCGCGAAGCACG
 CGGTTGCGAAGGCGATTACCGATGCGCACCGTGGTCTGACCGGCACCCAGCACTTCCTGG
 CGCAAGTGAAC TTTCAGGAGCAACCGGCGGGTAACTGTTCCTGGGTGGCGTTCAGCAAG
 GTGGCGACACCATCTTTGTTTCATGGTCTGCACCGTGAGGGCCGTAGCGCGGATCTGAAGG
 GCCAGCTGGCGCAACGTATTGTTGACGATGTGAGCGTTGCGGCGGAAATCGACCGTAAAC
 ACATTTGGGTGTACTTCGGCGAGATGCCGGCGCAGCAAATGGTTGAATATGGCCGTTTCCT
 G **GAAAATCTGTACTTCCAAGGT** – 3'

Figure 4.10. Codon-optimized nucleotide sequence for expression of *tcis*-CaaD in *E. coli*. Highlighted nucleotides encode a TEV recognition sequence. Underlined ATG methionine codon is cleaved by MAP.

Plasmid pAS1.046

Previously, the gene encoding for wild-type Cg10062 from *Corynebacterium glutamicum* was codon optimized to be expressed in *E. coli* and modified with a TEV protease recognition site (ENLYFQG) (Figure 4.11).⁸ This was cloned into pET-21a(+) to contain a C-terminal His₆- tag flanked by 5'-NdeI and 3'-XhoI restriction sites. This plasmid encoding wild-type Cg10062, pAS1.046, was then mutated to form the variants studied including Cg10062 (E114N).



Figure 4.11. Plasmid map of pAS1.046.

Plasmid pAS2.031

MSAD from *Corynebacterium bacterium* FG41 was synthesized similarly to wild-type Cg10062 (Figure 4.12).⁸

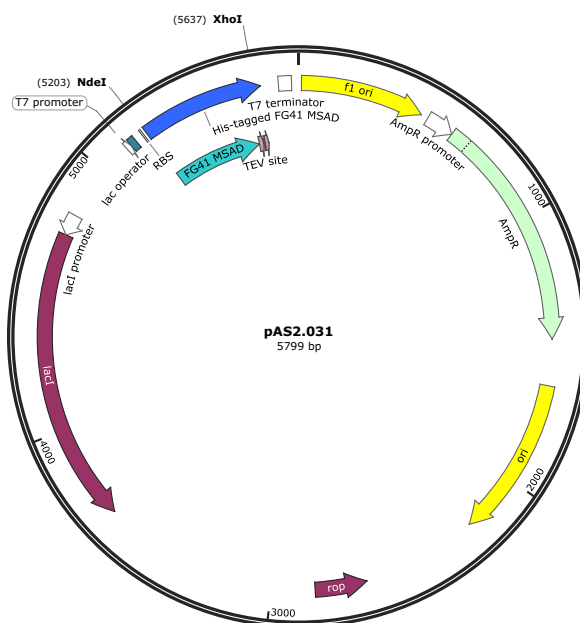


Figure 4.12. Plasmid map for pAS2.031.

Plasmid pKK1.1025

MmsB was obtained by the amplification of the *mmsB* gene in *P. putida* KT2440 genomic DNA and cloning into the pET-21a(+) vector at the NdeI and XhoI sites yielding plasmid pAS2.031 (Figure 4.13).⁹



Figure 4.13. Plasmid map for pKK1.1025.

Plasmid pAS5.003

The gene encoding Cg10062 (E114N) from plasmid pAS2.100 was ligated into pET-21a(+) vector through NdeI and XhoI cut sites after restriction digest as previously described.⁸ Bgl-Brick vector pBbA1a-RFP was purchased from Addgene. The biosensor reporter plasmid, pAS5.003, was constructed as described elsewhere.⁸ The *mmsR* gene as well as a *mmsA* promoter was amplified from *Pseudomonas denitrificans*. A J23119 auto promoter and ribosomal binding site was synthesized by Integrated DNA Technologies. The pBbA1a-RFP vector was digested with BglII and SalI and purified. HiFi DNA assembly was performed according to manufacturer instructions to build plasmid pAS5.003 (Figure 4.13).

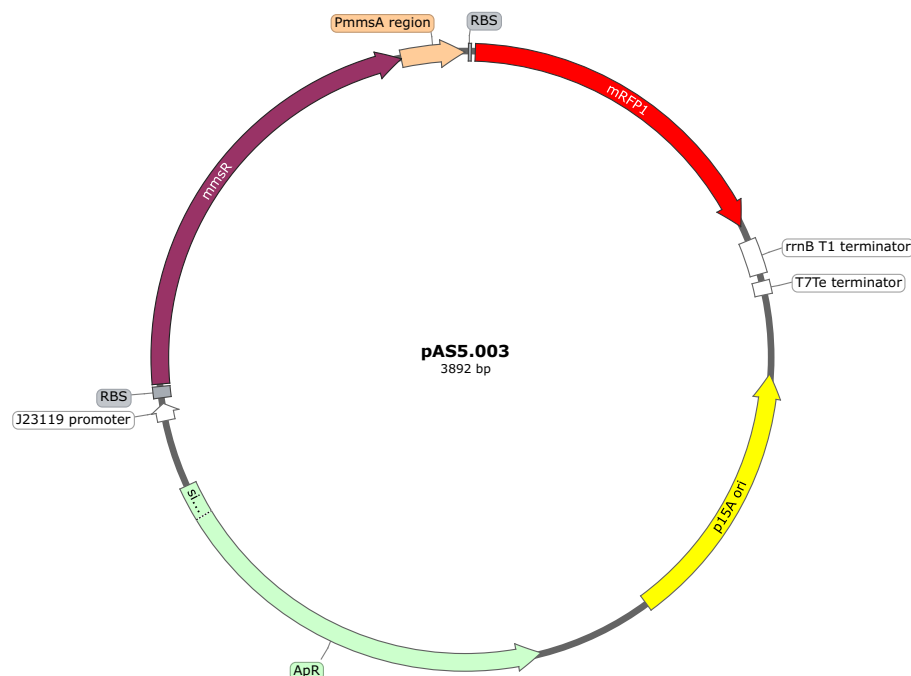


Figure 4.14. Plasmid map of pAS5.003.

Plasmid pAS5.005

Plasmid pAS5.005 for the production of 3HP was constructed using NEB HiFi DNA Assembly as described elsewhere.⁸ The *cg10062* gene was amplified from pAS1.046 and *ydfg* was amplified from the previously designed plasmid pAS2.084 and a stop codon was introduced to the end of each gene. J23119 and J23102 promoters each with ribosomal binding sites were designed and synthesized by Integrated DNA Technologies. The pET-28a(+) vector was digested with BglIII and EcoRV and HiFi DNA assembly was used according to manufacturer's instructions to assemble the pAS5.005 plasmid (Figure 4.14).

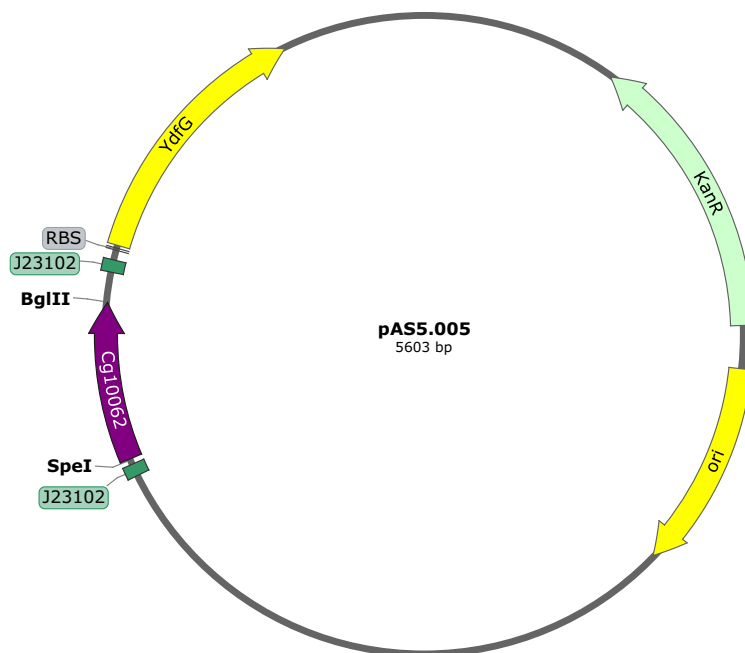


Figure 4.15. Plasmid map of pAS5.005.

Plasmid pAS5.008

pAS5.008 was designed in the same way, except the gene *Cg10062* (*E114N*) was amplified from pAS2.100 instead of *cg10062* (Figure 4.16).⁸ Due to unsuccessful cloning attempts, molecular optimization was performed to obtain correct plasmids without error. Electrocompetent *E. coli* Mg1655 cells were transformed (as previously described) with reporter plasmid pAS5.003 and plated on LB plates with kanamycin. Electrocompetent Mg1655/pAS5.003 cells were prepared from the transformants. These were then transformed with either plasmid pAS5.005 or pAS5.008 and plated onto LB plates containing ampicillin.

4.12.2 Q5 Site-Directed Mutagenesis

To create variants of *cis*-CaaD and *Cg10062*, Q5 site-directed mutagenesis was performed using pNJD1.002 as the template, unless otherwise indicated. PCR was carried out in a Bio-Rad DNA Engine Peltier Thermal Cycler (Hercules, CA). Q5 site-directed mutagenesis was completed in 3 steps. Step 1 includes exponential amplification from parent template (Tables 4.2 and 4.3) using the primers listed in Table 4.6. Step 2 is Kinase, Ligase and DpnI (KLD) treatment

(Table 4.7) of the PCR product from step 1. The final step is the transformation of the KLD product to isolate the plasmid with the desired modification. KLD treatments were incubated at room temperature for 5 minutes.

Table 4.6. Primers used for Q5 site-directed mutagenesis.

Variant	Primer	Sequence
Cg10062 (E114Q)	AS001 (F)	CAACATGACCC <u>AGT</u> ATGGCCGTC
	AS002 (R)	CTACCCGGGATTT <u>CGG</u> TA
Cg10062 (E114D)	AS003 (F)	CAACATGACCGATTATGGCCGTC
	AS002 (R)	CTACCCGGGATTT <u>CGG</u> TA
Cg10062 (E114N)	AS026 (F)	CAACATGACCA <u>ACT</u> ATGGCCGTCTG
	AS002 (R)	CTACCCGGGATTT <u>CGG</u> TA
Cg10062 (Y103F)	AS023 (F)	AGTGTGGGTTTTTATTACCGAAATCCC
	AS022 (R)	TCCTCGTTCGGGATACCC
<i>cis</i> -CaaD (E114Q)	KS001 (F)	GCAAATGGTT <u>CAG</u> TATGGCCGTTTCCTGCCGC
	KS002 (R)	TGCGCCGGCATCTCGCCG
<i>cis</i> -CaaD (E114D)	KS003 (F)	GCAAATGGTT <u>GAC</u> TATGGCCGTTTCCTGCCGC
	KS002 (R)	TGCGCCGGCATCTCGCCG
<i>cis</i> -CaaD (E114N)	KS004 (F)	GCAAATGGTTA <u>ACT</u> ATGGCCGTTTCCTGCCG
	KS002 (R)	TGCGCCGGCATCTCGCCG
<i>cis</i> -CaaD (H28A)	KS005 (F)	TACCGATGCGG <u>CCC</u> GTGGTCTGACC
	KS006 (R)	ATCGCCTTCGCAACCGCG
<i>cis</i> -CaaD (R73A)	KS007 (F)	CCGTGAGGGCG <u>GCC</u> AGCGCGGATC
	KS008 (R)	TGCAGACCATGAACAAAGATG
<i>cis</i> -CaaD (Y103A)	KS009 (F)	CATTTGGGTG <u>GCC</u> TTCGGCGAGATG
	KS010 (R)	TGTTTACGGTTCGATTTCC
<i>cis</i> -CaaD (Y103F)	KS011 (F)	CATTTGGGTGTT <u>CTT</u> CGGCGAGA
	KS012 (R)	TGTTTACGGTTCGATTTCCG
<i>cis</i> -CaaD (T32A)	KS013 (F)	CCGTGGTCTGG <u>CCG</u> GCACCCAGC
	KS014 (R)	TGCGCATCGGTAATCGCCTTCGC
<i>cis</i> -CaaD (T34A)	KS015 (F)	TCTGACCGGCG <u>CCC</u> AGCACTTCC
	KS016 (R)	CCACGGTGCGCATCGGTAATC
<i>tcis</i> -CaaD	KS052 (F)	TTGAATATGGCCGTTTCCTGGAAAATCTGTACTTCCAAG
		GTCT
	KS053 (R)	CAGGAAACGGCCATATTC

^aUnderlined nucleotides encode variant amino acid.

Transformations were carried out using a Bio-Rad Gene Pulser II electroporation system (Hercules, CA). For the transformation of PCR products, 50 μ L *E. coli* DH5 α electrocompetent cells were thawed on ice and 5 μ L of the KLD reaction was added to the electrocompetent cells. The sample was transferred to a cold sterile Gene Pulser electroporation cuvette and the cells were pulsed at 2.5 kV (25 mF capacitance, 200 Ω resistance). The cells were resuspended in 1

mL SOC and shaken at 37 °C for 1 hour. The cells were pelleted at 13,000 rpm in a microcentrifuge and the SOC was decanted. The cells were resuspended in 100 mL SOC, spread onto LB plates containing ampicillin and incubated at 37 °C overnight.

Table 4.7. KLD Treatment.

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

Single colonies were inoculated into separate culture tubes containing 5 μ L of LB media with ampicillin and the cultures were shaken at 37 °C overnight. DNA extraction from the cell pellets was carried out using a QIAprep Miniprep kit, following manufacturer instructions. Isolated plasmids were sequenced by the Michigan State University Research Technology Support Facility (MSU RTSF) Genomics Core and the plasmids containing the desired mutations were transformed into electrocompetent *E. coli* BL21(DE3) for protein expression.

4.12.3 Activity Assay

Steady-state kinetics were measured at 25 °C using a Shimadzu UV2600 spectrophotometer. All assays were carried out in triplicate with a final volume of 1 mL. Enzyme activity was measured using a coupled enzyme assay shown previously.¹⁰⁻¹⁴ Reduction of acetaldehyde by NADH-dependent alcohol dehydrogenase (ADH) was monitored by following the oxidation of NADH at 340 nm ($\epsilon = 6220 \text{ M}^{-1} \text{ cm}^{-1}$).

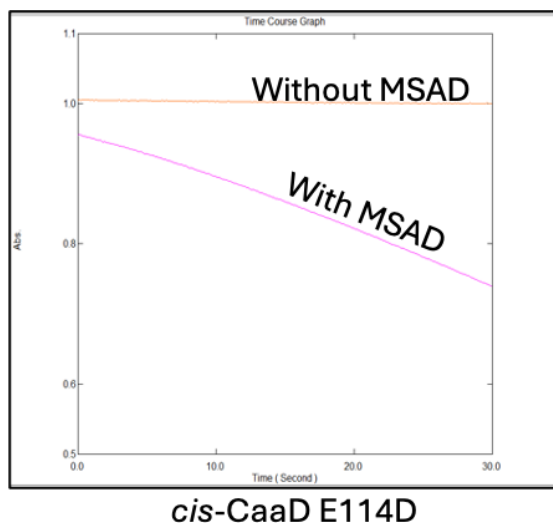


Figure 4.16. Kinetic trend example.

Kinetic assays used for determining 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. Stock solutions of lyophilized ADH were prepared in deionized water per manufacturer instructions. Stock solutions of acetylenecarboxylic acid (ACA, 5 M) and *cis*-3-chloroacrylic acid (CCA, 5 M) were prepared by diluting the appropriate concentration of each reagent in 100 mM sodium phosphate pH 8.0 and adjusting to pH 8.0 by addition of 10 N sodium hydroxide. Initial kinetic screens for *cis*-CaaD variants contained NADH (0.14 mM), ADH (12 U), MSAD (2.4 U), ACA or CCA (concentration varied) and *cis*-CaaD and/or variant (concentration varied). Assays were carried out in triplicate (1 mL) at 25 °C using a Shimadzu UV2600 spectrophotometer. Excess quantities of coupling enzymes, ADH and MSAD, were used to ensure the activities of *cis*-CaaD and variants were rate-limiting. Due to the varying concentration of *cis*-CaaD or variant used to observe measurable activity, these kinetic screens were used in the determination of product profiles only and not used in the direct comparison of specific activities. An example of the kinetic trends with and without MSAD for a hydratase only variant is shown in Figure 4.16. The calculation for determining specific activity of *cis*-CaaD and each variant is shown in Equation 4.1.

$$\text{Specific Activity} \frac{\mu\text{mol}}{\text{min} * \text{mg}}$$

$$= \text{Rate} \left(\frac{\text{mAu}}{\text{min}} \right) \times \frac{\text{mol} * \text{cm}}{6220 \text{ L}} \times \frac{1}{(\text{path length}) \text{ cm}} \times \frac{\text{mL}}{\text{mg}} \times \frac{\text{mL}}{[\text{enzyme}] \text{ mL}} \times \frac{(\text{final assay volume}) \text{ mL}}{(\text{volume of enzyme}) \text{ mL}} \times \frac{10^6 \mu\text{mol}}{1 \text{ mol}} \times \frac{1 \text{ L}}{1000 \text{ mL}}$$

Equation 4.1 Calculation for enzyme specific activity in kinetic assay.

Cg10062 assays contained NADH (0.34 mM final concentration), ADH (12 U), MSAD (1.2 U), and ACA pH 8 (0.5 mM). The final pH of each assay was 8. The initial rates of each reaction of Cg10062 variant relative to varied ACA concentrations (1-5000 mM) were plotted to fit the Michaelis-Menten model and analyzed using OriginPro 9.0 (Figure 4.17). Specific activities were determined as described previously. Kinetics were carried out using a Molecular Devices SpectraMax® iD3 multi-mode microplate reader.

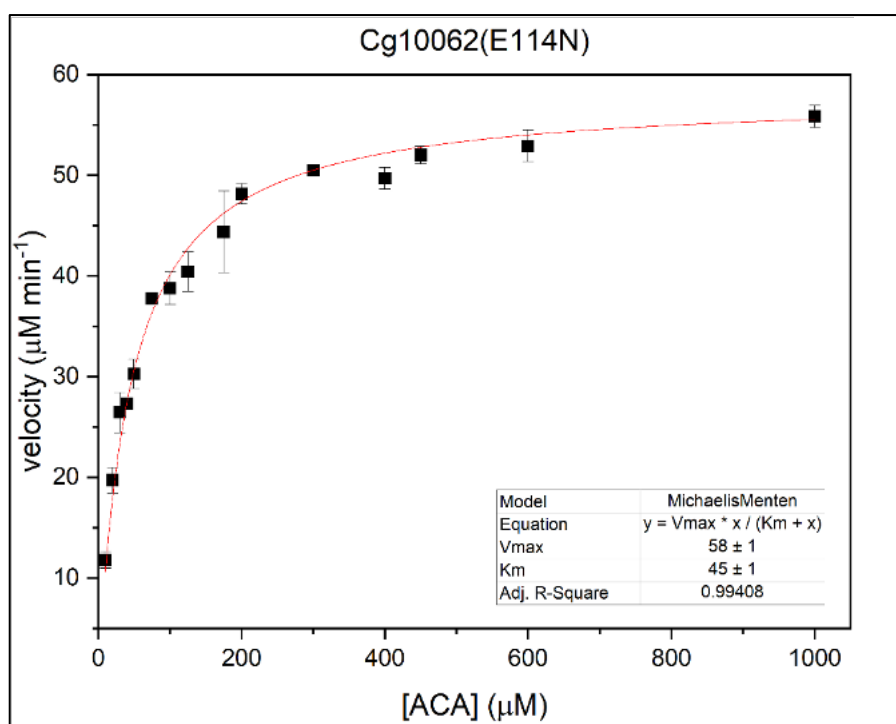


Figure 4.17. Michaelis-Menten kinetics of Cg10062 (E114N).⁹

Assays for truncated and chimera *cis*-CaaDs contained NADH (0.12 mM), ADH (12 U), MSAD (4.8 U), ACA pH 8 or CCA pH 8 (concentration varied) and *cis*-CaaD or variant (concentration varied). The final pH of each assay was 8. The initial rates of each reaction of *cis*-

CaaD variant relative to varied ACA or CCA concentrations (0 – 60 mM) were plotted to fit the Michaelis-Menten model and analyzed using GraphPad. Kinetics were carried out using a Shimadzu UV2600 spectrophotometer. Steady-state kinetics were carried out in triplicate at 25 °C in 100 mM sodium phosphate pH 8.0 to a final volume of 1 mL (Figure 4.18).

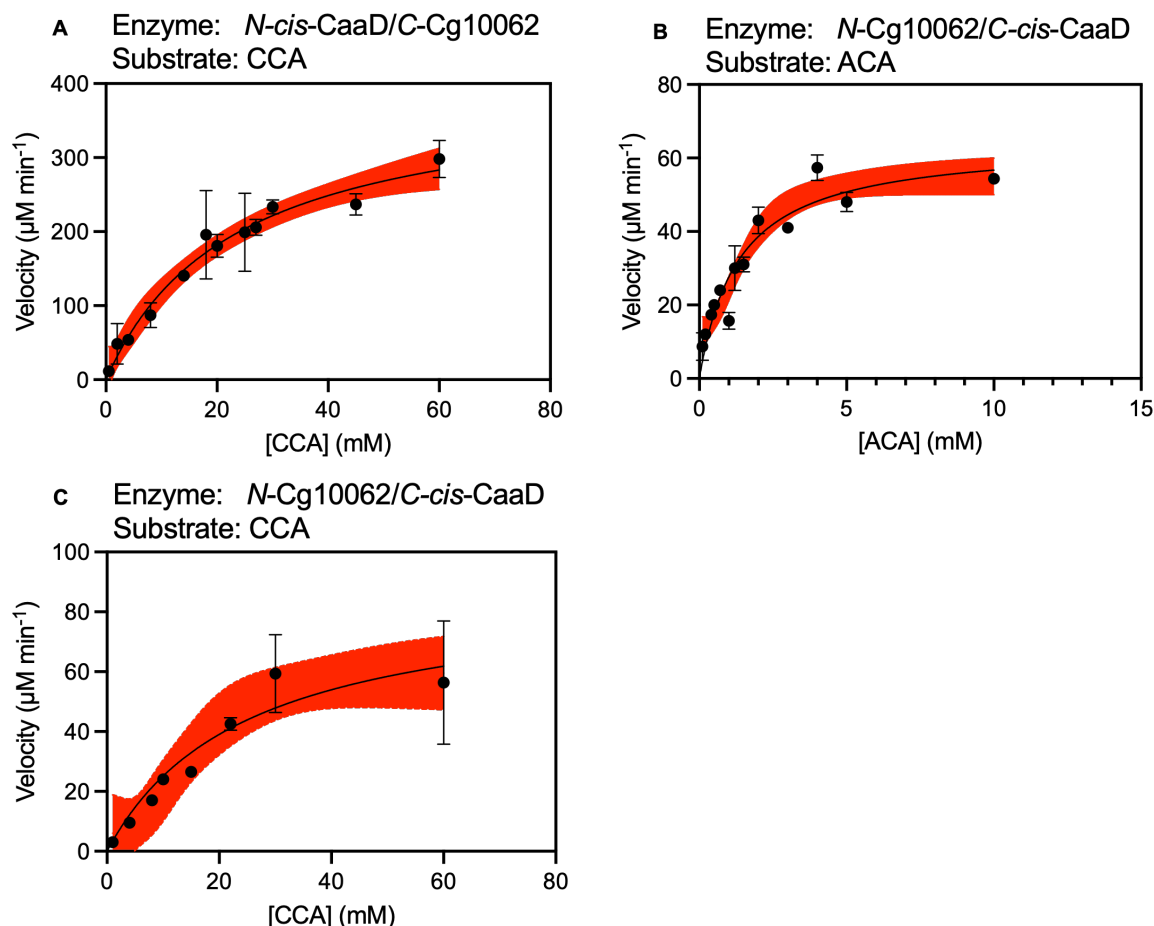


Figure 4.18. Steady state kinetics of (A) *N-cis-CaaD/C-Cg10062* with CCA, (B) *N-Cg10062/C-cis-CaaD* with ACA, and (C) *N-Cg10062/C-cis-CaaD* with CCA.

The specific activity of MSAD was measured using the same coupled enzyme assay. 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 added, except ACA and MSAD. The assays were initiated with the addition of ACA. After complete conversion of ACA to MSA (30 s), MSAD was added to the reaction and the activity of MSAD was measured.

4.12.4 Biosensor Fluorescence Assay

To determine the native fluorescence of the Bgl brick vector, pBbA1a-RFP, from induction by IPTG, a single colony of MG1655/pBbA1a-RFP was inoculated into a 5 mL culture of LB with antibiotic. Cells were grown at 37 °C shaking at 200 rpm for 12 hours overnight. A 5 mL culture was then reinoculated into LB to an OD₆₀₀ of 0.05 and grown at 37 °C until an OD₆₀₀ of 0.5-0.7 was reached. Aliquots of 190 µL of culture were then transferred to a Corning® costar black clear-bottomed 96-well plate. IPTG (10 µL) was added to each well to a final concentration of 0, 0.01, 0.1, and 1 mM. Each sample was prepared in triplicate and sealed with a MicroAmp™ Optical Adhesive film. Control wells containing 200 µL M9 salts were used as blanks. OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 hours.

To determine fluorescence of the biosensor system based on varying 3HP concentrations, a single colony of MG1655/pAS5.003, MG1655/pAS5.003/pAS5.005, and MG1655/pAS5.003/pAS5.008 was inoculated into 5 mL LB media with respective antibiotic.¹⁵ Cells were grown at 37 °C shaking at 200 rpm for 12 hours overnight. Cultures were then pelleted by centrifugation (17,000 rpm for 1 minute) and resuspended in M9 salts to remove residual media and repeated twice more. The cell pellet was then resuspended in 2 mL of M9 salts and used to reinoculate M9/Glu with respective antibiotic to an OD₆₀₀ of 0.1. Aliquots of 180 µL of each culture were transferred to wells of a Corning® costar 96-well black clear-bottomed plate and sealed with a Genetix BreatheSeal film. The plate was incubated, shaking at 200 rpm, at 37 °C until an OD₆₀₀ of 0.5-0.7 was reached. Alternatively, cultures were able to reach an OD₆₀₀ of 0.5-0.7 in culture tubes before aliquoting into the 96-well plate. A 1 M stock solution of 3HP was prepared in a 5 mL volumetric flask by diluting 1.39 mL of 30% (w/v) 3HP in M9 salts and pH adjusted to 7 using 10N NaOH. From this 0, 5, 10, 25, 50, 100, 250, 500, 750, and 1000 mM 3HP stock solutions were prepared in M9/Glu media. Each 3HP stock solution was added (20 µL) to wells to obtain final concentrations of 0, 0.5, 1, 2.5, 5, 10, 25, 50, 75, and 100 mM, respectively. Each sample was

prepared in triplicate and sealed with a MicroAmp™ Optical Adhesive film and incubated shaking at 200 rpm at 37 °C. Control wells containing 200 µL M9 salts were used as blanks. OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 6, 12, and 24 hours. The sample experiment was performed using varying concentrations of ACA in place of 3HP. To determine RFP fluorescence as a function of ACA, the OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 hours. The same protocol was also followed using higher concentrations of 3HP and ACA to a final concentration of 0, 75, 100, 125, 150, 175, 200, 250 and 300 mM. OD₆₀₀ and fluorescence (ex/em = 535/585 nm) were measured at 0, 3, 6, and 12 hours.

4.12.5 Error-Prone Mutagenesis

Error-prone mutagenesis was carried out on wild-type Cg10062 as well as Cg10062 (E114N) using a GeneMorph II Random Mutagenesis kit according to manufacturer's instruction.¹⁶ Error-prone mutagenesis was carried out using forward primer (CTTTAAGAAGGAGATATACATATG) and reverse primer (TATATGCTCGAGGCCTT) on pAS2.100 and pAS1.046 (Table 4.8) flanking the Cg10062 gene. PCR was carried out using the suggested conditions (Table 4.9). Varying template DNA of 100, 500, and 1000 ng was used to screen for high, medium, and low mutation frequency (Table 4.10). Sanger sequencing was performed to confirm the difference in mutation rate. PCR products were isolated using a Zymogen PCR clean up kit according to manufacturer instructions and a restriction digest using NdeI and XhoI was performed to isolate mutated Cg10062 vector as well as on vector pET-21a(+).

Table 4.8 Error-prone PCR reaction components.

Component	Volume	Final Concentration
Mutazyme Buffer (10X)	5 µL	1X
dNTP mix (40 mM)	1 µL	200 µM each
Forward Primer (75 ng µL ⁻¹)	1 µL	1.5 ng
Reverse Primer (75 ng µL ⁻¹)	1 µL	1.5 ng
Mutazyme II DNA Polymerase	1 µL	2.5 U
Template DNA (varies)	1 µL	-
Deionized water	40 µL	-

Table 4.9. Error-prone PCR conditions.

Step	Temperature	Time
Initial Denaturation	95 °C	2 min
30 cycles	95 °C	30 sec
	56 °C ^a	30 sec
	72 °C	1 min ^b
Final Extension	72 °C	10 min
Hold	4 °C	-

^aT_m – 5 °C. ^bOne minute for less than or equal to one kb targets and 1 minute per kb for greater than one kb targets.

Table 4.10. Template DNA to add for desired mutation frequency.

Mutation Frequency (Mutations/kb)	Initial Target Quantity
0-4.5 (low)	500-1000 ng
4.5-9 (medium)	100-500 ng
9-16 (high)	0.1-100 ng

The amplified PCR products were ligated into pET-21a(+) vector with a vector to insert ratio of 1:3 using T4 ligase for two hours at room temperature and transformed into BL21(DE3) cells and plated on LB with ampicillin and grown overnight at 37 °C (Table 4.11). From the culture plate, 96 colonies were selected and inoculated into a deep well 96 well plate, each well containing 1 mL of LB with antibiotic. Cultures were allowed to grow at 37 °C overnight shaking. Cultures were then reinoculated to an OD₆₀₀ of 0.05 and induced with IPTG to a final concentration of 1 mM after 1.5 hours. Cells were then pelleted and resuspended in 500 µL of cell lysis buffer containing 0.25 mg mL⁻¹ lysozyme and incubated for 10 minutes at 37 °C shaking. The plate was then centrifuged and 10 µL of supernatant was added to fresh 1 mL 96 well plates containing 10 µL NADH (10 mg mL⁻¹) and 0.5 mg mL⁻¹ MmsB. The reaction was initiated by the addition of 50 µL ACA (1 mM final concentration). Absorbance was measured at 340 nm following the oxidation of NADH. ACA that has been converted to MSA will be converted to 3HP by the NADH dependent enzyme MmsB. Conversion of MSA to 3HP is monitored by NADH depletion. Thus, indicating high MSA production and a promising variant. Cg10062 (E114N) was also screen similarly without undergoing error-prone mutagenesis to compare activities and screen for a more active mutant.

Table 4.11. T4 ligation conditions.

Component	Volume	Final Concentration
Ligation Buffer (10X)	2 μ L	1X
Vector DNA (4 kb)	2 μ L	50 ng
Insert DNA (1 kb)	2 μ L	37.5 ng
Deionized Water	13 μ L	-
T4 DNA Ligase	1 μ L	20 U

4.12.6 ^1H NMR Studies

Identification of products from wild-type Cg10062 and Cg10062 variant catalyzed reactions with acetylene carboxylate were determined by ^1H NMR spectroscopy on a 500 MHz Varian NMR Spectrophotometer and analyzed using MestreNova 14.2.0 software. ^1H wet1D was used for solvent suppression. 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.

All stock solutions were prepared in 100 mM sodium phosphate pH 8.0. To prepare a 10 mL, 1 M stock solution of ACA pH 8, an appropriate volume of ACA was diluted in 100 mM sodium phosphate pH 8.0 and neutralized with 10 N sodium hydroxide until pH 8. ACA (111 mM, 20 μ L of 5 M stock) was dissolved in 830 μ L of 100 mM sodium phosphate pH 8.0. Solutions of CCA were prepared similarly. The reactions were initiated with the addition of purified *cis*-CaaD or variant (50 μ L of 4.8 mg mL^{-1}). After incubating the samples at 25 $^\circ\text{C}$, 150 μ L aliquots of each sample were quenched by the addition of 5 M H_2SO_4 (2 μ L) after 0 h 15 min, 30 min, 1 h, 2 h, and 16 h. The samples were spun in a microcentrifuge to remove precipitated protein. 100 μ L of each sample was combined with TSP (10 mM, 70 μ L of a 100 mM stock), and DMSO- d_6 (30 μ L). The final volume was adjusted to 700 μ L using 100 mM sodium phosphate pH 8.0 for NMR spectroscopy.

NMR spectra were obtained for each sample (Figure 2.5). The resonance at δ 2.91 (s, 1H) corresponds to ACA. The resonance at δ 6.3 (d, 1H) and 6.4 (d, 1H) corresponds to CCA. Resonances at δ 3.20 (d, 2H), δ 9.50 (t, 1H) and δ 2.30 (d, 2H), 5.13 (t, 1H) correspond to

malonate semialdehyde and its hydrate, respectively. Resonances at δ 2.03 (d, 3H), 9.47 (q, 1H) and δ 1.12 (d, 3H), 5.05 (q, 1H) correspond to acetaldehyde and its hydrate, respectively.

4.12.7 Protein Crystallization

Screening for crystallization conditions for *cis*-CaaD and variants was performed with a mosquito robot (TTP Labtech) using the screens PEG/Ion and PEG/Ion 2 by the hanging drop vapor diffusion method in 96 well plates. Proteins were screened for crystallization at room temperature or 4 °C in 3 μ L drops that contained 1.5 μ L of the enzyme solution (18 mg mL⁻¹ in 10 mM Tris-SO₄ pH 8.0) and 1.5 μ L of 4-fold diluted crystallization screen solution. *cis*-CaaD and mutant crystals grew under multiple conditions and were used in the following steps without further optimization. Table 4.12 lists several conditions that yielded suitable crystals. Other than the native protein, all mutant crystals were soaked for various lengths of time in 2 μ L of substrate solutions, which contained 1 μ L of 2.5 to 10 mM ACA or CCA and 1 μ L of 2x concentrated appropriate crystallization solution. Crystals were protected in Paratone or glycerol with crystallization solution before they were flash-frozen in liquid nitrogen.

Screening for crystallization conditions for Cg10062 (E114N) was performed using commercial screens PEG/Ion and PEG/Ion 2 and the hanging drop vapor diffusion method in 24 well plates or was performed with a mosquito robot (TTP Labtech) in 96 well plates. Proteins were screened for crystallization at room temperature or 4 °C in 2 or 3 μ L drops containing 1:1 ratio of the enzyme solution (18 mg mL⁻¹ in 10 mM Tris-SO₄, pH 8.0) to crystallization screen solution. Crystals of the E114N variant of Cg10062 grew under multiple conditions. Conditions used for soaking experiments was 2% v/v TacsimateTM (pH 5.0) and 5% w/v Polyethylene glycol 3,350 or 0.2 M Ammonium sulfate and 20% w/v Polyethylene glycol 3,350. After washing the crystals in well solution, they were soaked for 1 or 25 minutes in 4 μ L of substrate solution which contained 1 μ L of 10 mM ACA and 3 μ L of well solution or 2 μ L of substrate solution which contained 1 μ L of 10 mM ACA and 1 μ L of well solution. Subsequently, crystals were protected in Paratone or

27% glycerol in well solution before being flash-frozen in liquid nitrogen. More information of crystallization conditions used in for Cg10062 variants is reported elsewhere.¹³

For chimeras and truncated *cis*-CaaD, screening for crystallization conditions was performed using commercial screens PEG/Ion and PEG/Ion 2 and the hanging drop vapor diffusion method in 24 well plates. Proteins were screened for crystallization at room temperature in 3 μ L drops containing 1.5 μ L of the enzyme solution (18 mg mL⁻¹ in 10 mM Tris-SO₄, pH 8.0) and 1.5 μ L of crystallization screen solution. *N-cis-CaaD/C-Cg10062* grew with 0.2 M Lithium chloride, 20% w/v Polyethylene glycol 3,350 and *tcis*-CaaD grew with 0.2 M Ammonium sulfate, 20% w/v Polyethylene glycol 3,350. After washing the crystals in well solution, they were soaked at room temperature for 15 minutes in 2 μ L of substrate solution which contained 1 μ L of well solution and 1 μ L of 10 mM ACA for *N-cis-CaaD/C-Cg10062* or CCA for *tcis*-CaaD. Subsequently, crystals were protected in 27% glycerol in screen solution before being flash-frozen in liquid nitrogen.

Diffraction data were collected from a single crystal at Argonne National Laboratory (APS) (Argonne, IL) LS-CAT, (Sector 21-ID-D, F, and G) using either a Dectris Eiger 9M, MAR300, or MAR350 detector. The individual images were extracted from an HDF5 container using the *eiger2cbf* program written by Takanori Nakane. Indexing, integration, rejection, and scaling data were performed with HKL2000. A summary of the data collection statistics is listed in Table 4.13. The crystal structure of *cis*-CaaD and variants were solved with molecular replacement techniques using the wild-type *cis*-CaaD (PDB ID: 3MF8) as a search model. Molecular replacement and subsequent refinement were performed with the PHENIX Suite (version 1.17.1).¹⁷ Thereafter, substrate structures and water molecules were modelled using COOT.¹⁸

4.12.8 X-Ray Crystal Structures of apo, ACA and CCA Soaked *cis*-CaaD (wild type)

The x-ray crystal structure of wild-type *cis*-CaaD that was soaked with ACA substrate was solved to a 2.0 Å resolution and 97% completeness by molecular replacements and refined to R and R_{free} values of 20% and 26%, respectively (Figure 4.19). The asymmetric unit contains three

chains. All residues from Pro-1 to position 145 fit within the electron density with the remaining four residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). Electron density accounted for three ACA molecules in each active site of each chain show to 1.5 rmsd.

The x-ray crystal structure of wild-type *cis*-CaaD that was soaked with CCA substrate was solved to a 1.3 Å resolution and 99% completeness by molecular replacements and refined to R and R_{free} values of 19% and 21%, respectively (Figure 4.20). The asymmetric unit contains one chain. All residues from Pro-1 to position 118 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). A single acetate molecule from crystal screening solution was seen in the active site shown at 1.5 rmsd.

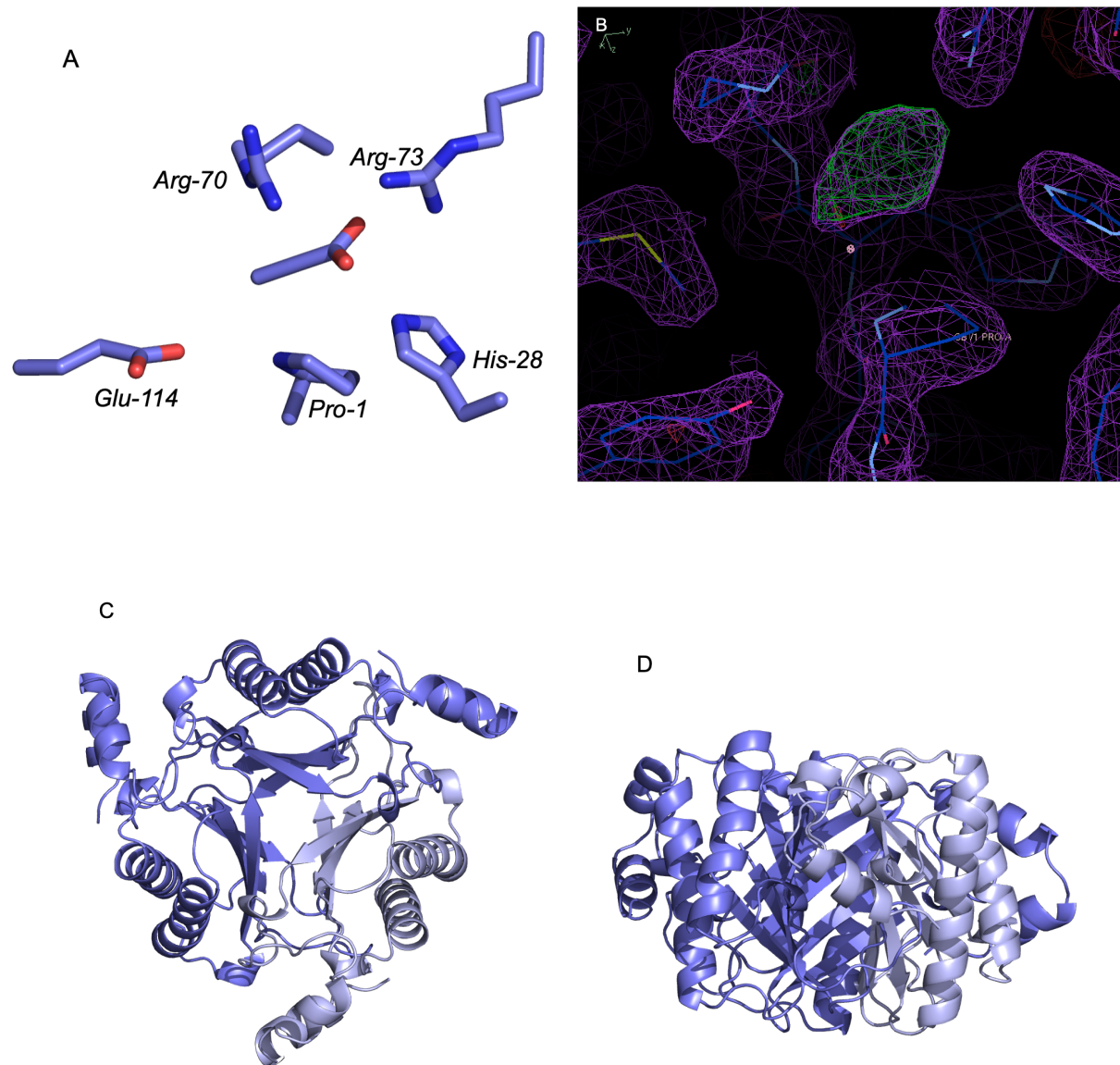


Figure 4.19. Crystal structure of *cis*-CaaD (wild-type) soaked with substrate showing an (A) unreacted ACA in the active site at 2.0 Å, (B) electron density before addition of substrate in active site, (C) trimer with one monomer in a lighter shade, (D) side view of trimer.

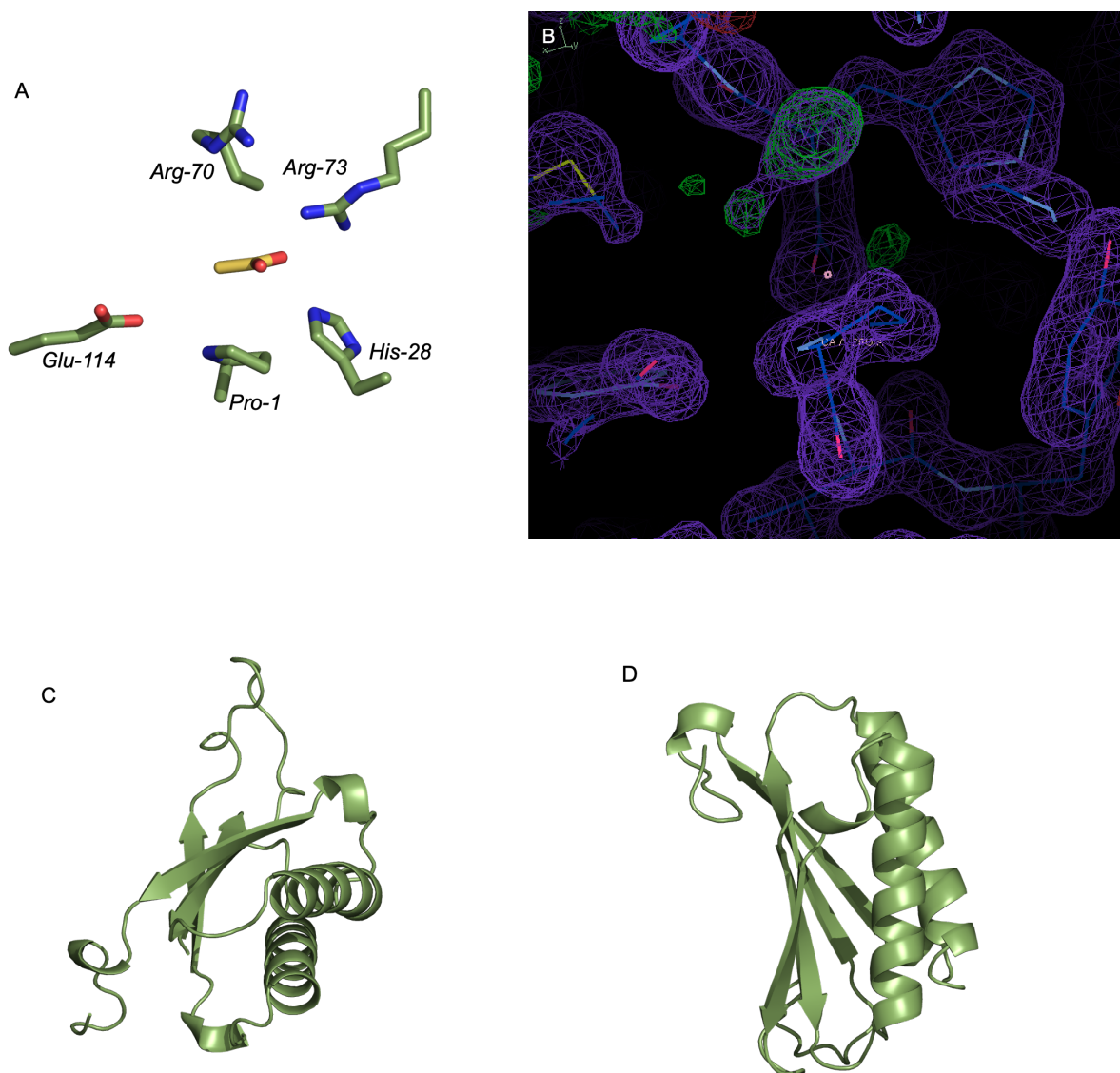
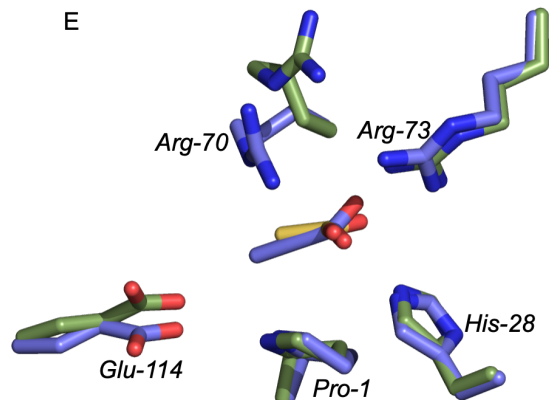


Figure 4.20. Crystal structure of wild-type *cis*-CaaD at 1.3 Å (A) with an acetate in the active site, (B) electron density before addition of acetate, (C) top view of trimer with one monomer in a lighter shade, (D) side view of trimer, and (E) aligned with the ACA-soaked structure showing ACA in the active site with acetate.

Figure 4.20. (cont'd)



4.12.9 X-Ray Crystal Structures of apo, ACA, and CCA Soaked *cis*-CaaD (E114D)

The x-ray crystal structure of apo *cis*-CaaD (E114D) was solved to a 1.6 Å resolution with 88% completeness by molecular replacements and refined to R and R_{free} values of 18% and 22%, respectively (Figure 4.21). There are three chains in the asymmetric unit. All residues from Pro-1 to position 146 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). Sulfate from crystal screening solution was in all three active sites.

The x-ray crystal structure of *cis*-CaaD (E114D) that was soaked with ACA substrate was solved to a 2.5 Å resolution and 95% completeness by molecular replacements and refined to R and R_{free} values of 19% and 25%, respectively (Figure 4.22). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). The structure of *cis*-CaaD E114D had partial occupancy of hydroxypropionate and product malonic semi aldehyde in two chains. The density of the third chain was fully occupied by substrate ACA.

The x-ray crystal structure of *cis*-CaaD (E114D) that was soaked with CCA substrate was solved to a 1.8 Å resolution and 91% completeness by molecular replacements and refined to R and R_{free} values of 18% and 22%, respectively (Figure 4.23). The asymmetric unit contains one chain. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). The third chain is occupied by both a sulfate ion and covalently attached (*N*-prolyl)-ethene.

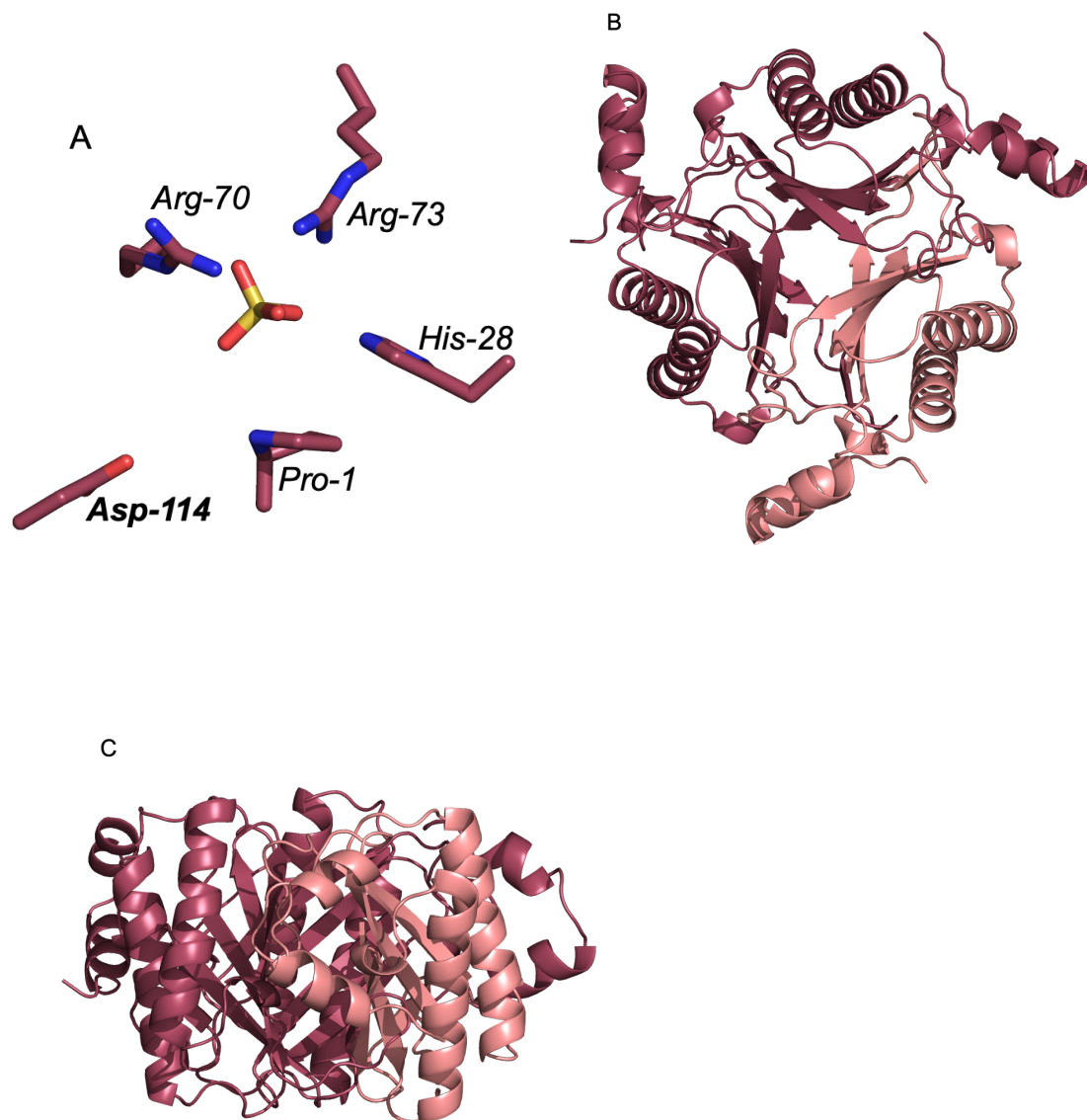


Figure 4.21. Crystal structure of apo *cis*-CaaD (E114D) with (A) sulfate shown in the active site, (B) a top view of the trimer with one monomer a lighter shade, and (C) a side view of the trimer.

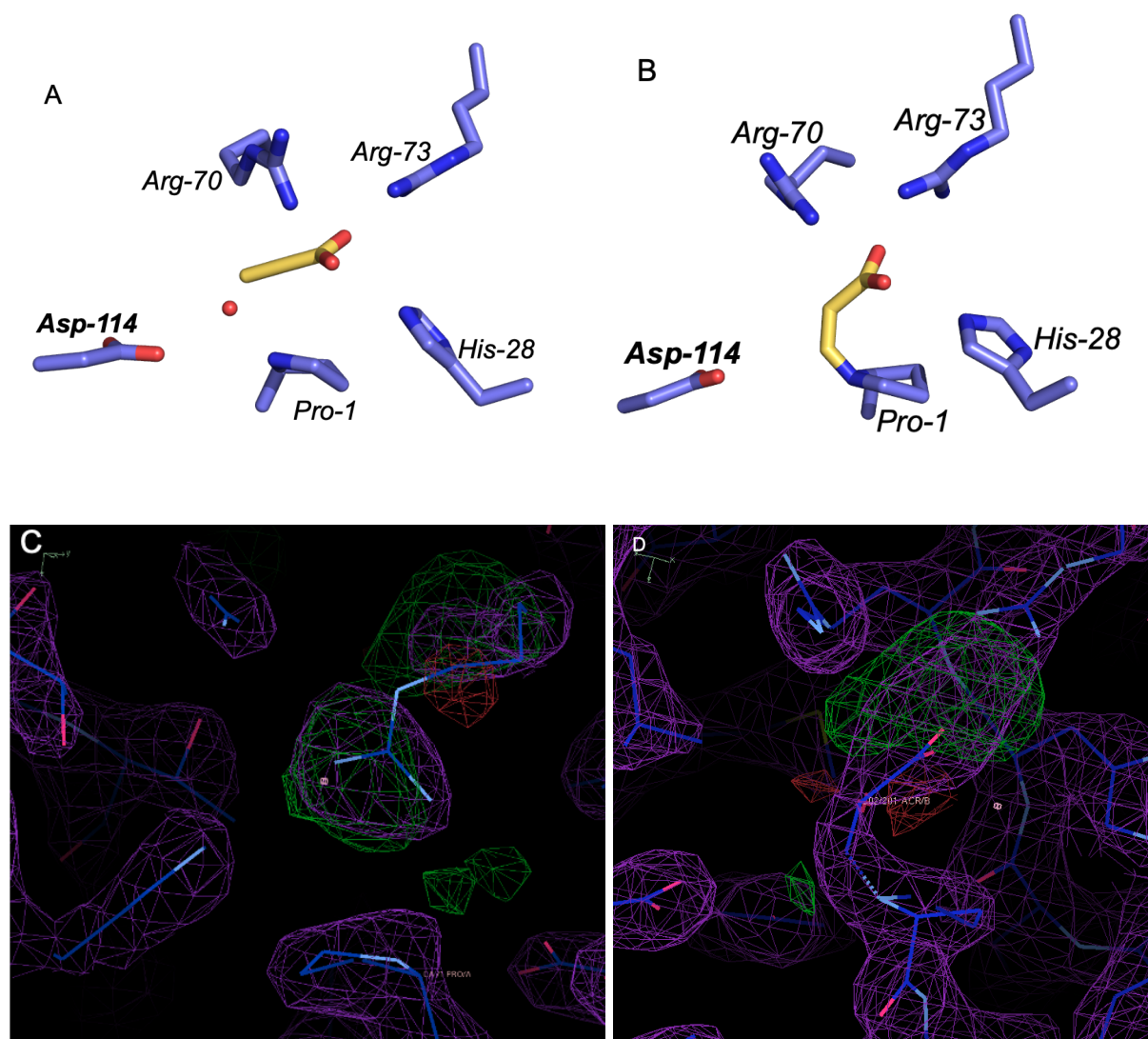
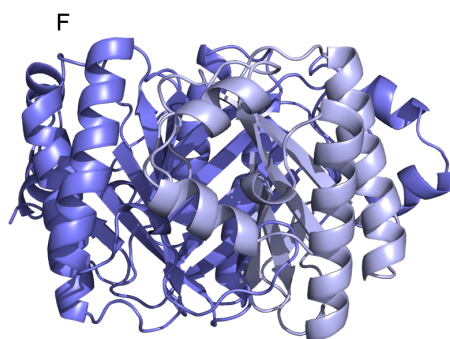
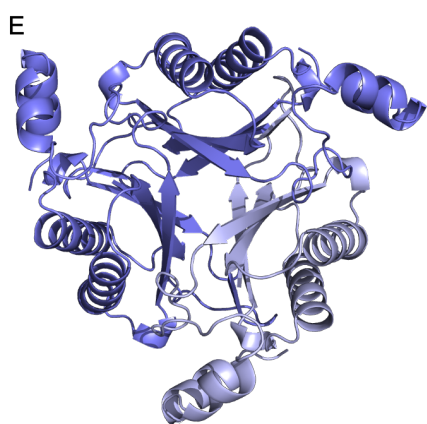


Figure 4.22. Crystal structures of ACA soaked *cis*-CaaD (E114D) at 2.5 Å showing (A) unreacted substrate and (B) covalent intermediate of 3-(*N*-prolyl)-acrylate linked to catalytic Pro-1. The electron density is shown (C) before and (D) after the addition of 3-(*N*-prolyl)-acrylate into the active site. The trimer is shown with (E) a top view and (F) side view.

Figure 4.22. (cont'd)



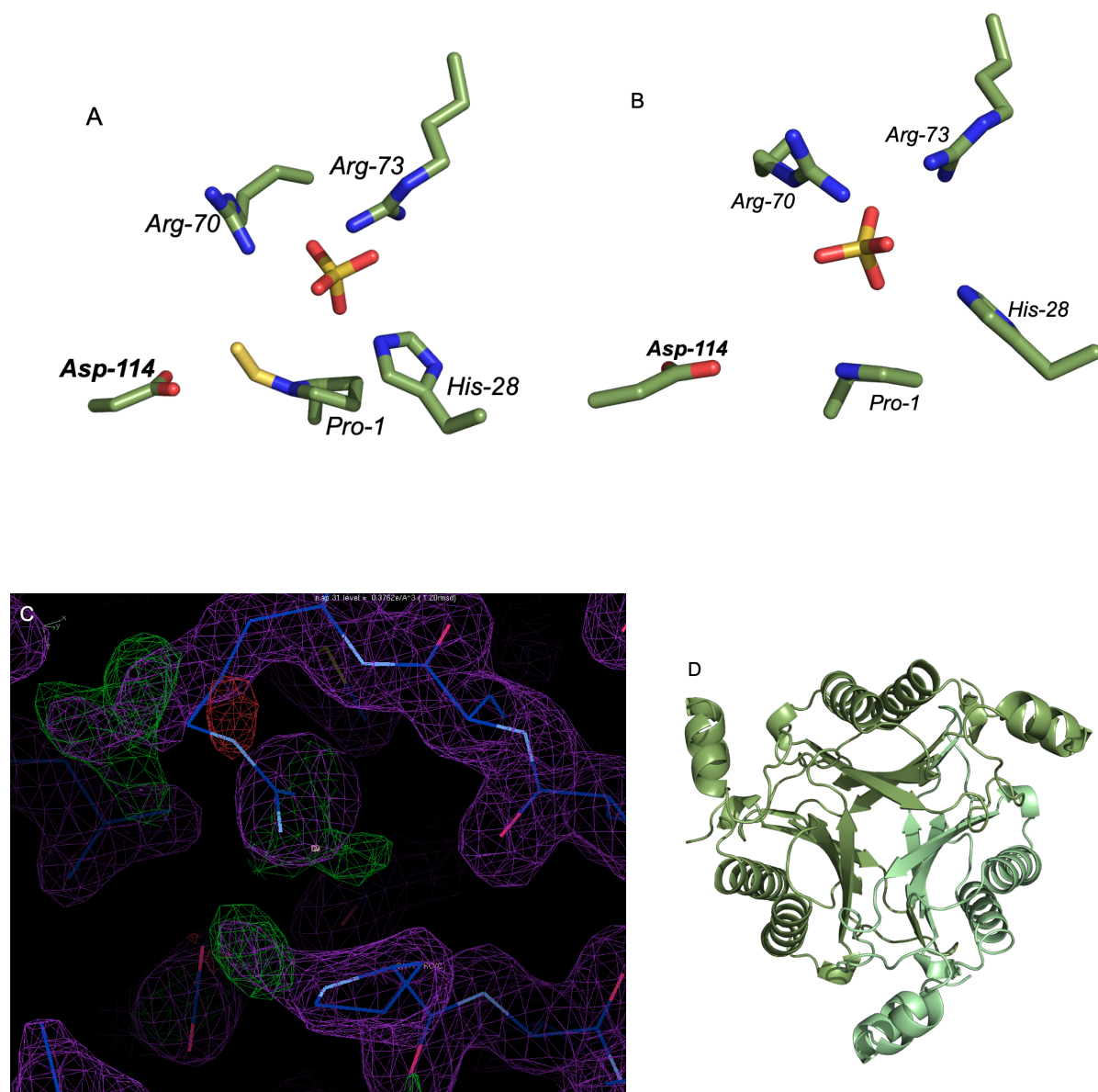
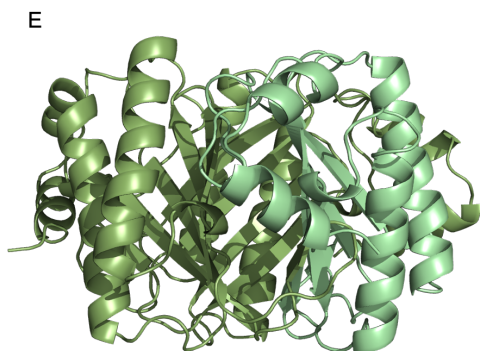


Figure 4.23. Crystal structures of CCA soaked *cis*-CaaD (E114D) at 1.8 Å showing the covalent intermediate of (*N*-prolyl)-ethene linked to catalytic Pro-1 with sulfate (A) and (B) sulfate alone product. Electron density shows active site before the addition of ligand (C). The trimer is shown at a top view (D) with one monomer a lighter shade and at a side view (E).

Figure 4.23. (cont'd)



4.12.10 X-Ray Crystal Structures of apo and ACA Soaked *cis*-CaaD (E114N)

The x-ray crystal structure of apo *cis*-CaaD (E114N) was solved to a 2.5 Å resolution with 89% completeness by molecular replacement and refined to R and R_{free} values of 22% and 29%, respectively (Figure 4.24). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). A sulfate from buffer solution was bound in the active site of each asymmetric unit.

The x-ray crystal structure of *cis*-CaaD (E114N) that was soaked with ACA substrate was solved to a 2.3 Å resolution to 83% completeness by molecular replacements and refined to R and R_{free} values of 20% and 28%, respectively (Figure 4.25). The asymmetric unit contains one chain. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). The first chain was occupied by unreacted ACA substrate. The other two chains contained density that was occupied by malonic semialdehyde product.

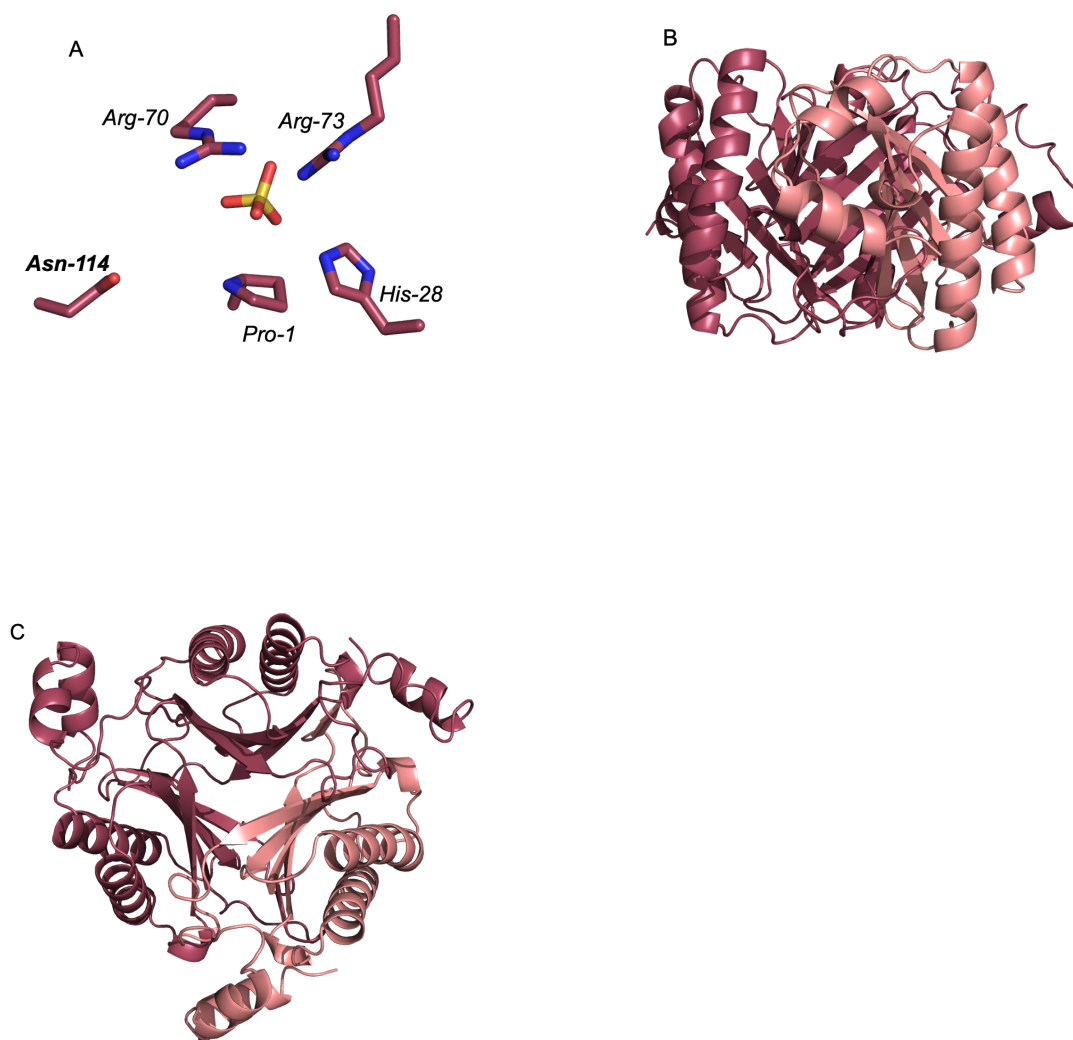


Figure 4.24. Crystal structure of apo *cis*-CaaD (E114N) is shown with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.

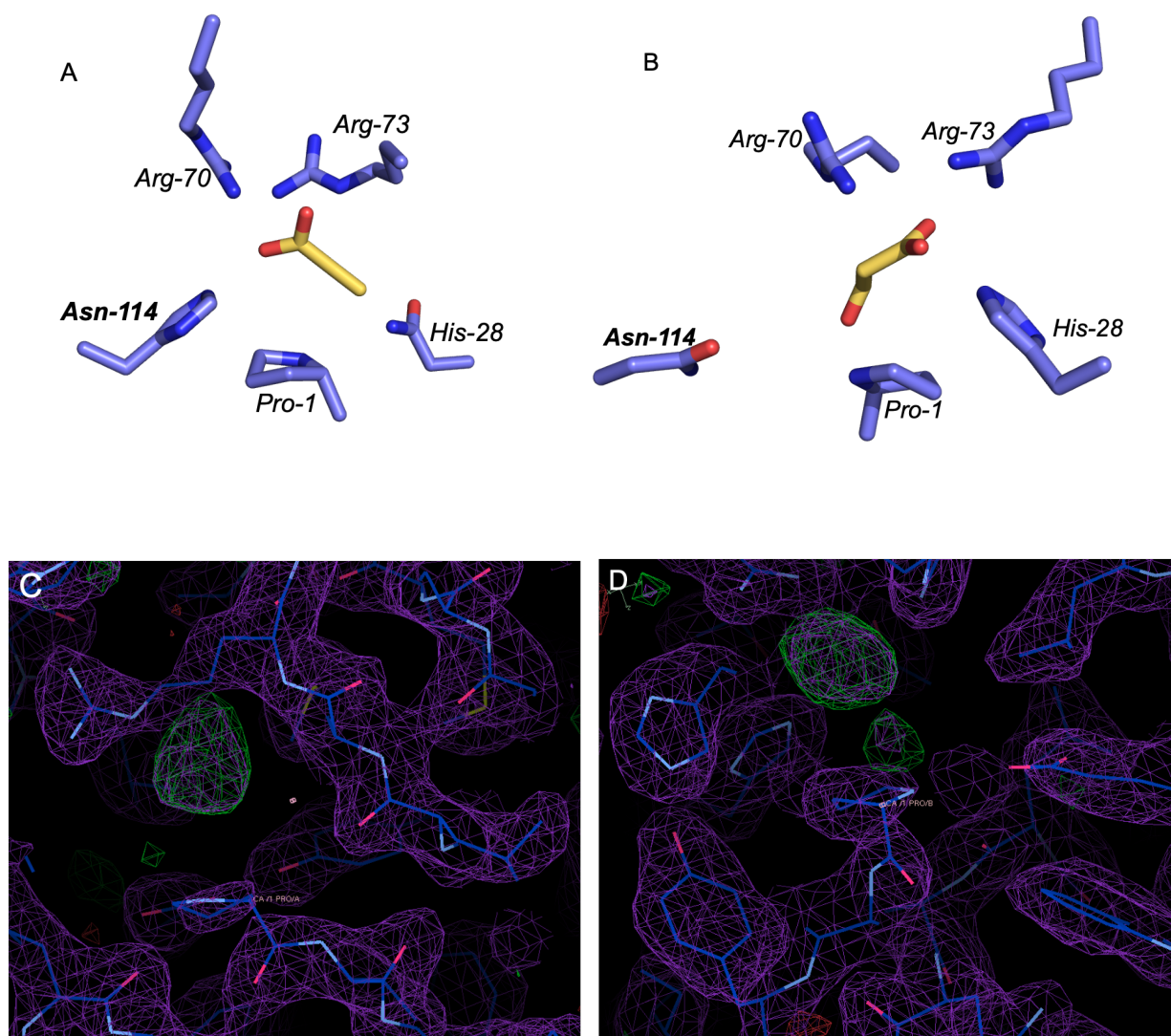
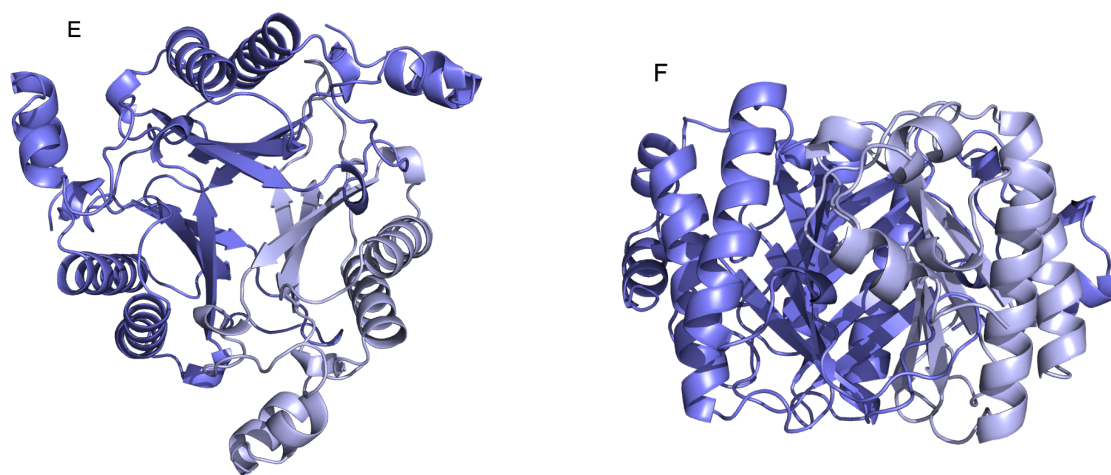


Figure 4.25. Crystal structure of shifted orientation of (A) ACA and (B) MSA product in the active site of *cis*-CaaD (E114N) is shown at 2.3 Å. Electron density is shown before the addition of (C) ACA and (D) MSA. The trimer is shown (E) at a top view with one monomer a lighter shade and (F) at a side view.

Figure 4.25. (cont'd)



4.12.11 X-Ray Crystal Structures of apo, ACA, and CCA Soaked *cis*-CaaD (E114Q)

The x-ray crystal structure of apo *cis*-CaaD (E114Q) was solved to a 1.8 Å resolution and 99% completeness by molecular replacements and refined to R and R_{free} values of 17% and 22%, respectively (Figure 4.26). The asymmetric unit contains one chain. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). Sulfate was seen in the active site of all chains in the asymmetric unit.

The x-ray crystal structure of *cis*-CaaD (E114Q) that was soaked with ACA substrate was solved to a 1.9 Å resolution with 89% completeness by molecular replacements and refined to R and R_{free} values of 21% and 24%, respectively (Figure 4.27). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). The first chain had a covalently attached 3-(*N*-prolyl)-3-hydroxypropionate intermediate, while the other two chains showed malonic semi aldehyde product.

The x-ray crystal structure of *cis*-CaaD (E114Q) that was soaked with CCA substrate was solved to a 2.2 Å resolution and 75% completeness by molecular replacements and refined to R and R_{free} values of 17% and 24%, respectively (Figure 4.28). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). One chain contains ACA, one chain contains acrylate, and one chain contains both covalently linked 3-(*N*-prolyl)-ethene and sulfate.

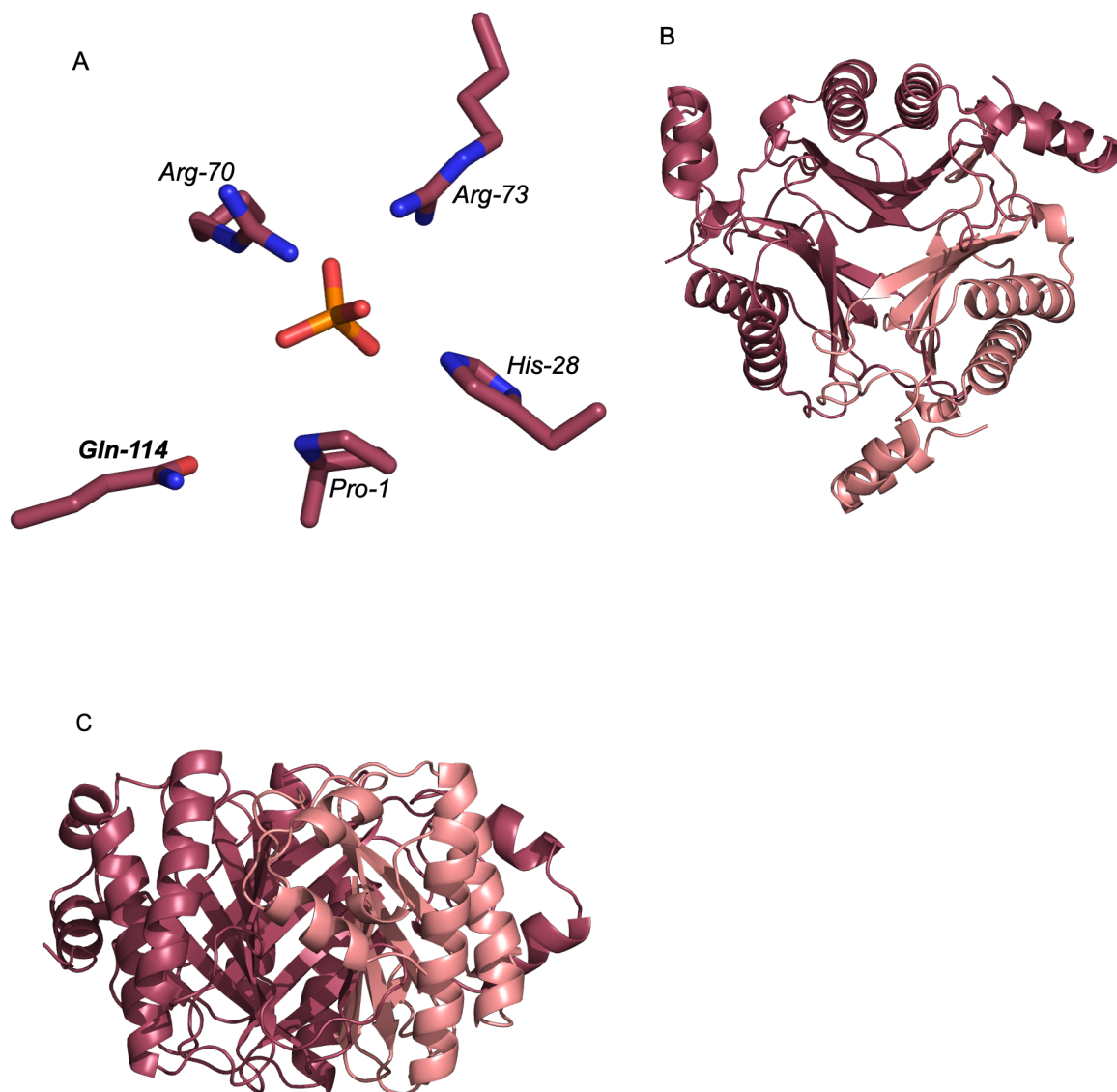


Figure 4.26. Crystal structure of apo *cis*-CaaD (E114Q) shown with (A) phosphate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.

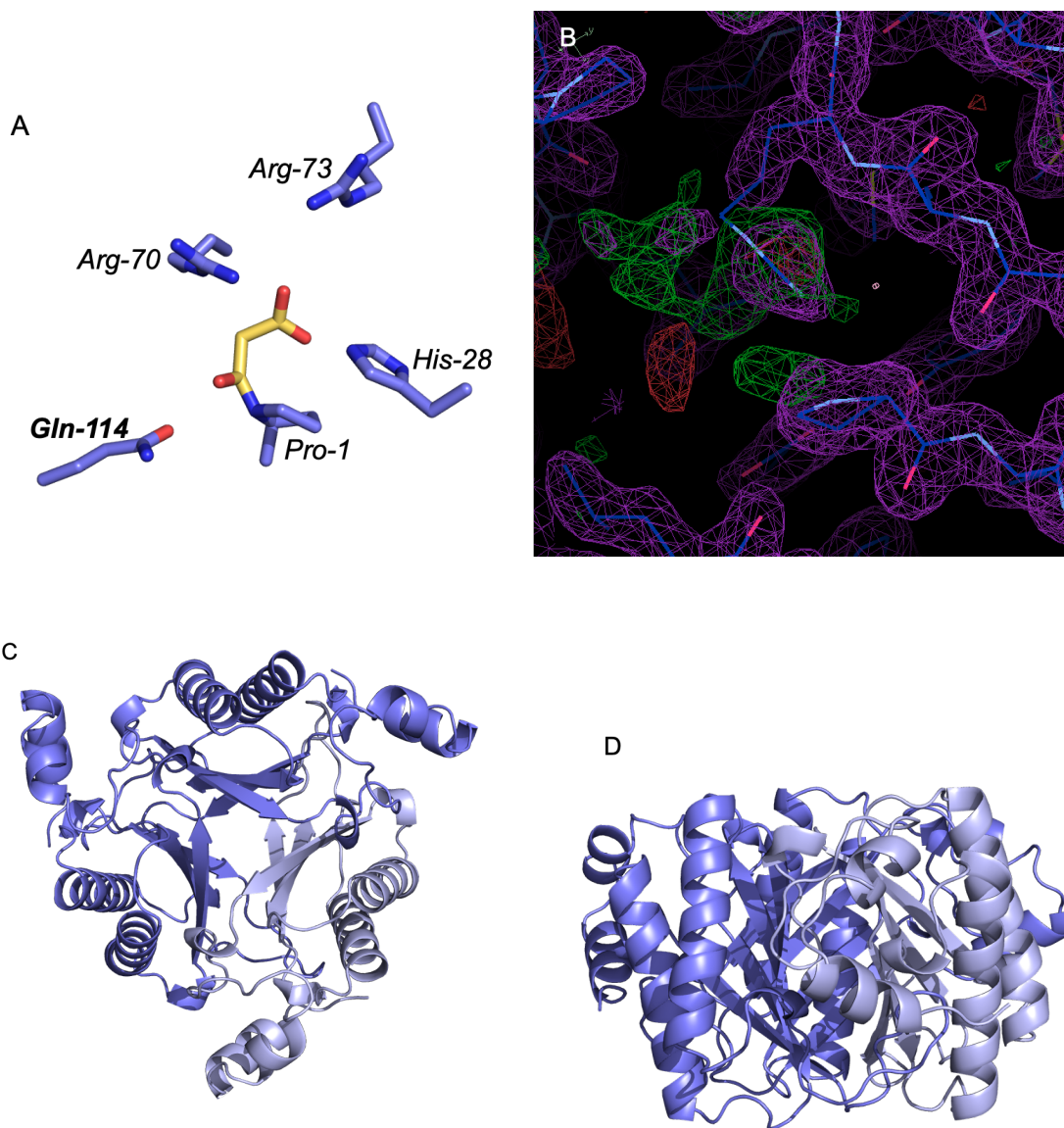


Figure 4.27. Crystal structures of ACA soaked *cis*-CaaD (E114Q) at 1.9 Å crystals showing (A) covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1. Electron density is shown (B) before the addition of 3-(*N*-prolyl)-3-hydroxypropionate. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view.

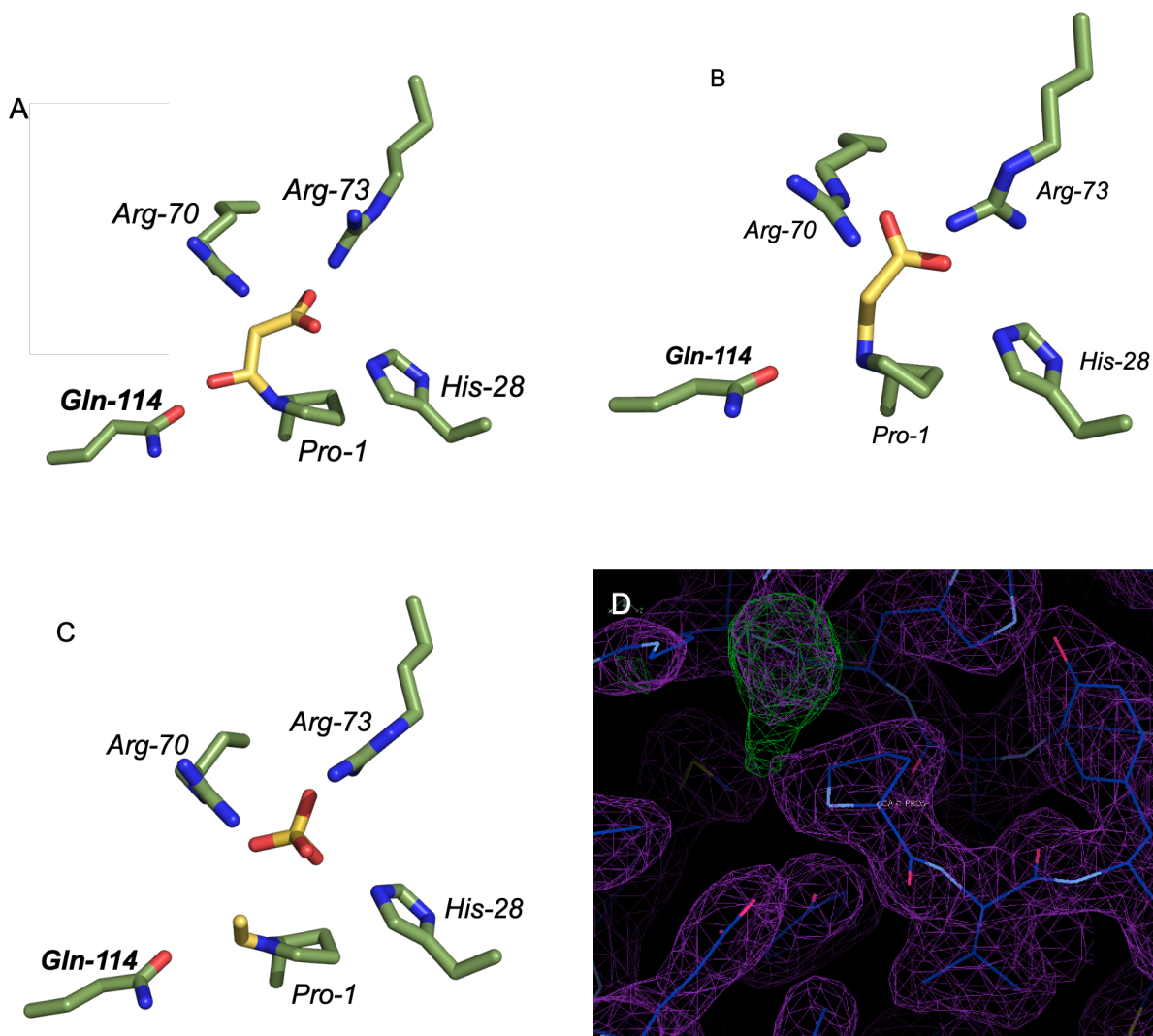
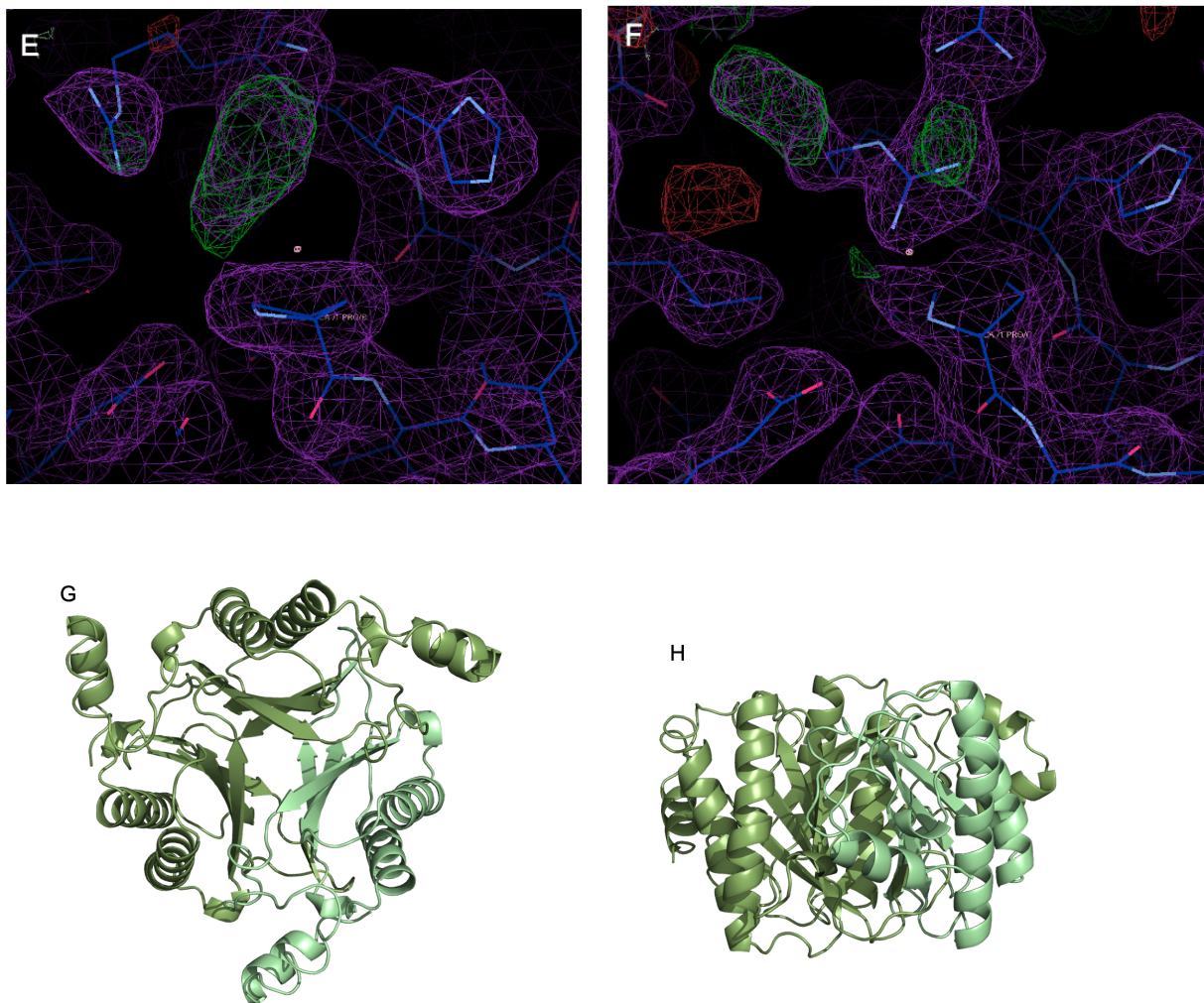


Figure 4.28. Crystal structures of CCA soaked *cis*-CaaD (E114Q) at 2.2 Å crystals showing (A) covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1, (B) covalent intermediate of 3-(*N*-prolyl)-acrylate linked to catalytic Pro-1 and (C) (*N*-prolyl)-ethene covalent intermediate with sulfate. Electron density is shown before the addition of (D) 3-(*N*-prolyl)-3-hydroxypropionate, (E) 3-(*N*-prolyl)-acrylate and (F) (*N*-prolyl)-ethene with sulfate. The trimer is shown (G) at a top view with one monomer a lighter shade and (H) at a side view.

Figure 4.28. (cont'd)



4.12.12 X-Ray Crystal Structures of apo, ACA, and CCA Soaked *cis*-CaaD (Y103F)

The x-ray crystal structure of apo *cis*-CaaD (Y103F) was solved to a 2.2 Å resolution and 85% completeness by molecular replacements and refined to R and R_{free} values of 21% and 27%, respectively (Figure 4.29). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8).

The x-ray crystal structure of *cis*-CaaD (Y103F) that was soaked with ACA substrate was solved to a 2.3 Å resolution and 83% completeness by molecular replacements and refined to R and R_{free} values of 23% and 35%, respectively (Figure 4.30). The asymmetric unit contains one chain. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). Crystallization screening buffer ion sulfate was in the active site of all chains.

The x-ray crystal structure of *cis*-CaaD (Y103F) that was soaked with CCA substrate was solved to a 2.2 Å resolution and 77% completeness by molecular replacements and refined to R and R_{free} values of 19% and 24%, respectively (Figure 4.31). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). All chains contain three waters in the active site. This is evident by electron density.

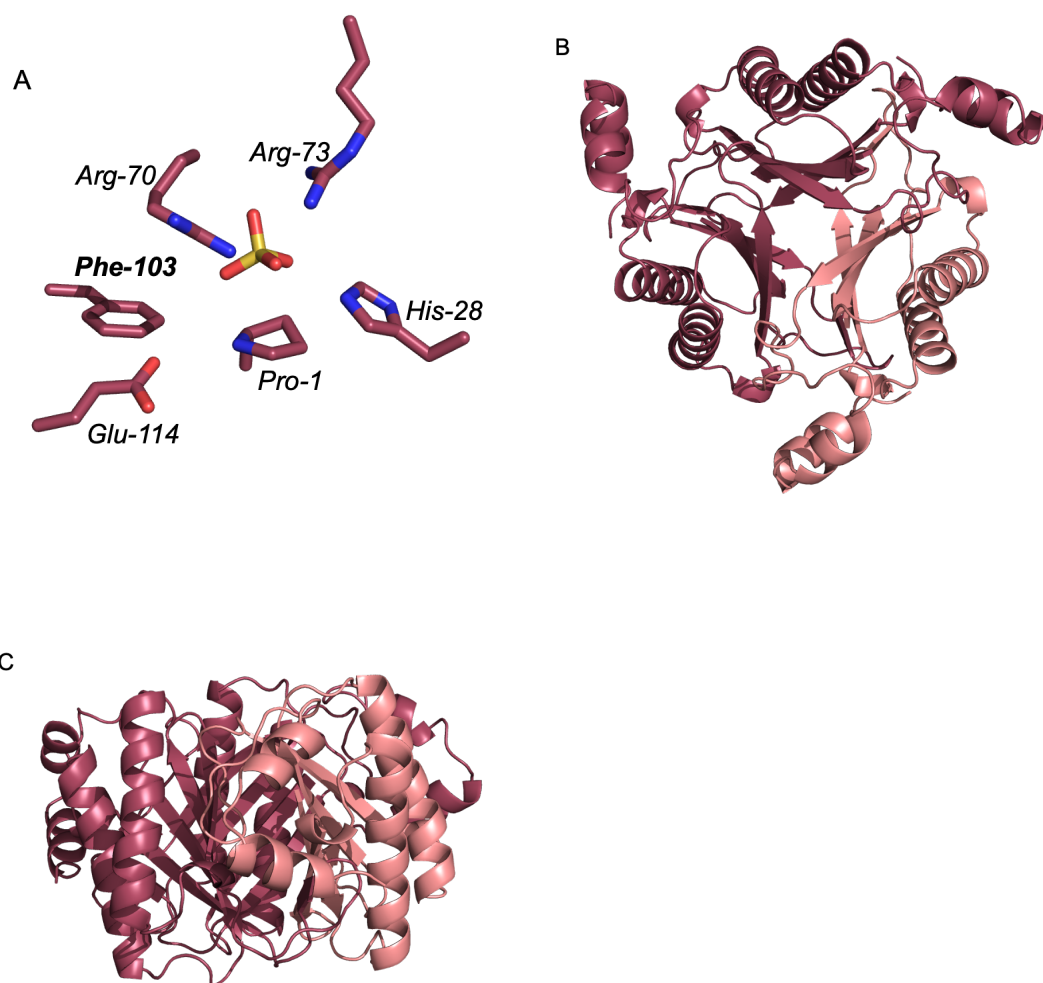


Figure 4.29. Crystal structure of apo *cis*-CaaD (Y103F) is shown with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.

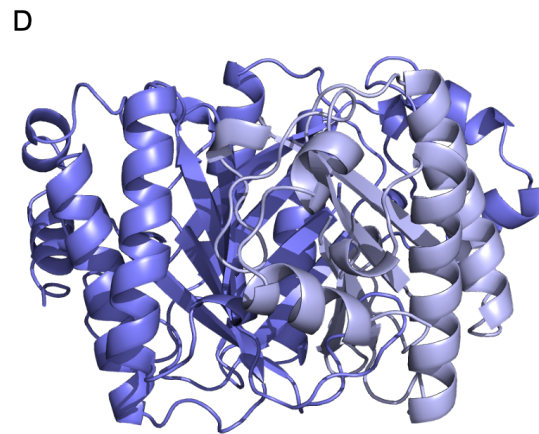
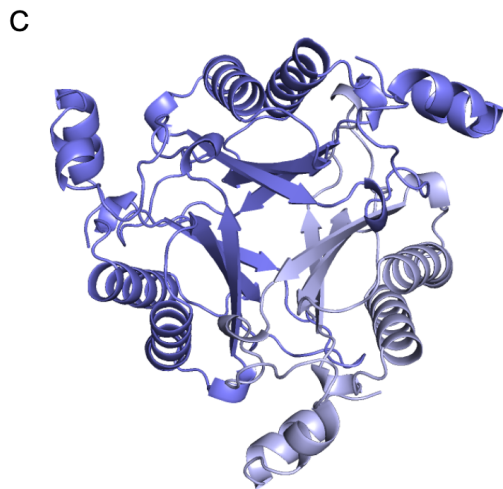
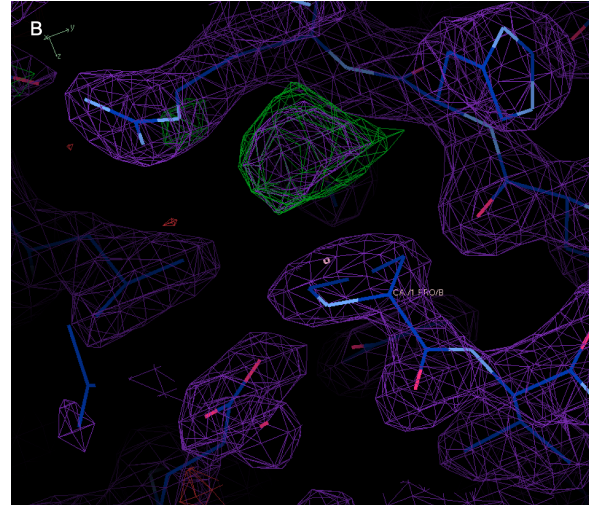
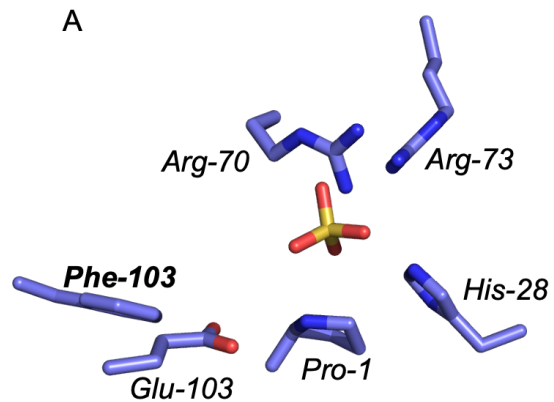


Figure 4.30. Crystal structures of ACA soaked *cis*-CaaD (Y103F) with (A) sulfate in the active site. Electron density is shown (B) before the addition of sulfate. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view.

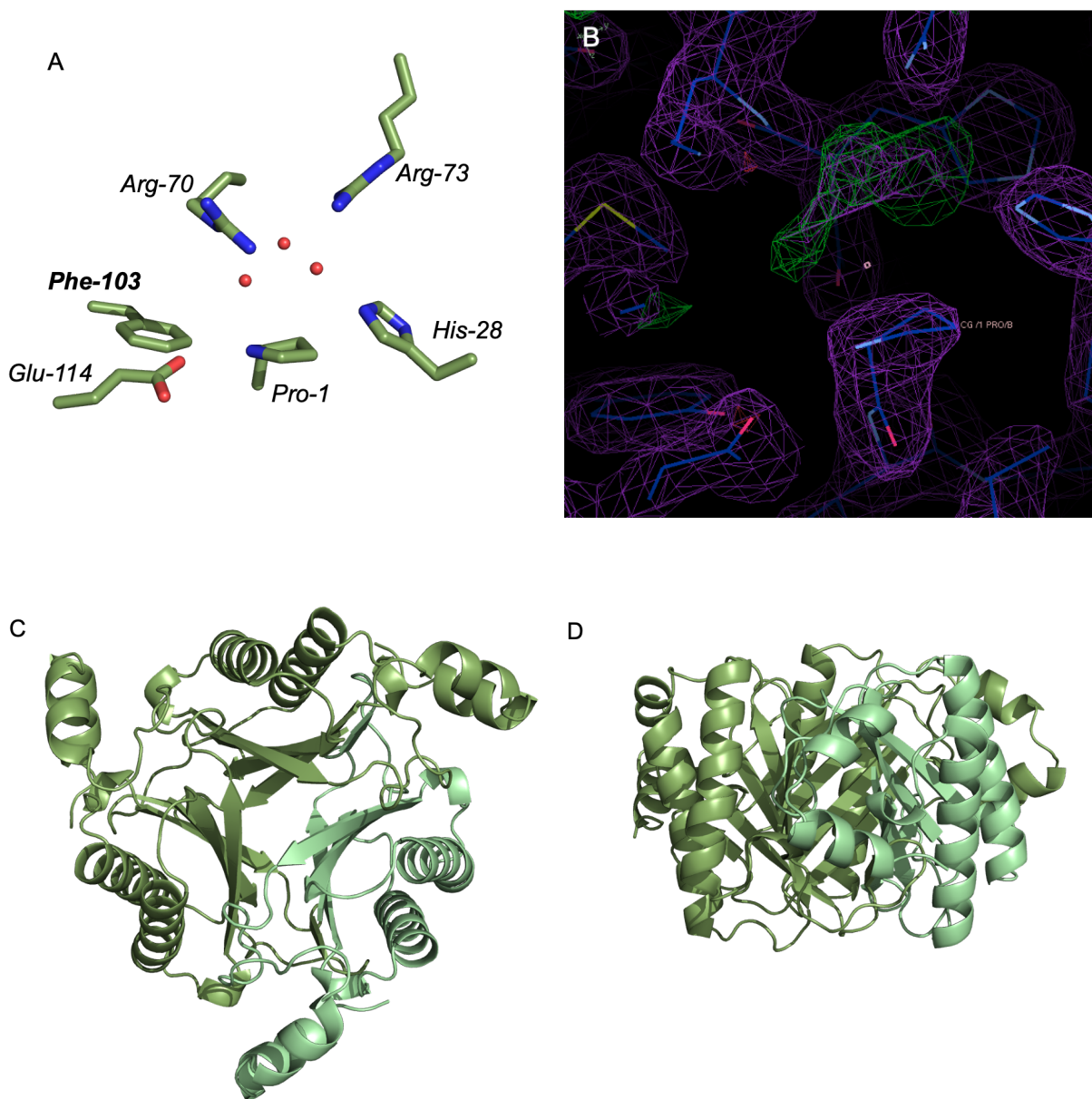


Figure 4.31. Crystal structures of CCA soaked *cis*-CaaD (Y103F) with (A) three waters in the active site. Electron density is shown (B) before the addition of waters. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view.

4.12.13 X-Ray Crystal Structures of ACA soaked *cis*-CaaD (H28A)

The x-ray crystal structure of *cis*-CaaD (H28A) that was soaked with ACA substrate was solved to a 1.8 Å resolution and 76% completeness by molecular replacements and refined to R and R_{free} values of 19% and 23%, respectively (Figure 4.32). The asymmetric unit contains three

chains. All residues from Pro-1 to position 122, 130, and 129 in each of the three chains fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). Two molecules of acetate and one water are in the three active sites of the three chains in the asymmetric unit. The C-terminal was disordered, and the last alpha helices were unable to be built by the electron density.

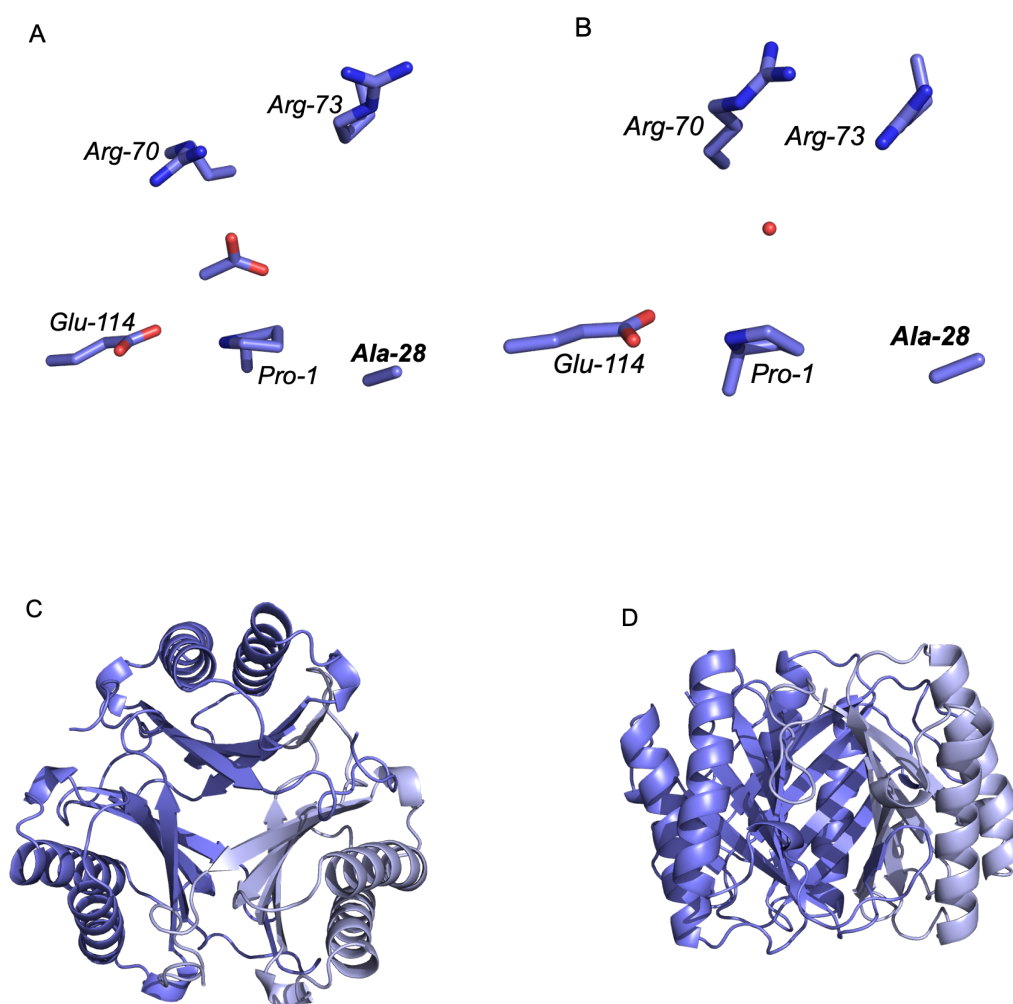


Figure 4.32. Crystal structures of ACA soaked *cis*-CaaD (H28A) with (A) acetate and (B) water in the active site. The trimer is shown (C) at a top view with one monomer a lighter shade and (D) at a side view.

4.12.14 X-Ray Crystal Structures of apo, ACA, and CCA Soaked *cis*-CaaD (T34A)

The x-ray crystal structure of apo *cis*-CaaD (T34A) was solved to a 2.5 Å resolution and 87% completeness by molecular replacements and refined to R and R_{free} values of 20% and 28%, respectively (Figure 4.33). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). All chains in the asymmetric unit contain sulfate from crystallization screen solution in the active site.

The x-ray crystal structure of *cis*-CaaD (T34A) that was soaked with ACA substrate was solved to a 2.3 Å resolution and 83% completeness by molecular replacements and refined to R and R_{free} values of 21% and 28%, respectively (Figure 4.34). The asymmetric unit contains three chains. All residues from Pro-1 to position 146 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). All chains in the asymmetric unit contain sulfate from crystallization screen solution in the active site.

The x-ray crystal structure of *cis*-CaaD (T34A) that was soaked with CCA substrate was solved to a 2.4 Å resolution and 70% completeness by molecular replacements and refined to R and R_{free} values of 21% and 31%, respectively (Figure 4.35). The asymmetric unit contains three chains. All residues from Pro-1 to position 145 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). All chains in the asymmetric unit contain sulfate from crystallization screen solution in the active site.

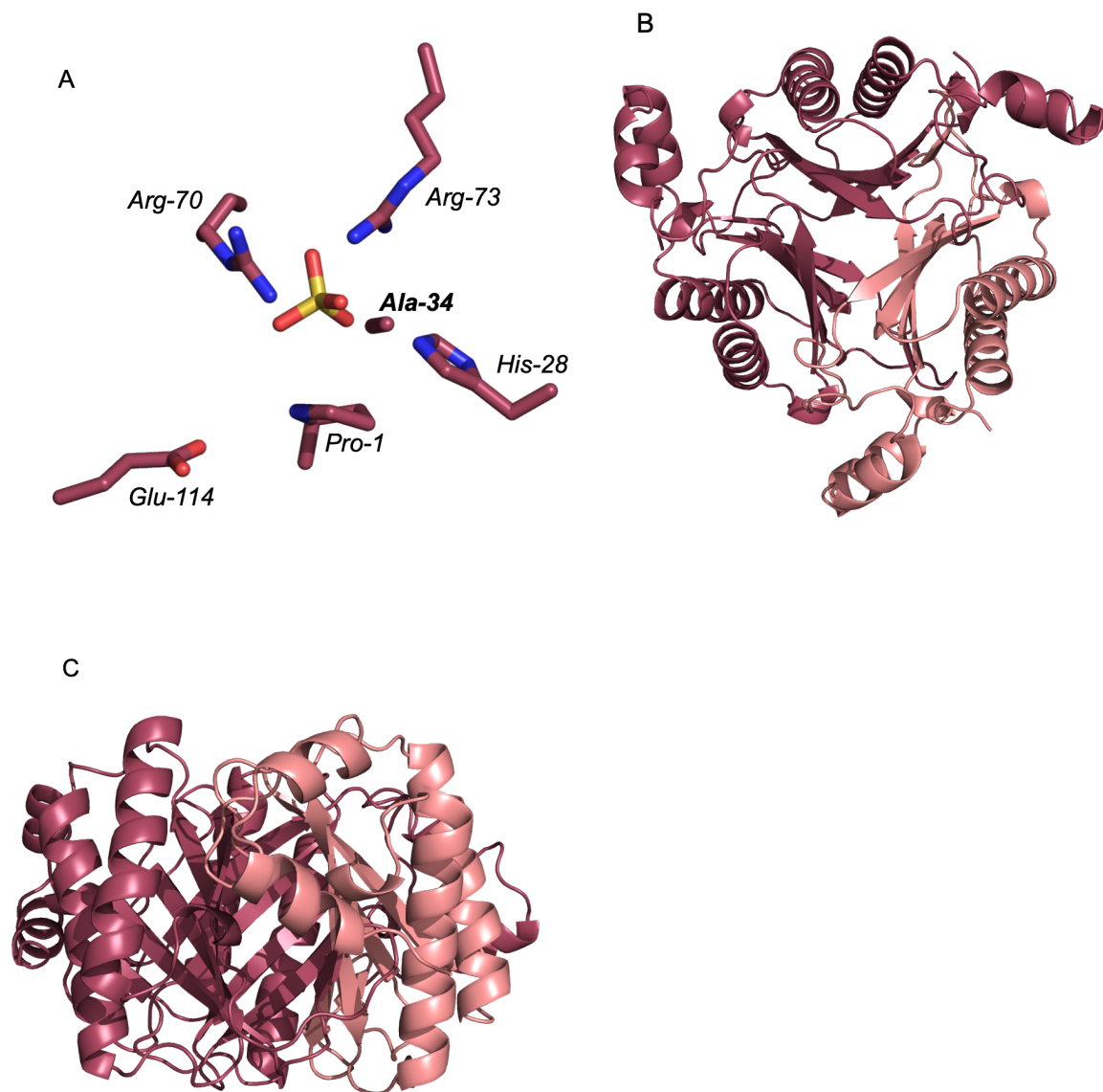


Figure 4.33. Crystal structure of apo *cis*-CaaD (T34A) with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.

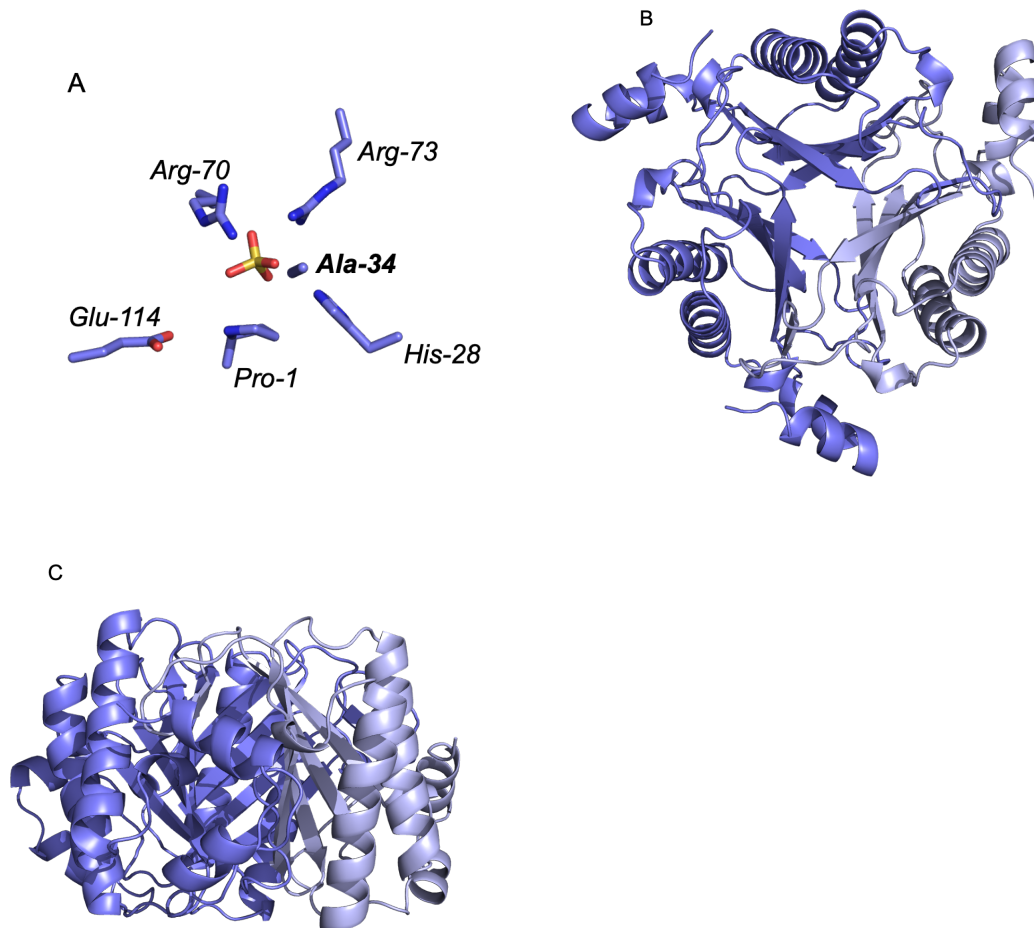


Figure 4.34. Crystal structures of ACA soaked *cis*-CaaD (T34A) shown with (A) sulfate in the active site (A). The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.

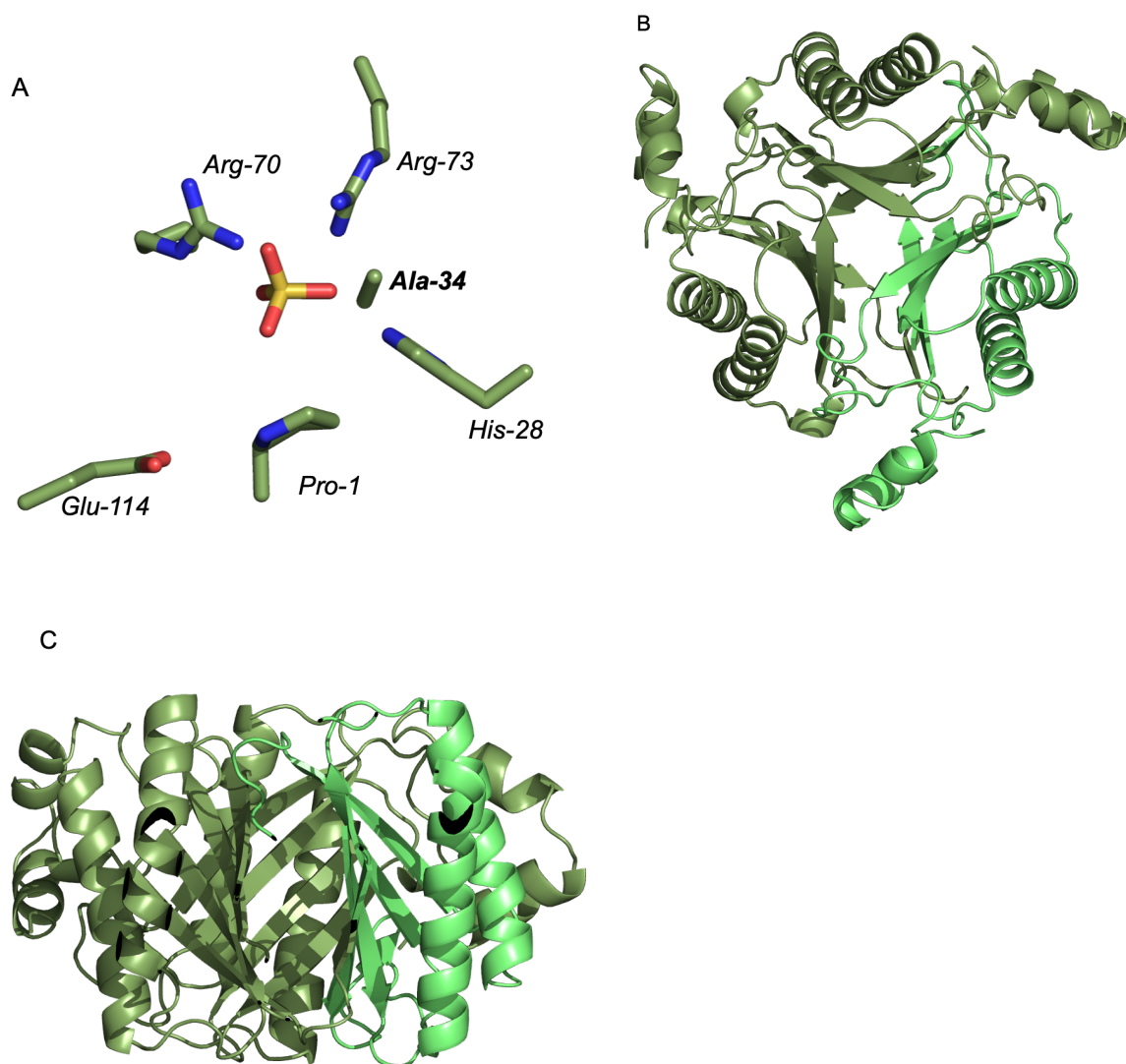


Figure 4.35. Crystal structures of CCA soaked *cis*-CaaD (T34A) shown with (A) sulfate in the active site. The trimer is shown (B) at a top view with one monomer a lighter shade and (C) at a side view.

4.12.15 X-Ray Crystal Structures of apo Cg10062 (E114N)

The x-ray crystal structure of apo Cg10062 (E114N) was solved to a 2.5 Å resolution and 66% completeness by molecular replacements and refined to R and R_{free} values of 21% and 31%, respectively (Figure 4.36). There are six chains in the asymmetric unit consisting of two complete trimers. All residues from Pro-1 to position 147 fit within the electron density with remaining two

residues of the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of Cg10062 (PDB: 3N4G).

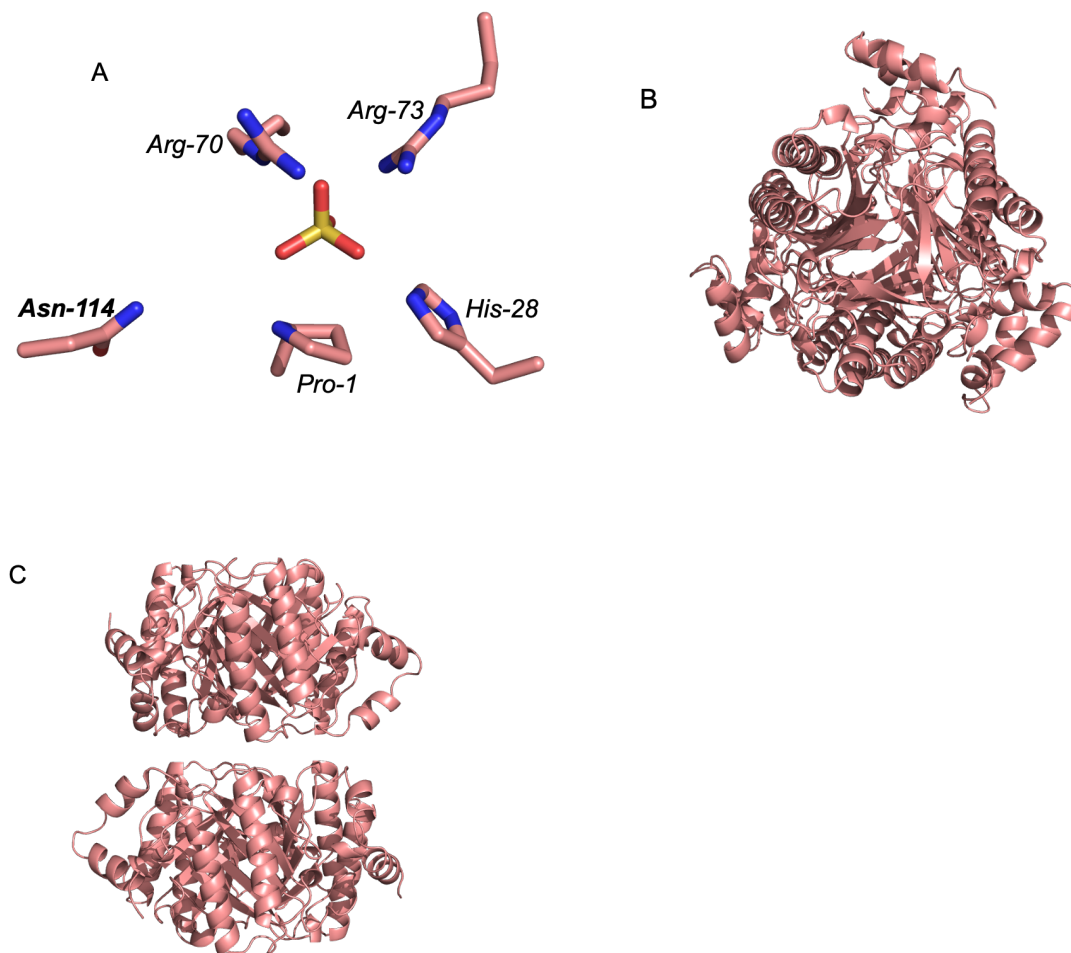


Figure 4.36. Crystal structure of apo Cg10062 (E114N) with (A) sulfate in the active site. The two trimers are shown (B) at a top view and (C) at a side view.

4.12.16 X-Ray Crystal Structures of ACA Soaked Cg10062 (E114N)

Two structures of ACA soaked Cg10062 (E114N) were obtained and used for evaluation. The first x-ray crystal structure of ACA-soaked Cg10062 (E114N) was solved to a 2.9 Å resolution in the P4132 space group with 89% completeness by molecular replacements and refined to R and R_{free} values of 25% and 31%, respectively (Figure 6.37A-C). There was only one chain in the asymmetric unit with unreacted ACA substrate trapped in the active site. All residues from Pro-1

to position 146 fit within the electron density with remaining residues of the C-terminal not accounted for by electron density. Two conformers fit the electron density for Arg-70. The structure was resolved based on the previously published structure of Cg10062 (PDB: 3N4G). The second crystal structure of ACA-soaked Cg10062 (E114N) was solved to a 2.6 Å resolution in the C121 space group with 67% completeness by molecular replacements and refined to R and R_{free} values of 19% and 29%, respectively (Figure 6.37D-G). There were six chains in the asymmetric unit containing trapped ACR and HPA intermediates in two active sites with sulfate in the remaining four active sites. All residues from Pro-1 to position 147 fit within the electron density with remaining residues of the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of Cg10062 (PDB: 3N4G).

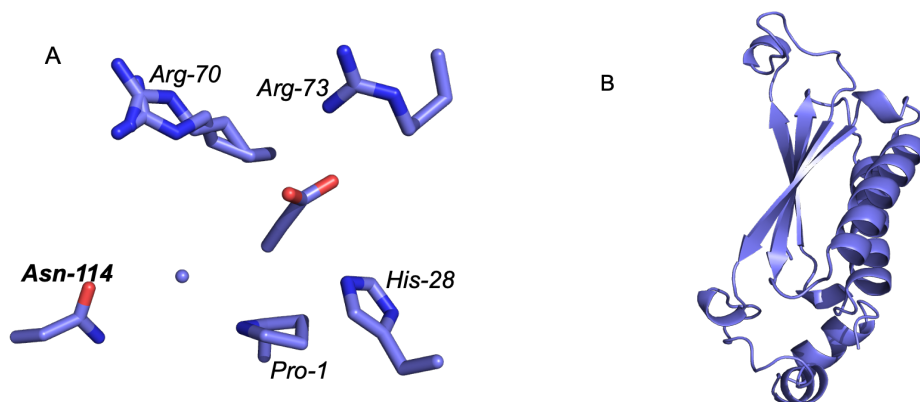
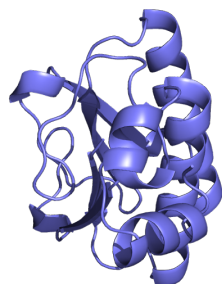


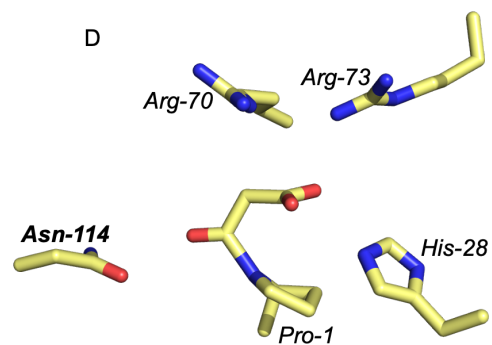
Figure 4.37. Crystal structure of the first obtained structure of ACA-soaked Cg10062 (E114N) with (A) ACA in the active site. The single chain shown (B) at a top view and (C) at a side view. The second crystal structure of ACA-soaked Cg10062 (E114N) shows (D) a covalent intermediate of 3-(*N*-prolyl)-3-hydroxypropionate linked to catalytic Pro-1 and (E) a covalent intermediate of 3-(*N*-prolyl)-acrylate linked to catalytic Pro-1. The two trimers are shown (F) at atop view and (G) at a side view.

Figure 4.37. (cont'd)

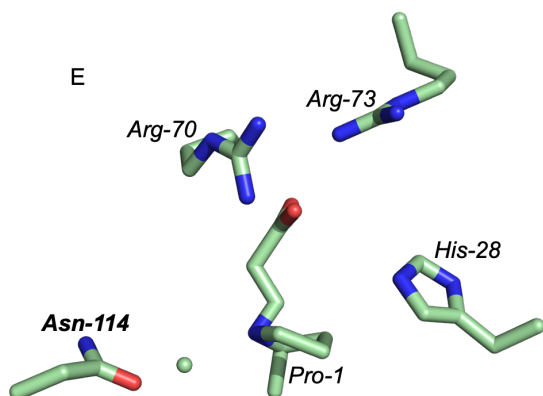
C



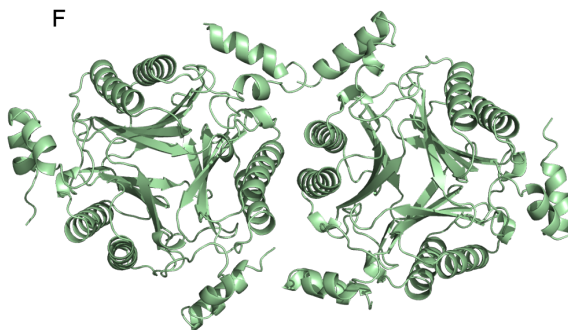
D



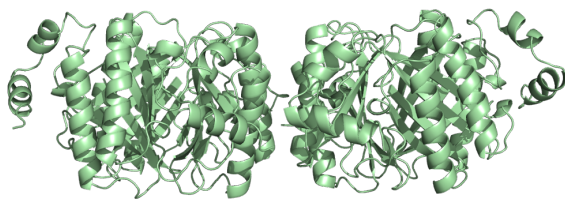
E



F



G



4.12.17 X-Ray Crystal Structure of *tcis*-CaaD

The x-ray crystal structure of *tcis*-CaaD that was soaked with CCA substrate was solved to a 2.5 Å resolution and 87% completeness by molecular replacements and refined to R and R_{free} values of 21% and 30%, respectively (Figure 6.38). There is one chain in the asymmetric unit. All residues from Pro-1 to position 119 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). Sulfate is seen trapped in the active site of the asymmetric unit.

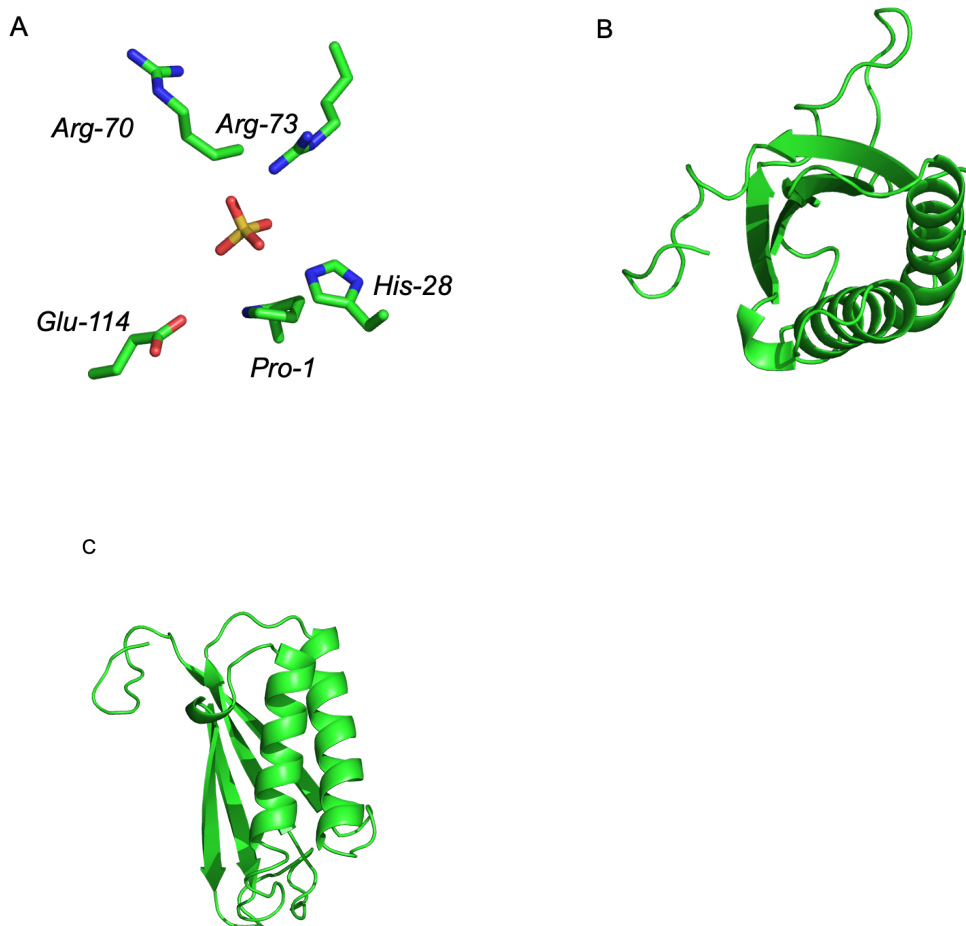


Figure 4.38. The crystal structure of truncated *cis*-CaaD has a (A) sulfate in the active site. The single chain shown (B) at atop view and (C) at a side view.

4.12.18 X-Ray Crystal Structure of *N-cis*-CaaD/C-Cg10062

The x-ray crystal structure of *N-cis*-CaaD/C-Cg10062 that was soaked with ACA substrate was solved to a 2.7 Å resolution and 74% completeness by molecular replacements and refined to R and R_{free} values of 24% and 34%, respectively (Figure 4.39). The asymmetric unit contains only one chain. All residues from Pro-1 to position 123 fit within the electron density with remaining residues to the C-terminal not accounted for by electron density. The structure was resolved based on the previously published structure of *cis*-CaaD (PDB: 3MF8). A chloride from crystallization screen solution is seen in the active site of the asymmetric unit.

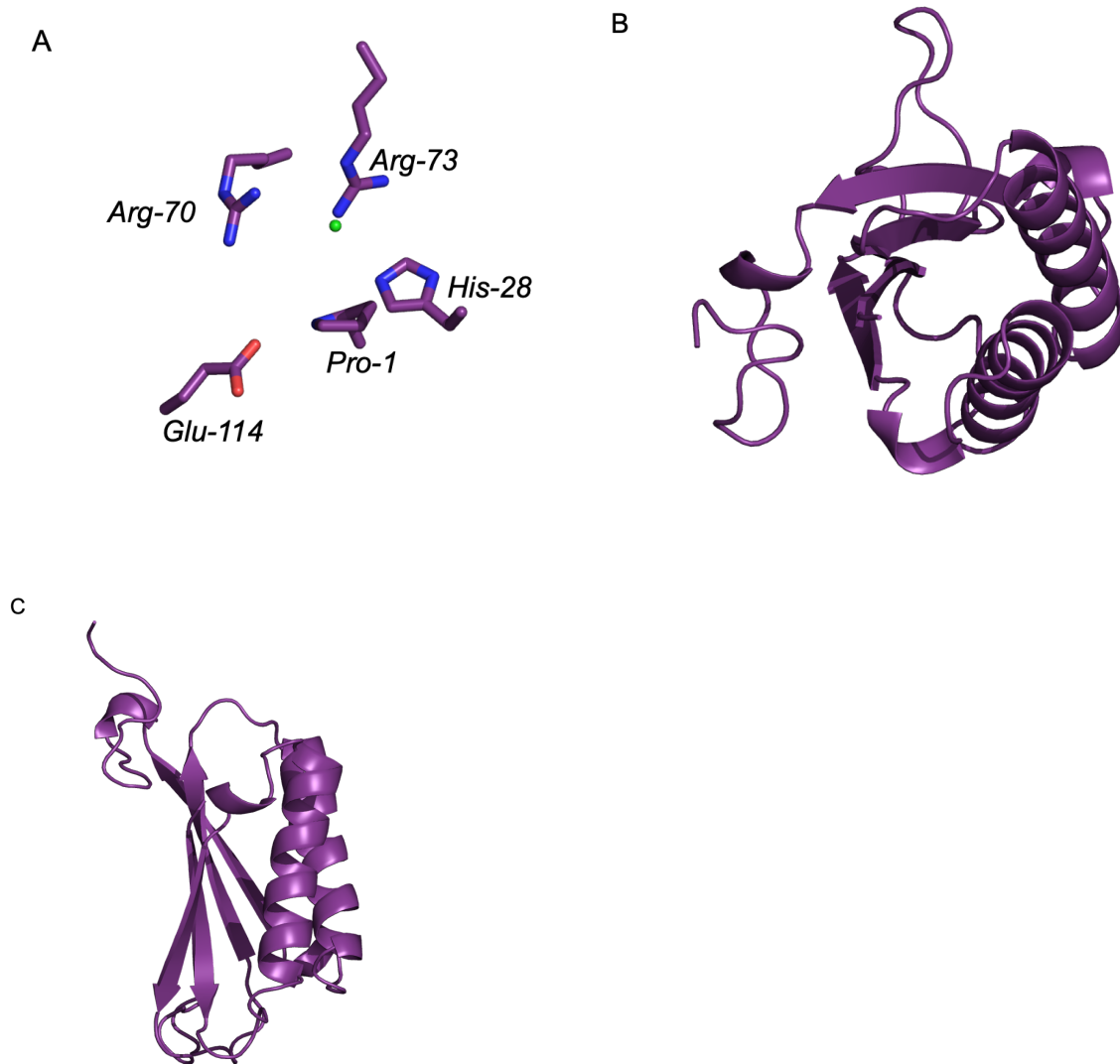


Figure 4.39. *N-cis-CaaD/C-Cg10062* shows a (A) chloride ion trapped in the active site. The single chain shown (B) at atop view and (C) at a side view.

Table 4.12. Crystallographic soaking parameters for wild-type and variants of *cis*-CaaD and Cg10062 (E114N).

Enzyme	Soak	Reservoir (well)	Soak Time	Temp (°C)	Cryoprotectant
<i>cis</i> -CaaD (wild-type)	CCA	0.18 M ammonium acetate 16% PEG 3,352	1 min	23	Paratone
<i>cis</i> -CaaD (wild-type)	ACA	0.2 M ammonium tartrate dibasic, 20% w/v PEG 3,353	6 sec	23	Paratone
<i>cis</i> -CaaD (E114D)	Apo	0.2 M ammonium sulfate 20% w/v PEG 3,351	-	23	Paratone
<i>cis</i> -CaaD (E114D)	ACA	0.2 M ammonium sulfate 20% w/v PEG 3,360	15 min	4	Paratone
<i>cis</i> -CaaD (E114D)	CCA	0.2 M ammonium sulfate 20% w/v PEG 3,353	6 min	23	Paratone
<i>cis</i> -CaaD (E114N)	Apo	(0.03 M citric acid, 0.07 M BIS-TRIS propane) pH 7.6 20% w/v PEG 3,350	-	23	Paratone
<i>cis</i> -CaaD (E114N)	ACA	(0.05 M citric acid, 0.05 M BIS-TRIS propane) pH 6 17% w/v PEG 3,355	25 min	23	Paratone
<i>cis</i> -CaaD (E114Q)	Apo	(0.05 M citric acid, 0.05 M BIS-TRIS propane) pH 5 20% w/v PEG 3,354	-	4	Paratone
<i>cis</i> -CaaD (E114Q)	ACA	(0.05 M citric acid, 0.05 M BIS-TRIS propane) pH 5 20% w/v PEG 3,350	15 min	4	Paratone
<i>cis</i> -CaaD (E114Q)	CCA	(0.05 M citric acid, 0.05 M BIS-TRIS propane) pH 5 20% w/v PEG 3,352	25 min	4	Paratone
<i>cis</i> -CaaD (Y103F)	Apo	0.2 M sodium fluoride 20% w/v PEG 3,350	-	23	27% Glycerol in Screen Solution
<i>cis</i> -CaaD (Y103F)	ACA	0.2 M magnesium sulfate heptahydrate 20% w/v PEG 3,350	25 min	4	24% Glycerol in Screen Solution
<i>cis</i> -CaaD (Y103F)	CCA	0.2 M sodium fluoride 20% w/v PEG 3,350	10 min	23	27% Glycerol in Screen Solution
<i>cis</i> -CaaD (H28A)	ACA	0.2 M ammonium acetate 20% w/v PEG 3,350	7 min	23	27% Glycerol in Screen Solution
<i>cis</i> -CaaD (T34A)	Apo	0.2 M ammonium sulfate 20% w/v PEG 3,350	-	4	Paratone
<i>cis</i> -CaaD (T34A)	ACA	0.2 M ammonium sulfate 20% w/v PEG 3,350	15 min	4	Paratone
<i>cis</i> -CaaD (T34A)	CCA	0.2 M ammonium sulfate 20% w/v PEG 3,350	20 min	4	Paratone
<i>N-cis</i> -CaaD/C-g10062	ACA	0.2 M lithium chloride 20% w/v PEG 3,350	15 min	23	27% Glycerol in Screen Solution
<i>tcis</i> -CaaD	CCA	0.2 M ammonium sulfate 20% w/v PEG 3,350	14 min	23	27% Glycerol in Screen Solution

Table 4.12. (cont'd)

Cg10062 (E114N)	Apo	0.2 M Ammonium sulfate 20% w/v PEG 3,350	-	4	27% Glycerol in Screen Solution
Cg10062 (E114N)	ACA	2% v/v Tacsimate™ pH 5.0 5% w/v PEG 3,350	1 min	23	Paratone
Cg10062 (E114N)	ACA	0.2 M Ammonium sulfate 20% w/v PEG 3,350	25 min	4	27% Glycerol in Screen Solution

Table 4.13. Crystallographic data collection and structure refinement statistics. ^aApo and ^bACA soaked, ^cCCA soaked. ACA, acetylenecarboxylic acid. HPA, 3-(*N*-prolyl)-3-hydroxypropionate. ACR, 3-(*N*-prolyl)-acrylate. ETY, (*N*-prolyl)-ethene. MSA, malonic semialdehyde. ACT, acetate.

	<i>cis</i> -CaaD (wild-type) ^b	<i>cis</i> -CaaD (wild-type) ^c
	(ACA)	(ACT)
PDB ID	9NFF	9NFH
Data statistics		
Space group	C 2 2 2 ₁	P 6 3 1
No. of chains/ASU	3	1
Unit cell dimensions (Å, °)	a = 60.027	a = 59.964
	b = 100.296	b = 59.964
	c = 142.954	c = 58.565
	α = 90, β = 90, γ = 90	α = 90, β = 90, γ = 120
Total reflections	514078	563948
Resolution range (Å) (outer shell)	50.00 – 2.00	50.00 – 1.32
	(2.03-2.00)	(1.34-1.32)
Unique reflections (outer shell)	29776 (1474)	27976 (1052)
Completeness (outer shell)	100.0 (99.9)	99.8 (97.8)
Average I/σ	16.76	32.48
Refinement Statistics		
Reflections (outer shell)	28706 (1461)	27781 (1910)
Reflections (R-free) (outer shell)	1993 (97)	1981 (134)
R-work (outer shell)	0.1958 (0.2351)	0.1900 (0.2894)

Table 4.13. (cont'd)

R-free (outer shell)	0.2615 (0.3196)	0.2114 (0.3136)
Structure Statistics		
Non-hydrogen atoms	3695	1009
Solvent atoms	246	82
Ligand atoms	15	4
Protein residues	435	118
RMS (Bonds) (Å)	0.008	0.007
RMS (angles) (°)	1.05	1.33
Ramachandran plot (%) favored, allowed, outliers	98.37, 1.63, 0.00	98.28, 0.86, 0.86
Rotamer outliers (%)	0.85	1.06
B-factor, Average	26.45	19.65
Macromolecules, Ligands, Solvent	26.19, -, 30.10	19.01, 20.18, 26.83
<hr/>		
	<i>cis</i>-CaaD (E114D)^a	<i>cis</i>-CaaD (E114D)^b
	(APO)	(ACA, ACR)
PDB ID	9NF1	9NER
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 59.997	a = 60.256
	b = 100.294	b = 98.828
	c = 147.226	c = 142.409
	α = 90, β = 90, γ = 90	α = 90, β = 90, γ = 90

Table 4.13. (cont'd)

Total reflections	5056410	1281563
Resolution range (Å) (outer shell)	50.00 – 1.60	50.00 – 2.50
	(1.63-1.60)	(2.54-2.50)
Unique reflections (outer shell)	52366 (2769)	14718 (691)
Completeness (outer shell)	88.9 (95.0)	98.1 (96.6)
Average I/ σ	26.63	13.78

Refinement Statistics

Reflections (outer shell)	51555 (3461)	14196 (1066)
Reflections (R-free) (outer shell)	1993 (130)	1428 (110)
R-work (outer shell)	0.1755 (0.2403)	0.1859 (0.2626)
R-free (outer shell)	0.2189 (0.3105)	0.2519 (0.3638)

Structure Statistics

Non-hydrogen atoms	4026	3534
Solvent atoms	555	96
Ligand atoms	15	15
Protein residues	439	435
RMS (Bonds) (Å)	0.007	0.007
RMS (angles) (°)	0.97	0.83
Ramachandran plot (%) favored, allowed, outliers	97.69, 1.85, 0.46	97.20, 2.80, 0.00
Rotamer outliers (%)	0.56	2.54
B-factor, Average	18.66	37.56

Table 4.13. (cont'd)

Macromolecules, Ligands,	17.13, 23.36, 28.06	37.62, 40.12, 35.34
Solvent		
	<i>cis</i> -CaaD (E114D) ^c	<i>cis</i> -CaaD (E114N) ^a
	(ETY, SO4)	(APO)
PDB ID	9NF3	9NF9
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 59.534	a = 58.222
	b = 99.939	b = 99.02
	c = 142.794	c = 137.557
	α = 90, β = 90, γ = 90	α = 90, β = 90, γ = 90
Total reflections	5473420	2678303
Resolution range (Å) (outer shell)	50.00 – 1.80	50.00 – 2.50
	(1.83-1.80)	(2.54-2.50)
Unique reflections (outer shell)	36388 (1251)	13174 (664)
Completeness (outer shell)	91.5 (64.3)	94.3 (100.0)
Average I/σ	13.3	12.53
Refinement Statistics		
Reflections (outer shell)	36018 (1807)	12430 (1111)
Reflections (R-free) (outer shell)	1973 (104)	1265 (119)

Table 4.13. (cont'd)

R-work (outer shell)	0.1809 (0.2863)	0.2150 (0.2447)
R-free (outer shell)	0.2237 (0.3666)	0.2901 (0.3444)

Structure Statistics

Non-hydrogen atoms	3866	3490
Solvent atoms	407	57
Ligand atoms	12	15
Protein residues	437	435
RMS (Bonds) (Å)	0.007	0.014
RMS (angles) (°)	1.00	1.67
Ramachandran plot (%) favored, allowed, outliers	98.37, 1.16, 0.47	95.10, 3.03, 1.86
Rotamer outliers (%)	0.84	3.97
B-factor, Average	27.72	38.17
Macromolecules, Ligands, Solvent	26.83, 32.38, 35.14	38.18, 39.89, 37.09

	<i>cis</i> -CaaD (E114N) ^b (ACA, MSA)	<i>cis</i> -CaaD (E114Q) ^a (APO)
PDB ID	9NFD	9NGL
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 60.405 b = 100.972 c = 148.185	a = 59.977 b = 101.078 c = 144.071

Table 4.13. (cont'd)

	$\alpha = 90, \beta = 90, \gamma = 90$	$\alpha = 90, \beta = 90, \gamma = 90$
Total reflections	1849912	2576016
Resolution range (Å) (outer shell)	40.00 – 2.30 (2.34-2.30)	50.00 – 1.80 (1.83-1.80)
Unique reflections (outer shell)	17606 (1024)	40869 (2024)
Completeness (outer shell)	85.5 (99.9)	99.5 (99.4)
Average I/ σ	9.67	27.89
Refinement Statistics		
Reflections (outer shell)	17140 (1478)	40586 (2805)
Reflections (R-free) (outer shell)	1730 (136)	1992 (143)
R-work (outer shell)	0.2037 (0.2268)	0.1735 (0.2514)
R-free (outer shell)	0.2759 (0.3221)	0.2153 (0.3219)
Structure Statistics		
Non-hydrogen atoms	3594	3855
Solvent atoms	154	387
Ligand atoms	5	15
Protein residues	437	436
RMS (Bonds) (Å)	0.008	0.008
RMS (angles) (°)	0.92	1.00
Ramachandran plot (%) favored, allowed, outliers	96.97, 2.56, 0.47	98.37, 1.40, 0.23
Rotamer outliers (%)	4.52	1.12

Table 4.13. (cont'd)

B-factor, Average	35.10	20.40
Macromolecules, Ligands,	35.03, -, 36.55	19.37, 24.17, 29.38
Solvent		
<hr/>		
	<i>cis</i> -CaaD (E114Q) ^b	<i>cis</i> -CaaD (E114Q) ^c
	(HPA)	(ACR, HPA, ETY)
PDB ID	9O2O	9O2P
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 60.374	a = 60.608
	b = 100.420	b = 101.228
	c = 147.177	c = 148.09
	α = 90, β = 90, γ = 90	α = 90, β = 90, γ = 90
Total reflections	1612303	21779
Resolution range (Å) (outer shell)	50.00 – 1.93	50.00 – 2.20
	(1.96-1.93)	(2.24-2.20)
Unique reflections (outer shell)	30885 (1638)	21779 (964)
Completeness (outer shell)	90.0 (94.1)	92.4 (83.4)
Average I/σ	19.59	19.14
Refinement Statistics		
Reflections (outer shell)	30472 (1930)	21185 (771)
Reflections (R-free) (outer shell)	2039 (134)	1970 (74)

Table 4.13. (cont'd)

R-work (outer shell)	0.2144 (0.2546)	0.1724 (0.2897)
R-free (outer shell)	0.2380 (0.2849)	0.2362 (0.3791)
Structure Statistics		
Non-hydrogen atoms	3677	3587
Solvent atoms	215	143
Ligand atoms	18	18
Protein residues	436	435
RMS (Bonds) (Å)	0.007	0.009
RMS (angles) (°)	0.86	0.96
Ramachandran plot (%) favored, allowed, outliers	98.84, 0.93, 0.23	97.44, 2.10, 0.47
Rotamer outliers (%)	0.84	0.47
B-factor, Average	20.84	32.13
Macromolecules, Ligands, Solvent	20.49, 24.03, 26.21	32.06, 39.76, 32.98
<hr/>		
	<i>cis</i>-CaaD (Y103F)^a	<i>cis</i>-CaaD (Y103F)^b
	(APO)	(SO₄)
PDB ID	9NG4	9NG5
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 60.333	a = 60.259
	b = 101.281	b = 101.341

Table 4.13. (cont'd)

	c = 146.953	c = 137.573
	$\alpha = 90, \beta = 90, \gamma = 90$	$\alpha = 90, \beta = 90, \gamma = 90$
Total reflections	3738870	2634634
Resolution range (Å) (outer shell)	50.00 – 2.20	50.00 – 2.30
	(2.24-2.20)	(2.34-2.30)
Unique reflections (outer shell)	21107 (1056)	17389 (917)
Completeness (outer shell)	89.9 (93.0)	89.2 (96.3)
Average I/ σ	9.4	7.13
Refinement Statistics		
Reflections (outer shell)	20059 (1051)	16171 (1224)
Reflections (R-free) (outer shell)	1976 (103)	1605 (119)
R-work (outer shell)	0.2128 (0.2387)	0.2343 (0.2931)
R-free (outer shell)	0.2742 (0.3006)	0.3538 (0.4335)
Structure Statistics		
Non-hydrogen atoms	3537	3519
Solvent atoms	99	81
Ligand atoms	15	15
Protein residues	435	435
RMS (Bonds) (Å)	0.009	0.008
RMS (angles) (°)	0.97	0.95
Ramachandran plot (%) favored, allowed, outliers	97.90, 1.63, 0.47	95.57, 3.50, 0.93

Table 4.13. (cont'd)

Rotamer outliers (%)	2.26	5.37
B-factor, Average	32.67	29.48
Macromolecules, Ligands,	32.53, 47.41, 35.40	29.49, 27.19, 29.37
Solvent		
<hr/>		
	<i>cis</i> -CaaD (Y103F) ^c	<i>cis</i> -CaaD (H28A) ^b
	(H ₂ O)	(ACT)
PDB ID	9O2N	9NG7
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 60.641	a = 50.579
	b = 100.527	b = 87.614
	c = 147.025	c = 154.290
	α = 90, β = 90, γ = 90	α = 90, β = 90, γ = 90
Total reflections	3235874	1509835
Resolution range (Å) (outer shell)	50.00 – 2.20	50.00 – 1.80
	(2.24-2.20)	(1.83-1.80)
Unique reflections (outer shell)	20691 (1026)	32217 (1575)
Completeness (outer shell)	88.7 (90.2)	99.5 (98.7)
Average I/σ	14.18	27.66
Refinement Statistics		
Reflections (outer shell)	20511 (777)	31857 (1285)

Table 4.13. (cont'd)

Reflections (R-free) (outer shell)	2008 (68)	2017 (85)
R-work (outer shell)	0.1940 (0.2083)	0.1897 (0.2658)
R-free (outer shell)	0.2370 (0.2227)	0.2309 (0.2677)
Structure Statistics		
Non-hydrogen atoms	3539	2962
Solvent atoms	116	155
Ligand atoms	0	8
Protein residues	435	360
RMS (Bonds) (Å)	0.008	0.007
RMS (angles) (°)	0.96	0.88
Ramachandran plot (%) favored, allowed, outliers	97.90, 1.63, 0.47	96.87, 1.13, 0.00
Rotamer outliers (%)	0.56	2.11
B-factor, Average	32.32	25.17
Macromolecules, Ligands, Solvent	32.24, -, 34.71	24.90, 40.61, 29.32

	<i>cis</i> -CaaD (T34A) ^a (APO)	<i>cis</i> -CaaD (T34A) ^b (SO ₄)
PDB ID	9NG8	9NG9
Data statistics		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
No. of chains/ASU	3	3
Unit cell dimensions (Å, °)	a = 59.749	a = 59.617

Table 4.13. (cont'd)

	b = 100.399	b = 100.242
	c = 145.774	c = 146.836
	$\alpha = 90, \beta = 90, \gamma = 90$	$\alpha = 90, \beta = 90, \gamma = 90$
Total reflections	1213132	2233750
Resolution range (Å) (outer shell)	50.00 – 2.50	50.00 – 2.35
	(2.54-2.50)	(2.39-2.35)
Unique reflections (outer shell)	15048 (723)	16841 (825)
Completeness (outer shell)	98.7 (96.7)	88.8 (89.5)
Average I/ σ	5.4	3.19

Refinement Statistics

Reflections (outer shell)	13477 (772)	15740 (1236)
Reflections (R-free) (outer shell)	1360 (84)	1576 (124)
R-work (outer shell)	0.2014 (0.3081)	0.2073 (0.2634)
R-free (outer shell)	0.2835 (0.4432)	0.2799 (0.3728)

Structure Statistics

Non-hydrogen atoms	3501	3491
Solvent atoms	59	56
Ligand atoms	15	15
Protein residues	436	435
RMS (Bonds) (Å)	0.008	0.011
RMS (angles) (°)	0.96	1.02

Table 4.13. (cont'd)

Ramachandran plot (%) favored, allowed, outliers	94.65, 4.88, 0.47	96.97, 2.80, 0.23
Rotamer outliers (%)	2.84	5.98
B-factor, Average	33.51	32.63
Macromolecules, Ligands, Solvent	33.55, 36.50, 30.30	32.59, 41.27, 32.72

	<i>cis</i> -CaaD (T34A) ^c (SO ₄)	<i>N-cis</i> -CaaD/C-Cg10062 ^b (Cl ⁻)
PDB ID	9NGA	9NGD
Data statistics		
Space group	C 2 2 2 ₁	P 6 3
No. of chains/ASU	3	1
Unit cell dimensions (Å, °)	a = 59.552 b = 98.326 c = 142.892 $\alpha = 90, \beta = 90, \gamma = 90$	a = 59.783 b = 59.783 c = 57.126 $\alpha = 90, \beta = 90, \gamma = 120$
Total reflections	829911	4561954
Resolution range (Å) (outer shell)	50.00 – 2.40 (2.44-2.40)	50.00 – 2.70 (2.75-2.70)
Unique reflections (outer shell)	15076 (554)	2758 (160)
Completeness (outer shell)	88.7 (66.5)	83.3 (98.8)
Average I/ σ	7.89	12.27
Refinement Statistics		

Table 4.13. (cont'd)

Reflections (outer shell)	11836 (573)	2751 (1079)
Reflections (R-free) (outer shell)	1183 (52)	273 (106)
R-work (outer shell)	0.2114 (0.2779)	0.2391 (0.2870)
R-free (outer shell)	0.3078 (0.3664)	0.3363 (0.3105)
Structure Statistics		
Non-hydrogen atoms	3474	970
Solvent atoms	39	10
Ligand atoms	15	1
Protein residues	435	123
RMS (Bonds) (Å)	0.008	0.009
RMS (angles) (°)	1.08	1.19
Ramachandran plot (%) favored, allowed, outliers	93.94, 4.43, 1.63	90.91, 7.44, 1.65
Rotamer outliers (%)	3.99	3.06
B-factor, Average	37.62	46.73
Macromolecules, Ligands, Solvent	37.63, 40.82, 35.57	46.79, 52.37, 40.22

	<i>tcis</i>-CaaD^c	Cg10062 E114N^a
	(SO₄)	(APO)
PDB ID	9NGB	3NF4
Data statistics		
Space group	P 6 3	C 1 2 1

Table 4.13. (cont'd)

No. of chains/ASU	1	6
Unit cell dimensions (Å, °)	a = 59.266	a = 153.555
	b = 59.266	b = 88.834
	c = 57.816	c = 81.655
	$\alpha = 90, \beta = 90, \gamma = 120$	$\alpha = 90, \beta = 100.058, \gamma = 90$
Total reflections	1328281	1542266
Resolution range (Å) (outer shell)	50.00 – 2.50	40.00 – 2.50
	(2.54-2.50)	(2.54-2.50)
Unique reflections (outer shell)	3682 (192)	35409 (1774)
Completeness (outer shell)	91.3 (100.0)	96.8 (98.8)
Average I/ σ	13.48	7.19

Refinement Statistics

Reflections (outer shell)	3661 (779)	33933 (1240)
Reflections (R-free) (outer shell)	371 (77)	1995 (71)
R-work (outer shell)	0.2089 (0.2301)	0.2136 (0.3248)
R-free (outer shell)	0.2991 (0.2817)	0.3082 (0.3585)

Structure Statistics

Non-hydrogen atoms	963	7392
Solvent atoms	27	177
Ligand atoms	5	30
Protein residues	119	881

Table 4.13. (cont'd)

RMS (Bonds) (Å)	0.008	0.009
RMS (angles) (°)	0.90	1.17
Ramachandran plot (%) favored, allowed, outliers	97.44, 1.71, 0.85	92.17, 5.41, 2.42
Rotamer outliers (%)	3.16	6.32
B-factor, Average	38.28	40.15
Macromolecules, Ligands, Solvent	38.20, 49.02, 38.99	40.69, 40.00, 34.87

	Cg10062 E114N^b (ACA)	Cg10062 E114N^b (HPA, ACR)
PDB ID	9NGH	3NF6
Data statistics		
Space group	P 41 3 2	C 1 2 1
No. of chains/ASU	1	6
Unit cell dimensions (Å, °)	a = 147.120 b = 147.120 c = 147.120 $\alpha = 90, \beta = 90, \gamma = 90$	a = 152.396 b = 88.062 c = 79.318 $\alpha = 90, \beta = 100.339, \gamma = 90$
Total reflections	2535003	2365179
Resolution range (Å) (outer shell)	50.00 – 2.60 (2.64 - 2.60)	40.00 – 2.60 (2.64-2.60)
Unique reflections (outer shell)	17302 (843)	29108 (1475)
Completeness (outer shell)	99.8 (100.0)	96.6 (99.4)

Table 4.13. (cont'd)

Average I/ σ	13.5	8.8
---------------------	------	-----

Refinement Statistics

Table 4.13. (cont'd)

Reflections (outer shell)	10828 (584)	25850 (482)
Reflections (R-free) (outer shell)	1082 (85)	1995 (37)
R-work (outer shell)	0.2518 (0.2562)	0.1940 (0.2722)
R-free (outer shell)	0.3117 (0.2745)	0.2921 (0.3370)

Structure Statistics

Non-hydrogen atoms	1216	7294
Solvent atoms	1199	81
Ligand atoms	5	31
Protein residues	12	883
RMS (Bonds) (Å)	0.011	0.009
RMS (angles) (°)	1.21	1.15
Ramachandran plot (%) favored, allowed, outliers	95.14, 4.17, 0.69	90.01, 7.92, 2.07
Rotamer outliers (%)	0.78	7.48
B-factor, Average	38.02	46.68
Macromolecules, Ligands, Solvent	37.85, 61.13, 44.87	46.75, 46.25, 41.03

4.13 Chapter Three: Application of Cg10062 (E114N) to 3-Hydroxypropionic Acid

Synthesis

4.13.1 Bacterial Strains, Genes, and Plasmids

Escherichia coli strains DH5 α and MG1655 were obtained from Invitrogen. Cells were grown at 37 °C in LB or M9 media containing an appropriate antibiotic where necessary. Bgl-Brick vector pBbA1a-RFP was purchased from Addgene.¹⁵ A list of bacterial strains used in chapter three is shown in table 4.14.

Table 4.14. Strains and plasmids used in Chapter Three of this Study.

Strain or Plasmid	Description	Reference/ Source
<i>E. coli</i> DH5 α	K-12 B F ⁻ ϕ 80/ <i>lacZ</i> Δ M15 Δ (<i>lacZ</i> YA- <i>argF</i>)U169 <i>recA1 endA1 hsdR</i> 17(<i>r</i> _K ⁻ , <i>m</i> _K ⁺) <i>phoA supE</i> 44 λ - <i>thi</i> -1 <i>gyrA</i> 96 <i>relA</i> 1	Invitrogen
<i>E. coli</i> BL21	K-12 B <i>fhuA</i> 2 [<i>lon</i>] <i>ompT gal</i> [<i>dcm</i>] Δ <i>hsdS</i>	Invitrogen
<i>E. coli</i> BL21(DE3)	K-12 B F ⁻ <i>ompT hsdS</i> _B (<i>r</i> _B ⁻ , <i>m</i> _B ⁻) <i>gal dcm</i> (DE3)	Invitrogen
<i>E. coli</i> MG1655	F ⁻ λ ⁻ <i>ilvG</i> ⁻ <i>rfb</i> -50 <i>rph</i> -1	Invitrogen
pET-21a(+)	Ap ^R , <i>lacI</i> , <i>P</i> _{T7} pMB1 replicon	Invitrogen
pBbA1a-RFP	Ap ^R , <i>lacI</i> ^q , <i>P</i> _{trc} <i>rfp</i> p15A replicon	Addgene
pNJD1.002	<i>P</i> _{T7} <i>cis</i> - <i>CaaD</i> in pET-21a(+)	Devito
pAS1.046	<i>P</i> _{T7} <i>cg10062</i> in pET-21a(+)	Sirinimal
pAS2.100	<i>P</i> _{T7} <i>Cg10062 (E114N)</i> in pET-21a(+)	Sirinimal
pKK1.1025	<i>P</i> _{T7} <i>mmsB</i> in pET-21a(+)	Kwiatkowski
pAIDA1	Cmr ^R <i>lacUV5 aidA</i> in pKM1D	Addgene
pKS8.0411	<i>P</i> _{lacuv} <i>Cg10062 (E114N)</i> in pAIDA1	This Study
<i>Corynebacterium glutamicum</i>	NCBI 10025	ATCC 13032
<i>Pseudomonas putida</i> KT2440	<i>rmo</i> ⁻ <i>mod</i> ⁺	ATCC 47054
<i>Pseudomonas denitrificans</i>	NCBI 9496	ATCC 13867
<i>Pseudomonas fluorescens</i> Pf-5	NCTC 10038	ATCC BAA-477
<i>Pseudomonas fluorescens</i> BL915	NCPPB 316	ATCC 17556
<i>Rhodopseudomonas palustris</i>	CGA009	ATCC BAA-98
<i>Zymomonas mobilis</i>	NCIB 11199	ATCC 29191
<i>Enterobacter cloacae</i>	CDC 442-68	ATCC 35549

Table 4.14. (cont'd)

<i>Pseudomonas chlororaphis</i> subsp. <i>aureofaciens</i>	NRRL B-1576	ATCC 13985
<i>Bacillus subtilis</i>	NCIB 3610	ATCC 6051
<i>Aspergillus flavus</i>	WB 1957	ATCC 16883
<i>Bacillus pumilus</i>	NRS 272	ATCC 7061
<i>Bacillus cereus</i> UW85	BCRC 10603	ATCC 53522
<i>Candida parapsilosis</i>	CBS 604	ATCC 22019
<i>Klebsilla oxytoca</i>	479-2	ATCC 43165
<i>Bacillus circulans</i>	NCIB 9374	ATCC 21558
<i>Citrobacter braakii</i>	NCIB 8173	ATCC 6750

Plasmid pAS2.100

The plasmid, pAS2.100, encoding Cg10062 (E114N) was designed as described previously (Figure 4.40).⁸

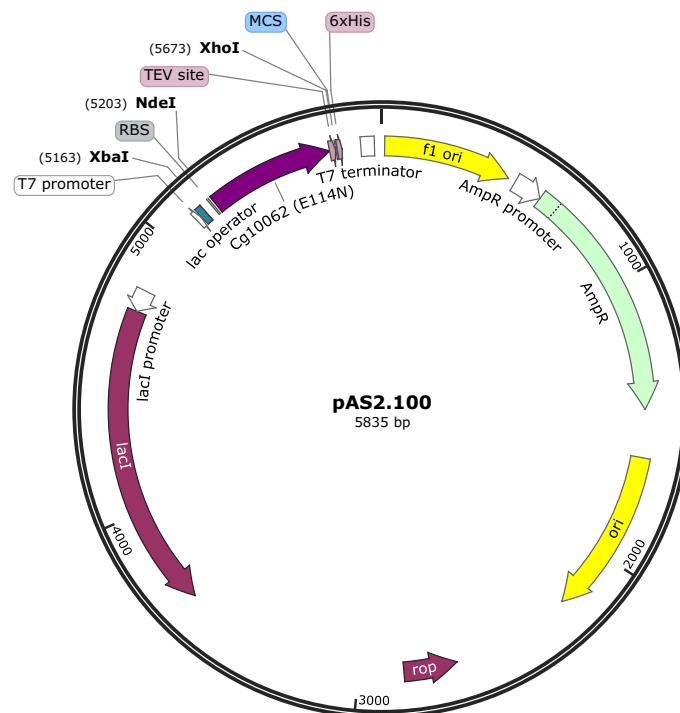


Figure 4.40. Plasmid map of pAS2.100.

pAIDA1

The pAIDA1 plasmid containing necessary promoter, operator, inducer, antibiotic resistance, signal peptide, recognition sites, linker, ori, and *aida* gene cloned into *E. coli* was obtained from Addgene (Figure 4.41).¹⁹ pAIDA plasmid DNA was isolated using a QIAprep Miniprep kit.

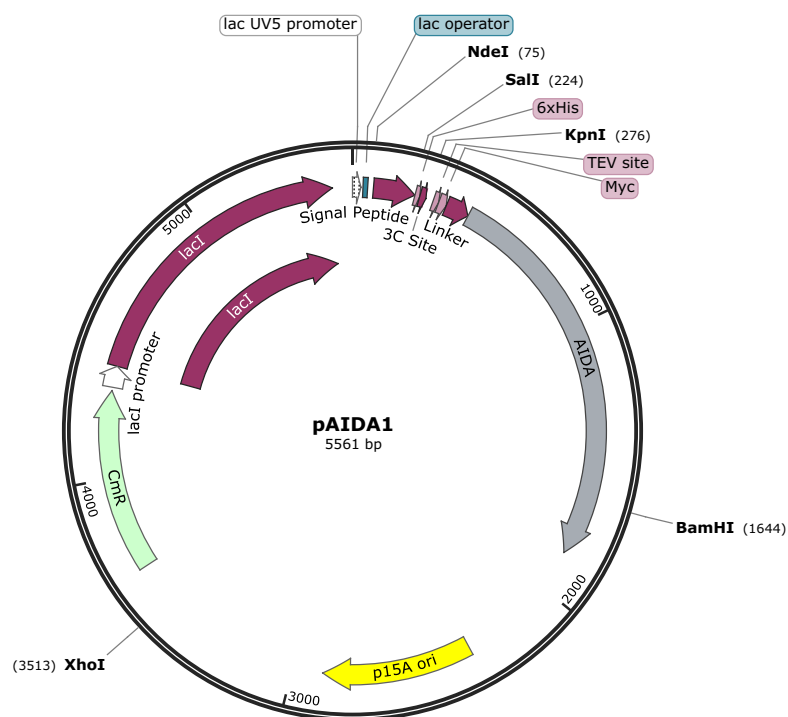


Figure 4.41. Plasmid map of pAIDA1.

Plasmid pKS8.0410

Q5 site-directed mutagenesis was used to move the His₆ tag in front of the TEV site in Cg10062 (E114N) named pKS8.0410 (Figure 4.42). This allows for easy detection of Cg10062 (E114N) through Western Blot confirmation analysis through anti-his antibody recognition, while preserving the option of TEV protease cleavage of the Cg10062 (E114N) passenger. Q5 site-directed mutagenesis was performed in three steps as described previously. Single colonies were inoculated into separate culture tubes containing 5 mL of LB media with respective antibiotic and the cultures were shaken at 37 °C overnight. DNA extraction from the cell pellets was carried out using a QIAprep Miniprep kit. Isolated plasmids were sequenced by the Michigan State University

Research Technology Support Facility (MSU RTSF) Genomics Core and the plasmids containing the desired mutations were transformed into electrocompetent *E. coli* BL21(DE3) and stored at -20°C.

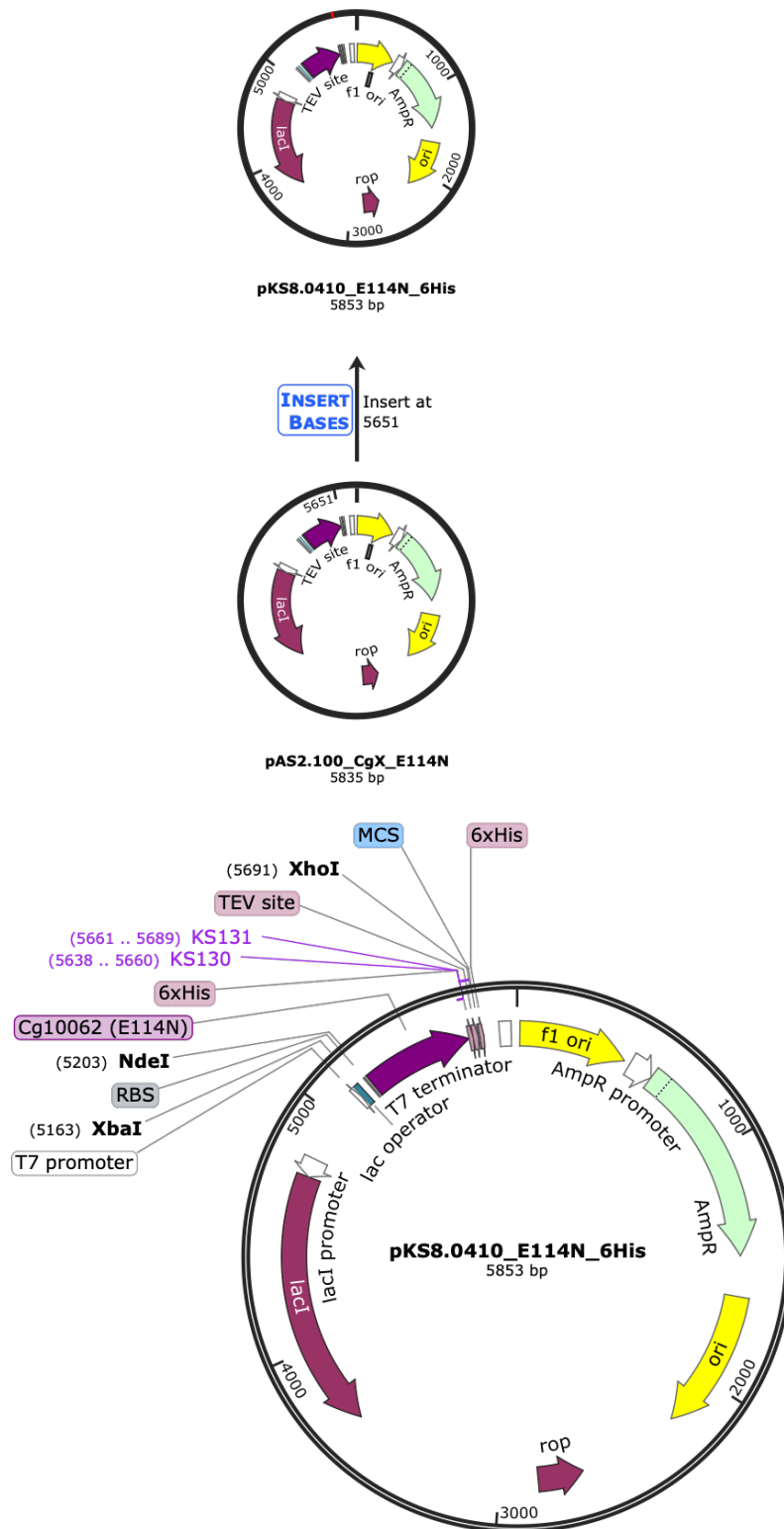


Figure 4.42. Plasmid design and map of pKS8.0410.

Plasmid pKS8.0411

Isolated DNA from pKS8.0410 was inserted into pAIDA1 through Gibson assembly to make pKS8.0411 (Figure 4.43). Here polymerase chain reactions (PCR) of template (pAIDA1) and insert (pKS8.0410) was performed using designed primers to amplify the template and insert, respectively (Tables 4.15, 4.16). The forward primer used on template (KS132) was TCTAGAGTGAATAACAATGGAAG, while the reverse primer (KS133) was TGCAAATGCATTTCCGATTG. The forward primer used on the insert DNA (KS134) was CAATCGGAAATGCATTTGCACCGACCTACACCTGCTGG, while the reverse primer (KS135) was CCATTGTTATTCACTCTAGAGCCTTGAAAATACAGGTTCTCATG. Separate PCR products were run on a 0.7% agarose gel with 50 μ L PCR product and 10 μ L loading dye. Bands corresponding to respective insert and vector were cut out. A Monarch Gel DNA Extraction protocol was used to isolate template and insert DNA. HiFi Gibson assembly was then performed to anneal the two fragments at 50 °C for one hour (Table 4.17). 5 μ L of product was then immediately transformed into chemically competent DH5 α cells (protocol described previously). Single colonies were inoculated into separate culture tubes containing 5 mL of LB media with respective antibiotic and the cultures were shaken at 37 °C overnight. DNA was extracted and sequenced. Plasmids containing the desired mutations were transformed into electrocompetent *E. coli* BL21(DE3) as previously described.

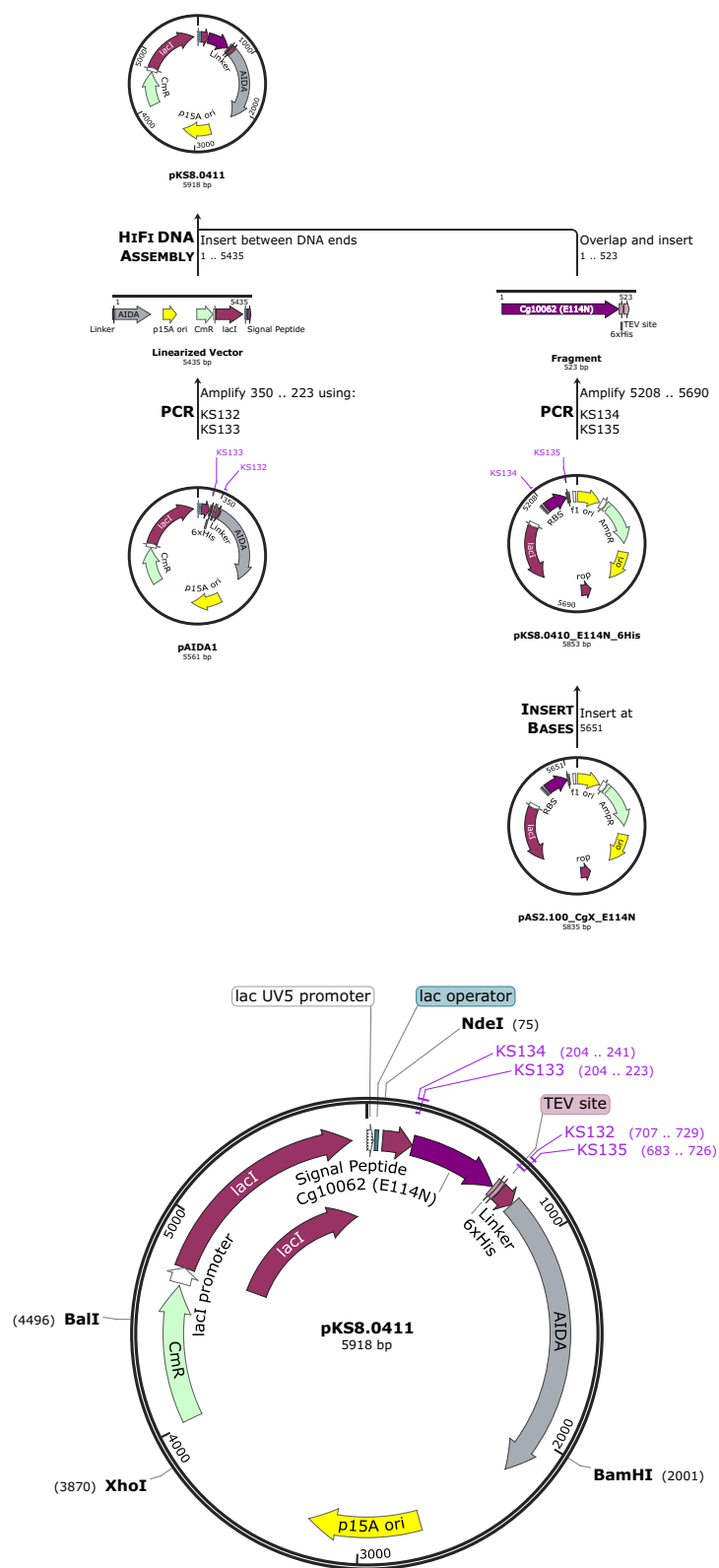


Figure 4.43. Gibson assembly design to make plasmid pKS8.0411.

Table 4.15. PCR reactions for the amplification of vector and insert in pAIDA/Cg10062(E114N) fusion protein design.

pAIDA1 Reaction			Cg10062 (E114N) Reaction		
Component	Volume	[Final]	Component	Volume	[Final]
5X Q5 Buffer (5X)	10 μ L	1 X	5X Q5 Buffer (5X)	10 μ L	1 X
dNTP (10 mM)	1 μ L	200 μ M	dNTP (10 mM)	1 μ L	200 μ M
Forward Primer (KS132) (10 μ M)	2.5 μ L	0.5 μ M	Forward Primer (KS134) (10 μ M)	2.5 μ L	0.5 μ M
Reverse Primer (KS133) (10 μ M)	2.5 μ L	0.5 μ M	Reverse Primer (KS135) (10 μ M)	2.5 μ L	0.5 μ M
pAIDA1 (2 ng/ μ L)	10 μ L	20 ng*	pKS8.0410 (2 ng/ μ L)	1 μ L	190 ng*
Q5 DNA Polymerase	0.5 μ L	-	Q5 DNA Polymerase	0.5 μ L	-
GC Enhancer (5X)	10 μ L	1 X	GC Enhancer (5X)	10 μ L	1 X
Deionized water	13.5 μ L	-	Deionized water	22.5 μ L	-

*NEB recommends 1ng-1 μ g

Table 4.16. Thermocycler settings for fusion protein design.

pAIDA1 Reaction			Cg10062 (E114N) Reaction		
Step	Temperature	Time	Step	Temperature	Time
Initial Denaturation	98 $^{\circ}$ C	30 Seconds	Initial Denaturation	98 $^{\circ}$ C	30 Seconds
35 Cycles	98 $^{\circ}$ C	10 Second	35 Cycles	98 $^{\circ}$ C	10 Second
	60 $^{\circ}$ C (Tm)	30 Seconds		60 $^{\circ}$ C (Tm)	30 Seconds
Final Extension	72 $^{\circ}$ C	2 Minutes*	Final Extension	72 $^{\circ}$ C	30 Seconds*
	72 $^{\circ}$ C	2 Minutes		72 $^{\circ}$ C	2 Minutes
Hold	4 $^{\circ}$ C	Forever	Hold	4 $^{\circ}$ C	Forever

* 30 seconds per kbp

Table 4.17. HiFi Gibson assembly reaction components for fusion protein.

Component	Reaction Volume
pAIDA1 (Vector)	3 μ L
pKS8.0410 (Insert)	1 μ L
HiFi Master Mix	10 μ L
Deionized water	6 μ L

4.13.2 Bacterial Growth on ACA

The following species were selected to screen for growth on ACA. *Pseudomonas fluorescens* Pf-5, *Pseudomonas fluorescence* BL915, *Rhodopseudomonas palustris*, *Pseudomonas putida* KT2440, *Zymomonas mobilis*, and *Pseudomonas dentrificans* were grown

at 30 °C according to their known growth conditions. *Enterobacter cloacae*, *Pseudomonas chloroaphis* subsp. *Aureofaciens*, *Bacillus subtilis*, *Asperigillus flavus*, *Bacillus pumilus*, *Corynebacterium glutamicum*, *Bacillus cereus* UW85, *Candida parapsilosis*, *Klebsiella oxytoca*, *Bacillus circulans*, and *Citrobacter braakii* were grown at 37 °C according to their known growth conditions. To determine the growth of bacterial species on ACA, agar plates were made containing M9 salts (366 mM Na₂HPO₄, 22 mM KH₂PO₄, 19 mM NH₄Cl, and 9 mM NaCl), minimal M9 media with 20 mM glucose, M9 with 20 mM ACA, and LB. These plates were divided into four quadrants. A separate bacterial species was streaked onto one quadrant of each media type and incubated at their known incubation temperatures. From the colonies that grew on M9/ACA plates, 2 colonies from each plate were selected and inoculated in 5 mL liquid cultures of both M9/ACA and M9/Glu. Liquid growth from these colonies were unsuccessful. Therefore, colonies were selected from LB plates of species that had grown on solid ACA media. It was thought that perhaps taking a starter colony of a species that we have seen grow on ACA from rich media may allow growth in liquid media. Again, 2 colonies from each LB plate as well as 2 colonies from the matching M9/ACA plate was inoculated in M9/ACA and M9/Glu. Growth in liquid ACA media was unsuccessful. Slight growth was seen for *Klebsiella oxytoca*, however continued growth was unsustainable. Thiamine was removed from M9 solid media, as it may contribute as a carbon source. However, experiments yielded the same results.

This brought attention back to *E. coli* as a microbe for use with ACA. To determine the amount of minimal glucose and maximum ACA to be used in media, a series of short-term *E. coli* adaptive growth studies were performed. A single overnight 5 mL culture of MG1655 *E. coli* cells were grown at 37 °C shaking at 200 rpm. The culture was then pelleted by centrifugation at 17,000 rpm for 1 minute and supernatant was removed. The pellet was then resuspended in M9 salts and pelleted again. This step was repeated twice. The final pellet was divided into two and resuspended in M9 salts and diluted to a stock OD₆₀₀ of 0.4. Of this stock, 10 µL of culture was

added to each well. Varying concentrations of glucose and ACA were used to find the optimal concentration of *E. coli* viability on glucose while maintaining ACA in the cell. This was also performed with glycerol. Controls were performed using M9/salts. A stock solution of M9 glucose, M9 glycerol, of M9 ACA was used in varying concentrations of 0, 4, 8, 12, 16, and 20 mM. Each was diluted with an alternate solution. A 96 well plate was filled with each well containing a different ratio of media. For example, one well contained M9/Glu 0mM: M9/ACA 20 mM, the next M9/Glu 4mM: M9/ACA 16mM, and so forth. The following row contained M9/Gly 0mM: M9/ACA 20 mM, the next M9/Gly 4 mM: M9/ACA 16 mM, etc. Media diluted with M9 salts served as a control. 190 μ L of media mixture was added to each well to obtain 200 μ L total in each well. All runs were performed in triplicate. This experiment was repeated two more times with two more separate MG1655 colonies to confirm results. M9 media containing 4mM glucose and 16 mM ACA maintained adequate cell growth while keeping high ACA concentration within media. This shows that ACA is not toxic to cells in high quantities and conversion of ACA to a transportable chemical, such as MSA, extracellularly can be a means to utilize enzymatic conversion of ACA *in vivo*.

4.13.3 Lysate Supplemented Growth on MSA

To determine growth on malonic semialdehyde, Cg10062 (E114N) lysate was obtained from a previously frozen cell mass of expressed protein. A BL21(DE3)/pAS2.100 cell pellet was lysed with 3 mL lysis buffer to 1 g mass pellet and passed through a French press as described previously. After confirmation of enzyme activity through the activity assay mentioned previously, specific activity of the cell lysate was confirmed to be in the expected range of approximately 7 U mg^{-1} . MG1655 *E. coli* cells were grown in M9/Glu overnight at 37 °C and washed with M9/salts twice. Cells were then reinoculated to an OD of 0.05 in M9 media containing 20 mM ACA. The amount of Cg10062 (E114N) lysate added to media was calculated based on specific activity and how much lysate would be needed to consume 20 mM of ACA within a 30-minute reaction. OD₆₀₀

was taken at various time points while simultaneously taking ^1H NMR reading during these same time points. At each time point 700 μL of sample was taken and quenched with 2 μL of concentrated H_2SO_4 . Samples were then spun down to remove precipitated protein and 600 μL of supernatant was added to 30 μL of DMSO and 70 μL of 100 mM TSP. This was then evaluated using ^1H Wet 1D NMR.

Additionally, a plasmid encoding the gene for MmsB (pKK1.1025), an enzyme that is able to convert MSA to 3HP was transformed into electrocompetent BL21(DE3) cells as described previously. This plasmid contains a T7 promoter which enhances expression of MmsB promoting production of 3HP. Overnight cultures of BL21(DE3)/ pKK1.1025 grown at 37 °C with necessary antibiotic was then reinoculated to an OD_{600} of 0.05 in M9 media containing 20 mM of ACA, Cg10062 (E114N) depending on the specific activity and amount needed to completely reduce ACA within 30 minutes, antibiotic, and IPTG for gene expression. A sample without IPTG was run simultaneously and served as a control. OD_{600} was taken at various time points while simultaneously taking ^1H NMR reading during these same time points. At each time point 700 μL of sample was taken and quenched with 2 μL of concentrated H_2SO_4 . Samples were then spun down to remove precipitation and 600 μL of supernatant was added to 30 μL of DMSO and 70 μL of 100 mM TSP. This was then evaluated using ^1H Wet 1D NMR. Additional timepoints were taken to evaluate reaction progress.

4.13.4 Fusion Protein Growth Studies and ^1H NMR

A 5 mL culture of fusion protein (BL21(DE3)/pKS8.0411) was grown overnight in M9 media containing 22 mM glucose and chloramphenicol antibiotic. The overnight culture was inoculated to an OD_{600} of 0.1 in 50mL M9 media containing 4 mM Glu and 16 mM ACA with antibiotic. The culture was induced with 0.2mM IPTG when OD_{600} reached 0.2 (approximately 2 hours). OD_{600} was taken at 0, 30 min, 2 h, 6 h, and 24 h time points.

While OD₆₀₀ was being taken at 0, 30 min, 2 h, 6 h, and 24 h time points, ¹H NMR spectra was also taken. Samples of 700 µL were taken and quenched with 2 µL of concentrated H₂SO₄ and spun down to remove any precipitation. 600 µL of supernatant was then removed and added to an Eppendorf tube containing 30 µL DMSO and 70 µL TSP (100mM). This was then transferred to a NMR tube for ¹H Wet 1D NMR analysis.

REFERENCES

1. Sambrook, J.; Russell, D. W. *Molecular Cloning: A Laboratory Manual*; CSHL Press, 2001.
2. 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.
3. Miller, J. H. *Experiments in Molecular Genetics*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N. Y, 1972
4. Gill, S. C.; von Hippel, P. H. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. *Anal. Biochem.* 1989, *182* (2), 319–326.
5. *Protein Calculator*. <http://protcalc.sourceforge.net/> (accessed 2022-04-21).
6. 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.
7. *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).
8. Mathes Hewage, A. Synthesis of 3-hydroxypropionic Acid From Acetylenecarboxylic acid. *Dissertation*. 2022
9. Kwiatkowski, K. Synthesis of lactic acid and 3-hydroxypropionic acid via 1,3- and 1,4-hydration of acetylenecarboxylic acid. *Dissertation*. 2023
10. 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. *Biochem.* 2008, *47*, 8139–8147.
11. 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*, 48674–48683.
12. Schroeder, G. K.; Huddleston, J. P.; Johnson, W. H.; Whitman, C. P. A Mutational Analysis of the Active Site Loop Residues in *cis*-3-Chloroacrylic Acid Dehalogenase. *Biochem.* 2013, *52*, 4204–4216.
13. 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. *Biochem.* 2021, *60*, 3879-3886.
14. Mathes Hewage, A., Silva, K., Kwiatkowski, K., Lee, H., Sreedhar, D., Gavgani, H. Geiger, J., Draths, K. An Original Biosynthetic Route to 3-Hydroxypropionic Acid, 2025, *in preparation*
15. 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.

16. Manual, I. GeneMorph II Random Mutagenesis Kit.
<https://www.agilent.com/cs/library/usermanuals/public/200550.pdf>.
17. Torices, R.; Muñoz-Pajares, A. J. PHENIX: An R Package to Estimate a Size-Controlled Phenotypic Integration Index1. *Appl. Plant Sci.* 2015, 3 (5), 1400104.
18. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 2010, 66, 486–501.
19. Jarmander, J.; Gustavsson, M.; Do, T.-H.; Samuelson, P.; Larsson, G. A Dual Tag System for Facilitated Detection of Surface Expressed Proteins in *Escherichia coli*. *Microb. Cell Fact.* 2012, 11, 118.