

Washington University in St. Louis

## Washington University Open Scholarship

---

All Theses and Dissertations (ETDs)

---

January 2009

### An Investigation of *Acetobacter aceti* N5-Carboxyaminoimidazole Ribonucleotide Mutase and Its *purE-purK* Operon

Charles Constantine

*Washington University in St. Louis*

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

---

#### Recommended Citation

Constantine, Charles, "An Investigation of *Acetobacter aceti* N5-Carboxyaminoimidazole Ribonucleotide Mutase and Its *purE-purK* Operon" (2009). *All Theses and Dissertations (ETDs)*. 74.  
<https://openscholarship.wustl.edu/etd/74>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact [digital@wumail.wustl.edu](mailto:digital@wumail.wustl.edu).

WASHINGTON UNIVERSITY

Department of Chemistry

Dissertation Examination Committee

T. Joseph Kappock, Co-Chair

Kevin D. Moeller, Co-Chair

Robert Blankenship

Michael L. Gross

Dewey J. Holten

Joseph M. Jez

An Investigation of *Acetobacter aceti*  $N^5$ -Carboxyaminoimidazole Ribonucleotide Mutase  
and Its *purE-purK* Operon

by

Charles Zoltan Constantine

A dissertation presented to the  
Graduate School of Arts and Sciences  
of Washington University in  
partial fulfillment of the  
requirements for the degree  
of Doctor of Philosophy

August 2009

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

An Investigation of *Acetobacter aceti*  $N^5$ -Carboxyaminoimidazole Ribonucleotide Mutase  
and

Its *purE-purK* Operon

by

Charles Zoltan Constantine

Doctor of Philosophy in Chemistry

Washington University in St. Louis

T. Joseph Kappock, Co-Chairperson

Kevin D. Moeller, Co-Chairperson

*Acetobacter aceti* oxidizes ethanol to acetic acid. While the membrane-permeable acetic acid is toxic to many bacteria, *A. aceti* survives exposure to acetic acid by tolerating cytoplasmic acidification. The ability to tolerate an acidic cytoplasm suggests that proteins from *A. aceti* are unusually suited to function in an acidic environment. The ability to tolerate an acidic cytoplasm raises additional questions about biosynthetic pathways that employ acid-labile intermediates. To examine how *A. aceti* metabolism may have adapted to function under acidic conditions, a biosynthetic conversion involving an acid-labile metabolite has been selected for study.

The enzyme  $N^5$ -carboxyaminoimidazole ribonucleotide mutase (PurE) catalyzes the reversible intramolecular transfer of the carboxylate of  $N^5$ -carboxyaminoimidazole ribonucleotide ( $N^5$ -CAIR) to the C-4 position of the imidazole ring to form 4-

carboxyaminoimidazole ribonucleotide (CAIR). A series of mutants was made to explore the role of two conserved histidines. A pH-rate comparison of *AaPurE* and the active mutant *AaPurE*-H59D was used to identify His59 as the key active site acid/base residue. The thermostability of *Escherichia coli* PurE (*EcPurE*) over a range of pH was also assessed and compared to *AaPurE*. *AaPurE* was found to be significantly more thermostable than *EcPurE* over the entire pH range surveyed. Comparison of the pH-rate profiles constructed for *AaPurE* with recently reported pH-rate profiles for *EcPurE* indicate that the two do not differ significantly, indicating there has been no adaptive change in enzyme mechanism. Also reported is a summary and analysis of a number of crystal structures that have been determined for *AaPurE*, which suggests a strategy by which proteins may have become resistant to acid-mediated inactivation.

Initial functional complementation studies using the *purE* auxotroph PC0135 suggested that *AaPurE* may require *AaPurK* to function. We constructed a set of stable *E. coli* deletion strains and insertion strains that replace the chromosomal copies of *purE<sub>Ec</sub>* or *purK<sub>Ec</sub>* with their counterparts from *A. aceti*. Functional complementation experiments suggest that a third protein, located upstream of *purE<sub>Aa</sub>* and conserved in Rhodospirillales, may be involved in the proper functioning of *AaPurE*.

The nonenzymatic decarboxylation of CAIR and the corresponding ribonucleoside (CAIR-s) were examined as models for the PurE reaction. The decarboxylation of CAIR-s was not acid-catalyzed in the pH range examined, and did not accelerate in lower-polarity solvents.

## ACKNOWLEDGEMENTS

I would like to thank my advisor Professor T. Joseph Kappock for his invaluable guidance over the years. Experiencing his unique scientific insight has been the highlight of my education.

I would like to thank my Ph. D. Advisory Committee members Professors Michael L. Gross and Kevin D. Moeller for their insight in the classroom and during the course of my research. I thank Professors Joseph Jez, Dewey Holten and Robert Blankenship for serving on my dissertation defense committee.

Thanks to the members of the Kappock group, past and present, Dr. Julie Francois, Dr. Hong Jiang, Elwood Mullins, and John Hung, and especially Dr. Courtney Starks for their aid and friendship. I would also like to thank members of the Gross group, Dr. David Hambly, Dr. Justin Sperry and Sandy Kerfoot for their aid and accessibility over the years. I would especially like to thank Dr. Ed Hiss; words cannot describe what an asset he is.

No list of acknowledgements would be complete without thanking those who got me where I am today. Thank you to my family for their support and encouragement over the years. Thank you to my father for answering the never ending string of whys I provided as a child. Thank you to the amazing teachers and professors I had before coming to Washington University, Cristina Geiger, Gayle Stout, and Dr. Harold Hoops. Thank you to those who are no longer with us: to my Uncle John Constantine for his support in the strange new city of St. Louis, and to my high school chemistry teacher Nancy Middleton; I now realize what a remarkable teacher you were.

Finally I would like to thank my friends. To Keith Taylor III and Patrick Lynn Clayton, thank you for being friends that I could count on through thick and thin at a moment's notice. To Dr. Elizabeth Elliott, Brian Barnes, Megan Daschbach, Matt VanDuzor and Dr. Dayna Tucker, without your friendship in graduate school, this would have been a far less enjoyable experience.

## TABLE OF CONTENTS

<b>Abstract of the Dissertation.....</b>	<b>ii</b>
<b>Acknowledgments.....</b>	<b>iv</b>
<b>Table of Contents.....</b>	<b>vi</b>
<b>List of Figures.....</b>	<b>viii</b>
<b>List of Tables.....</b>	<b>xii</b>
<b>List of Schemes.....</b>	<b>xiv</b>
<b>Chapter 1. Introduction</b>	
1.1 Perspective.....	2
1.2 Purpose.....	6
1.3 References.....	8
<b>Chapter 2. Non-enzymatic decarboxylation of 4-carboxy-5-aminoimidazole ribonucleotide (CAIR) and 4-carboxy-5-aminoimidazole ribonucleoside (CAIR-s).</b>	
2.1 Introduction.....	11
2.2 Methods and Materials.....	19
2.3 Results.....	22
2.4 Discussion.....	28
2.5 Future Directions.....	38
2.6 References.....	39
<b>Chapter 3. <i>N</i><sup>5</sup>-Carboxyaminoimidazole Ribonucleotide Mutase from <i>Acetobacter Aceti</i></b>	
3.1 Introduction.....	43
3.2 Methods and Materials.....	46
3.3 Results.....	53
3.4 Discussion.....	68
3.5 Future Directions.....	85
3.6 References.....	86
<b>Chapter 4. The <i>purE-purK</i> operon of <i>A. aceti</i></b>	
4.1 Introduction.....	90
4.2 Methods and Materials.....	95
4.3 Results.....	133

4.4 Discussion.....	160
4.5 Future Directions.....	183
4.6 References.....	184

**Appendix I. Citrate Synthase**

I.1 Introduction.....	188
I.2 Methods.....	190
I.3 Results.....	195
I.4 Discussion.....	206
I.5 References.....	208

<b>License for Figure 1.1.....</b>	<b>217</b>
------------------------------------	------------



## LIST OF FIGURES

<b>Figure 1.1</b>	Changes in the external and internal pH during growth of <i>A. acetii</i> with ethanol.....	5
<b>Figure 2.1</b>	Compounds used in this study.....	14
<b>Figure 2.2</b>	Nonenzymatic decarboxylation of CAIR and 2-(1-carboxy-1-hydroxyethyl)-3,4,-dimethylthiazolium.....	12
<b>Figure 2.3</b>	Direct decarboxylations.....	13
<b>Figure 2.4</b>	Decarboxylation of 3-carboxy-benzisoxazoles.....	14
<b>Figure 2.5</b>	Intramolecular hydrogen bonds.....	14
<b>Figure 2.6</b>	Ring protonation mechanism for CAIR-s at low pH.....	15
<b>Figure 2.7</b>	Ionizations of CAIR and CAIR-s.....	15
<b>Figure 2.8</b>	Illustration of ionization states of 4-carboxy-5-amino-imidazole ring.....	16
<b>Figure 2.9</b>	Possible alternate forms of CAIR-s in equilibrium with the zwitterion.....	16
<b>Figure 2.10</b>	Possible mechanism of decarboxylation for C4 protonated tautomer.....	17
<b>Figure 2.11</b>	Determination of $\Delta\epsilon_{260}$ for the conversion CAIR $\rightarrow$ $N^5$ -CAIR as a function of pH.....	23
<b>Figure 2.12</b>	pH-profile of rate of absorbance change at 260 nm for CAIR-s.....	26
<b>Figure 2.13</b>	C-4 protonation mechanism for CAIR zwitterion.....	31
<b>Figure 2.14</b>	C-4 protonation mechanism for decarboxylation of the CAIR cation.....	32
<b>Figure 2.15</b>	Direct ring protonation mechanism at low pH.....	33
<b>Figure 2.16</b>	One of the <i>many</i> possible mechanisms for formaldehyde decarboxylation.....	33

<b>Figure 2.17</b>	Hammick type mechanism.....	34
<b>Figure 2.18</b>	Hammick type mechanism, part 2.....	35
<b>Figure 2.19</b>	Possible role of PurE Ser57 and His59.....	36
<b>Figure 2.20</b>	Decarboxylation of $N^5$ -CAIR.....	37
<b>Figure 3.1</b>	Crystal structure of AaPurE with AIR bound.....	45
<b>Figure 3.2</b>	SDS-PAGE analysis (15% polyacrylamide) of an AaPurE purification.....	55
<b>Figure 3.3</b>	SDS-PAGE analysis of purified AaPurEs.....	55
<b>Figure 3.4</b>	Comparison of <i>E. coli</i> and <i>A. aceti</i> PurE $T_m$ values.....	57
<b>Figure 3.5</b>	pH dependence of wt AaPurE citrate affinity ( $K_{eq}$ ).....	59
<b>Figure 3.6</b>	Fluorescence titration of AaPurE with <i>cis</i> - and <i>trans</i> - aconitate.....	61
<b>Figure 3.7</b>	Fluorescence titration of inactive AaPurE mutants with CAIR.....	65
<b>Figure 3.8</b>	$k_{cat}/K_m$ pH-rate profiles.....	67
<b>Figure 3.9</b>	$k_{cat}$ pH-rate profiles.....	67
<b>Figure 4.1</b>	Illustration of method used to construct plasmid pJK389.....	106
<b>Figure 4.2</b>	Assembly of the insert of plasmid pJK421 (Part 1).....	109
<b>Figure 4.3</b>	Assembly of insert of pJK421 (Part 2).....	110
<b>Figure 4.4</b>	Assembly of insert of pJK448 (Part 1).....	113
<b>Figure 4.5</b>	Assembly of insert of pJK448 (Part 2).....	114
<b>Figure 4.6</b>	Schematic representation of strain CC0101 growth curves.....	145
<b>Figure 4.7</b>	Growth curves of strain CC0102 in minimal medium B.....	146
<b>Figure 4.8</b>	Schematic representation of strain CC0101 growth curves.....	147
<b>Figure 4.9</b>	Growth curves of strain CC0101 in minimal medium B.....	147

<b>Figure 4.10</b>	Schematic representation of strain CC0101 growth curves.....	150
<b>Figure 4.11</b>	Growth curves for strain CC0101 in minimal medium B.....	150
<b>Figure 4.12</b>	Schematic representation of strain W3110 growth curves.....	151
<b>Figure 4.13</b>	Growth of strain W3110 in minimal medium B.....	151
<b>Figure 4.14</b>	Illustration of plasmids used in growth experiments and summary of lag phases from Figure 4.13.....	155
<b>Figure 4.15</b>	Schematic representation of strain CC1202 growth curves.....	156
<b>Figure 4.16</b>	Growth curves for strain CC1202 in minimal medium B.....	156
<b>Figure 4.17</b>	Schematic representation of strain CC1203 and W3110 growth curves.....	158
<b>Figure 4.18</b>	Growth of strain CC1203 and W3110 in minimal medium B.....	159
<b>Figure 4.19</b>	Schematic representation of strain CC1204 and W3110 growth curves.....	160
<b>Figure 4.20</b>	Growth curves for strain CC1204 and W3110 in medium B.....	160
<b>Figure 4.21</b>	Illustration of plasmids used in growth experiments and summary of lag phases from Figure 4.20 and Figure 4.22.....	162
<b>Figure 4.22</b>	Schematic representation of strain CC1205 and W3110 growth curves.....	163
<b>Figure 4.23</b>	Growth curves for strain CC1205 and W3110 in medium B.....	163
<b>Figure 4.24</b>	Schematic representation of strain CC1205 growth curves.....	164
<b>Figure 4.25</b>	Growth curves for strain CC1205 in medium B.....	164
<b>Figure 4.26</b>	Comparison of growth curves of strains with a source of <i>EcPurE</i> .....	179
<b>Figure 4.27</b>	Comparison the effect of pJK173.....	180
<b>Figure 4.28</b>	Comparison of effects of pUC118 and pJK455.....	181

<b>Figure 4.29</b>	Comparison of the effect of pJK412.....	182
<b>Figure 4.30</b>	Comparison of the effects of pJK415 and pJK348.....	183
<b>Figure 4.31</b>	Comparison of the effect of pJK324.....	184
<b>Figure 4.32</b>	Sequence alignment of Orf1 from <i>A. aceti</i> and 12 close homologs.....	186
<b>Figure 4.33</b>	Sequence alignment of PurKs.....	188
<b>Figure I.1</b>	SDS-PAGE analysis of Cibracon Blue 3-GA agarose column purification of <i>TpCSTMD317G</i> .....	203
<b>Figure I.2</b>	CMX titration of <i>TpCSD317N</i> .....	205
<b>Figure I.3</b>	CMX titration of <i>TpCSD317G</i> .....	206
<b>Figure I.4</b>	CMX titration of <i>TpCSD317G</i> in 20% glycerol.....	206
<b>Figure I.5</b>	CMX titration of <i>TpCSD317G</i> in 30% sucrose.....	207
<b>Figure I.6</b>	CMCoA titration of <i>TpCSD317G</i> .....	207
<b>Figure I.7</b>	CMCoA titration of <i>TpCSD317G</i> in 20% glycerol.....	208
<b>Figure I.8</b>	CMCoA titration of <i>TpCSD317G</i> in 30% sucrose.....	208
<b>Figure I.9</b>	CMCoA titration of TM-D317G.....	209
<b>Figure I.10</b>	Stopped flow fluorescence of CMCoA binding TM-D317G-OAA at pH 8.0.....	211
<b>Figure I.11</b>	Values of $k_1$ and $k_2$ from double exponential fits of TM-D317G as a function of [CMCoA].....	211
<b>Figure I.12</b>	Stopped flow fluorescence of CMX binding TM-D317G-OAA at pH 8.0.....	212
<b>Figure I.13</b>	Gel filtration profile of TM-D317G.....	213

## LIST OF TABLES

<b>Table 2.1</b>	<i>k<sub>uncatalyzed</sub></i> of CAIR decarboxylation.....	24
<b>Table 2.2</b>	Activation energy for the decarboxylation of CAIR.....	25
<b>Table 2.3</b>	Solvent isotope effect on CAIR-s decarboxylation.....	27
<b>Table 2.4</b>	Decarboxylation of CAIR-s in mixed water-ethanol solvent.....	28
<b>Table 3.1</b>	Oligodeoxynucleotides (ODNs) used in this Chapter.....	47
<b>Table 3.2</b>	Purification chart for mutants with detectable activity throughout purification.....	56
<b>Table 3.3</b>	Results of ESI-MS of purified wild type and mutant <i>AaPurEs</i> .....	56
<b>Table 3.4</b>	Carboxylic acids screened as possible <i>AaPurE</i> ligands.....	60
<b>Table 3.5</b>	Properties of wt and mutant <i>AaPurEs</i> .....	62
<b>Table 3.6</b>	Activity detection threshold for residual activity of <i>AaPurE</i> -H59N by pH.....	63
<b>Table 3.7</b>	Summary of <i>AaPurE</i> crystal structures.....	81
<b>Table 3.8</b>	Comparison of <i>A. aceti</i> and <i>E. coli</i> crystal structures.....	83
<b>Table 4.1</b>	Oligodeoxynucleotides (ODNs) used in the construction of deletion strains.....	96
<b>Table 4.2</b>	ODNs used in the construction of plasmids and sequencing.....	96
<b>Table 4.3</b>	Description of strains.....	98
<b>Table 4.4</b>	Plasmids used in this Chapter.....	104
<b>Table 4.5</b>	Functional complementation studies of <i>purE</i> deficient strain <i>E. coli</i> PC0135.....	140
<b>Table 4.6</b>	Functional complementation studies of <i>purE<sub>Ec</sub></i> deficient strain PC0135 and <i>purE<sub>Ec</sub>purK<sub>Ec</sub></i> deficient strain CC0102 by <i>purE<sub>Ec</sub></i> constructs on solid media.....	143

<b>Table 4.7</b>	Functional complementation studies of strains CC0101, CC1201, and CC1202.....	149
<b>Table 4.8</b>	Assessment of the ability of strains CC0102, CC0103, CC1203, CC1204 and CC1205 to grow on solid medium B.....	153
<b>Table I.1.</b>	Acetyl-CoA analogue $K_d$ s Determined for <i>TpCS</i> mutants.....	204
<b>Table I.2</b>	Quantum yield of TM-D317G.....	214

## LIST OF SCHEMES

<b>Scheme 1.1</b>	Reaction catalyzed by Class I and Class II PurEs.....	3
<b>Scheme 3.1</b>	Reactions catalyzed by Class I and Class II PurEs.....	43
<b>Scheme 3.2</b>	Purification scheme for <i>AaPurE</i> .....	54
<b>Scheme 3.3</b>	Possible mechanism of <i>AaPurE</i> .....	74
<b>Scheme 4.1</b>	Organization of <i>purE-purK</i> operon in <i>A. aceti</i> and <i>E. coli</i> .....	90
<b>Scheme 4.2</b>	Other divergences in purine biosynthesis.....	92
<b>Scheme 4.3</b>	Method of gene deletion.....	99
<b>Scheme 4.4</b>	Illustration of cross-over PCR.....	103
<b>Scheme 4.5</b>	Method for precise gene insertion using pKOV vector system.....	117
<b>Scheme 4.6</b>	Diagram showing inserts of plasmids.....	134
<b>Scheme 4.7</b>	Schematic representation of the <i>purEpurK</i> region of strains used in this study.....	137
<b>Scheme 4.8</b>	Illustration of the inserts of plasmids used in the functional complementation studies of the <i>purE</i> strain PC0135.....	139
<b>Scheme I.1</b>	Reaction of citrate synthase.....	196

## **Chapter 1**

### Introduction



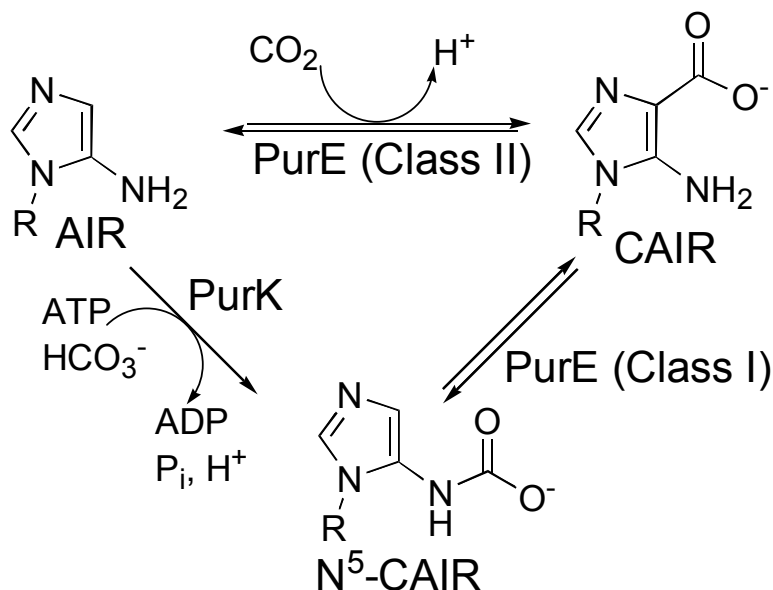
## 1.1 Perspective

The formation of carbon-carbon bonds represents some of the most interesting reactions in biosynthesis. Enzymes have evolved to catalyze many of these reactions through nucleophilic intermediates that are formed in the near neutral, aqueous environment of cells. This is especially interesting in the cases of enzymes that lack cofactors, because they cannot use the Lewis acidity of metals to stabilize reactive intermediates. To catalyze these reactions, enzymes employ a number of strategies including: ground state destabilization, transition state stabilization, and the sequestration of substrate and reactive intermediates from aqueous solvent. Many enzymes function simultaneously as both acid and base catalysts using distinct residue(s) in the active site.

One enzyme that catalyzes carbon-carbon bond formation by nucleophilic attack of a stabilized carbanion (or equivalent) on an electrophilic carbonyl is PurE. PurE catalyzes the formation of 4-carboxy-5-aminoimidazole ribonucleotide (CAIR), which represents the sole carbon-carbon bond formed in the *de novo* synthesis of purine nucleotides.

*PurE*. The *de novo* biosynthesis of inosine monophosphate (IMP) is accomplished in a total of 10 enzymatic transformations in animals, while bacteria and other eukaryotes accomplish this using 11 enzymatic transformations (1). The difference derives from how 4-carboxy aminoimidazole ribonucleotide is produced in these pathways (Scheme 1.1). In animals, Class II PurE converts 5-aminoimidazole ribonucleotide (AIR) + CO<sub>2</sub> to CAIR (2). In bacteria, yeast, and plants, the conversion of

AIR → CAIR requires two distinct steps catalyzed by two separate enzymes, PurK and Class I PurE (3). PurK catalyzes the carboxylation of the exocyclic amino group of AIR to form  $N^5$ -carboxyaminoimidazole ribonucleotide ( $N^5$ -CAIR), from AIR, ATP and bicarbonate likely through a carboxyphosphate intermediate (3). Class I PurE then catalyzes the reversible intramolecular transfer of the carboxylate to the C-4 position of the imidazole ring to form CAIR (4). The different substrate preferences make Class I PurEs  $N^5$ -carboxyaminoimidazole ribonucleotide mutases (EC 5.4.99.18), and Class II PurEs phosphoribosylaminoimidazole carboxylases (EC 4.1.1.21).



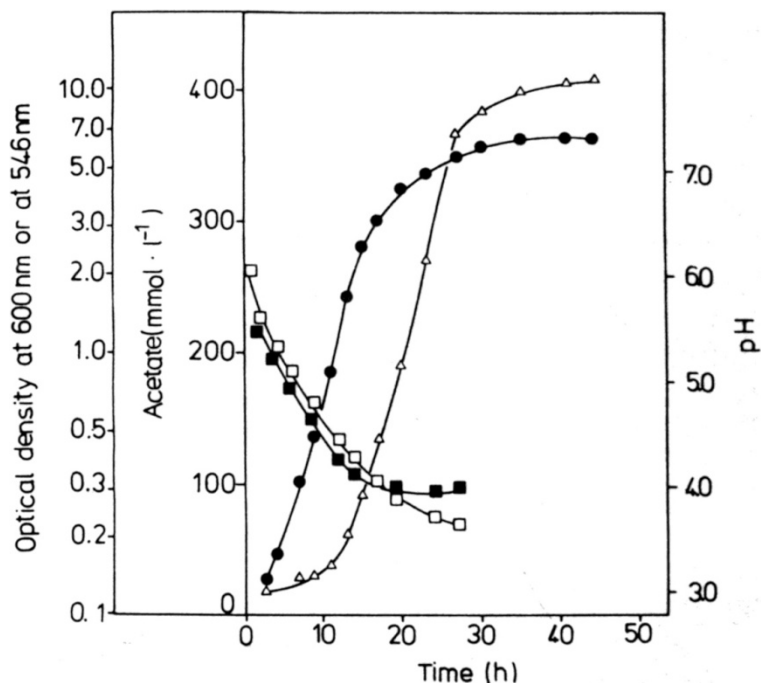
R = ribose 5-phosphate

**Scheme 1.1:** Reaction catalyzed by Class I and Class II PurEs

The identity of  $N^5$ -CAIR as the substrate of Class I PurEs remained obscure until 1994, owing to the unstable nature of  $N^5$ -CAIR (3).  $N^5$ -CAIR and an analogue, in which the ribose 5-phosphate group is replaced with a methyl group, undergo acid-mediated decarboxylation even at neutral pH (3, 5).

The instability of  $N^5$ -CAIR has inspired attempts to detect substrate channeling between PurK and PurE in *E. coli* (6, 7). Channeling has been proposed to occur during the transfer of another acid-labile intermediate in the purine biosynthesis pathway, phosphoribosylamine (8), between two successive pathway enzymes. Recent work in eukaryotic HeLa cells has shown a reversible co-localization of several fluorescently labeled purine biosynthesis enzymes, detected by fluorescence microscopy under periods of purine starvation (9).

*Acetobacter aceti*. *A. aceti* is a Gram-negative, obligately aerobic, rod-shaped bacteria which oxidizes ethanol to acetic acid, and is used in the commercial production of acetic acid. When grown under conditions to optimize the production of acetic acid, the culture in which *A. aceti* is grown can reach levels of acetic acid up to 14% w/v (10). Acetic acid is a membrane-permeable acid, and thus *A. aceti* has evolved to allow for the acidification of its cytoplasm as acetic acid is produced (Figure 1.1) (11). This distinguishes *A. aceti* from many apparently acidophilic bacteria, which resist external acidic conditions, but routinely maintain a cytoplasmic pH near neutrality. In this work the term “acidophilic” will be used to refer to bacteria such as *A. aceti* that tolerate an acidic cytoplasm, while the term “acid resistant” will refer to other bacteria that grow in acidic environments but maintain a relatively neutral cytoplasm. The term “neutralophile” will refer to bacteria and eukaryotes that grow in neutral environments and maintain a cytoplasm pH near neutrality.



**Figure 1.1:** Changes in the external and internal pH during growth of *A. aceti* with ethanol. The intracellular pH was measured with  $[C^{14}]$ -acetylsalicylic acid. Internal pH (filled squares); external pH (open squares); acetic acid (open triangles); optical density at 546 nm (filled circles). Reproduced with permission from (11).

The effect of an acidified cytoplasm raises several interesting questions, among them: how does an acidic cytoplasm affect the stability of acid-labile metabolites like  $N^5$ -CAIR, how is flux through pathways (like *de novo* purine biosynthesis) that form acid-labile metabolites affected, and how are enzymes adapted to function under acidic conditions. It was proposed that enzymes from an acidophile such as *A. aceti* must be adapted to function at low pH (12-14). As was previously posited, there are three non-exclusive ways that enzymes may be adapted for such conditions: (1) an enzyme may be catalytically optimized to function at low pH, (2) a protein may have improved stability at low pH, (3) a protein may work with the aid of chaperone proteins at low pH.

A comparison of crystal structures of proteins from *A. aceti* and from neutralophiles have identified a number of characteristics for proteins from *A. aceti*, including an increased number of hydrogen bonds, alterations in surface charge, or changes in the nature and number of salt bridges. A similar preference is observed when one compares thermophile and mesophile protein crystal structures. This suggests that there may be shared strategies in the design of proteins suited to extreme conditions. Proteins studied thus far from *A. aceti* have not only proven resistant to acid denaturation, but are also relatively resistant to thermal denaturation (12-14). Many extremophile proteins have “accidental” cross-resistance to different destabilizers.

*Citrate synthase.* An ongoing study of the enzyme citrate synthase (CS) from the bacteria *Thermoplasma acidophilum* was furthered by the creation, purification, and characterization of the mutant TM-D317G. This mutant was found to purify with partial occupancy of oxaloacetate, which hindered kinetic characterization of binding. The binding affinity for this mutant, and the parental mutants D317G and D317N were determined under a series of conditions.

## 1.2 Purpose

To explore these critical carbon-carbon bond forming reactions, the enzyme *A. aceti* PurE (*AaPurE*) and mutants of the enzyme *Thermoplasma acidophilum* citrate synthase (*TpCS*) have been characterized. In Chapter 2, studies of the compound CAIR and its corresponding ribonucleoside CAIR-s will be reported, including the change in absorbance in the reaction of CAIR  $\rightarrow$   $N^5$ -CAIR over a range of pH. This information

allowed for the construction of a pH-rate profile for the enzyme *AaPurE* and active mutant *AaPurE*-H59D (Chapter 3). In Chapter 3, the characterization of several mutants of *AaPurE* made to explore the roles of conserved residues His59 and His89 will be reported along with a summary of crystal structures of several mutants, highlighting a possible way that proteins from acidophiles are adapted to low pH. In Chapter 4, a series of stable *E. coli* deletion strains, and insertion strains that replace the chromosomal copies of *purE<sub>Ec</sub>* or *purK<sub>Ec</sub>* with their counterparts from *A. aceti* were made and the growth of these strains examined. In Appendix I, the *TpCS* mutant TM-D317G was created and purified. Characterization of this mutant, and the binding affinity the *TpCS* mutants D317G and D317N for substrate analogues are also reported.

### 1.3 References

1. Kappock, T. J., Ealick, S. E., and Stubbe, J. (2000) Modular evolution of the purine biosynthetic pathway, *Current Opinion in Chemical Biology* 4, 567-572.
2. Firestine, S. M., and Davisson, V. J. (1994) Carboxylases in de novo purine biosynthesis. Characterization of the *Gallus gallus* bifunctional enzyme, *Biochemistry* 33, 11917-11926.
3. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994)  $N^5$ -carboxyaminoimidazole ribonucleotide: Evidence for a new intermediate and two new enzymic activities in the de novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry* 33, 2269-2278.
4. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) Evidence for the direct transfer of the carboxylate of  $N^5$ -carboxyaminoimidazole ribonucleotide ( $N^5$ -CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by *Escherichia coli* PurE, an  $N^5$ -CAIR Mutase, *Biochemistry* 38, 3012-3018.
5. Alenin, V. V., Kostikova, T. R., and Domkin, V. D. (1987) Detection of products of  $N$ -carboxylation of  $N^1$ -substituted 5-aminoimidazoles in aqueous of potassium bicarbonate, *Zhurnal organicheskoi khimii* 57, 692-701.
6. Mathews, I. I., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1999) Crystal structure of *Escherichia coli* PurE, an unusual mutase in the purine biosynthetic pathway, *Structure* 7, 1395-1406.
7. Firestine, S. M., Misialek, S., Toffaletti, D. L., Klem, T. J., Perfect, J. R., and Davisson, V. J. (1998) Biochemical role of the *Cryptococcus neoformans* ADE2 protein in fungal de novo purine biosynthesis, *Archives of Biochemistry and Biophysics* 351, 123-134.
8. Rudolph, J., and Stubbe, J. (1995) Investigation of the mechanism of phosphoribosylamine transfer from glutamine phosphoribosylpyrophosphate amidotransferase to glycinamide ribonucleotide synthetase, *Biochemistry* 34, 2241-2250.
9. An, S., Kumar, R., Sheets, E. D., and Benkovic, S. J. (2008) Reversible compartmentalization of de novo purine biosynthetic complexes in living cells, *Science* 320, 103-106.
10. Waites, M. J., Morgan, N. L., Rockey, J. S., and Higton, G. (2001) *Industrial Microbiology, An introduction*, Blackwell Science.
11. Menzel, U., and Gottschalk, G. (1985) The internal pH of *Acetobacterium wieringae* and *Acetobacter aceti* during growth and production of acetic acid, *Archives of Microbiology* 143, 47-51.
12. Francois, J. A., and Kappock, T. J. (2007) Alanine racemase from the acidophile *Acetobacter aceti*, *Protein Expression and Purification* 51, 39-48.
13. Francois, J. A., Starks, C. M., Sivanuntakorn, S., Jiang, H., Ransome, A. E., Nam, J.-W., Constantine, C. Z., and Kappock, T. J. (2006) Structure of a NADH-insensitive hexameric citrate synthase that resists acid inactivation, *Biochemistry* 45, 13487-13499.
14. Constantine, C. Z., Starks, C. M., Mill, C. P., Ransome, A. E., Karpowicz, S. J., Francois, J. A., Goodman, R. A., and Kappock, T. J. (2006) Biochemical and structural studies of  $N^5$ -carboxyaminoimidazole ribonucleotide mutase from the acidophilic bacterium *Acetobacter aceti*, *Biochemistry* 45, 8193-8208.
15. Mullins, E. A., Francois, J. A., and Kappock, T. J. (2008) A specialized citric acid cycle requiring succinyl-Coenzyme A (CoA): Acetate CoA-transferase (AarC) confers acetic acid resistance on the acidophile *Acetobacter aceti*, *J. Bacteriol.* 190, 4933-4940.

16. Kurz, L. C., Shah, S., Crane, B. R., Donald, L. J., Duckworth, H. W., and Drysdale, G. R. (1992) Proton uptake accompanies formation of the ternary complex of citrate synthase, oxaloacetate, and the transition-state analog inhibitor, carboxymethyl-CoA. Evidence that a neutral enol is the activated form of acetyl-CoA in the citrate synthase reaction, *Biochemistry* 31, 7899-7907.
17. Gu, Z., Drucehammer, D. G., Kurz, L., Liu, K., Martin, D. P., and McDermott, A. (1999) Solid state NMR studies of hydrogen bonding in a citrate synthase inhibitor complex, *Biochemistry* 38, 8022-8031.
18. Kurz, L. C., Ackerman, J. J. H., and Drysdale, G. R. (1985) Evidence from carbon-13 NMR for polarization of the carbonyl of oxaloacetate in the active site of citrate synthase, *Biochemistry* 24, 452-457.
19. Kurz, L. C., and Drysdale, G. R. (1987) Evidence from Fourier transform infrared spectroscopy for polarization of the carbonyl of oxaloacetate in the active site of citrate synthase, *Biochemistry* 26, 2623-2627.
20. Kurz, L. C., Drysdale, G. R., Riley, M. C., Evans, C. T., and Srere, P. A. (1992) Catalytic strategy of citrate synthase: effects of amino acid changes in the acetyl-CoA binding site on transition-state analog inhibitor complexes, *Biochemistry* 31, 7908-7914.
21. Evans, C. T., Kurz, L. C., Remington, S. J., and Srere, P. A. (1997) Correction for: Active site mutants of pig citrate synthase: Effects of mutations on the enzyme catalytic and structural properties, *Biochemistry* 36, 9080-9080.
22. Remington, S. J. (1992) Structure and mechanism of citrate synthase, *Current Topics in Cellular Regulation* 33, 209-229.
23. Bayer, E., Bauer, B., and Eggerer, H. (1981) Evidence from inhibitor studies for conformational changes of citrate synthase, *European Journal of Biochemistry* 120, 155-160.

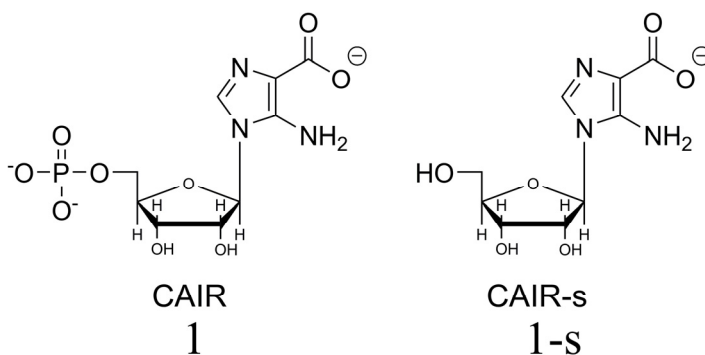


## **Chapter 2**

Non-enzymatic decarboxylation of 4-carboxy-5-aminoimidazole  
ribonucleotide (CAIR) and 4-carboxy-5-aminoimidazole  
ribonucleoside (CAIR-s)

## 2.1 Introduction

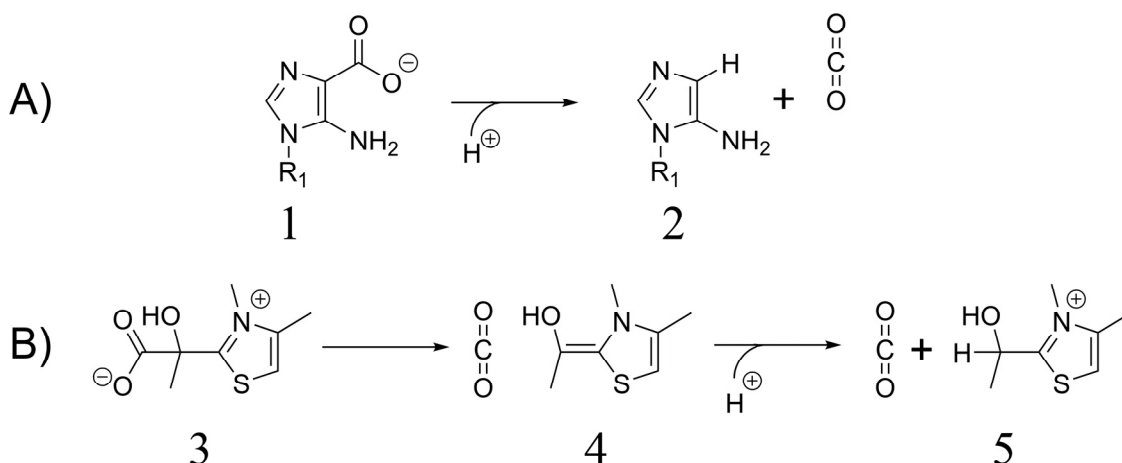
Although recent studies (1-7) have given much insight into PurE-mediated reactions, comparatively few studies have addressed the chemistry of 4-aminoimidazole ribonucleotide (AIR) and carboxylated form (6, 8-11). Understanding the nonenzymatic chemistry of AIR and 4-carboxy-5-aminoimidazole ribonucleotide (CAIR) is essential to a full understanding of enzyme action. Much of the work on the nonenzymatic carboxylation of AIR to CAIR (**1**) and the nonenzymatic decarboxylation of CAIR to AIR was performed over 30 years ago (8-10), long before the correct form of the reaction performed by Class I PurE was known (6). To gain further insight into the reaction mechanism of PurE, additional work on the nonenzymatic reactions of CAIR or its nucleoside CAIR-s (**1-s**) was performed (Figure 2.1). The reactions catalyzed by PurE (Class I or Class II) have an unknown relationship to non-enzymatic decarboxylation.



**Figure 2.1:** Compounds used in this study. The product of PurE CAIR(**1**) and the corresponding ribonucleoside CAIR-s (**1-s**)

It was suggested that PurE catalyzes the decarboxylation of CAIR (Figure 2.2 (A)) or  $N^5$ -CAIR by binding the carboxylate moiety of each in a hydrophobic pocket in the enzyme active site (2-4). This suggestion was based on analogies to other enzymes that

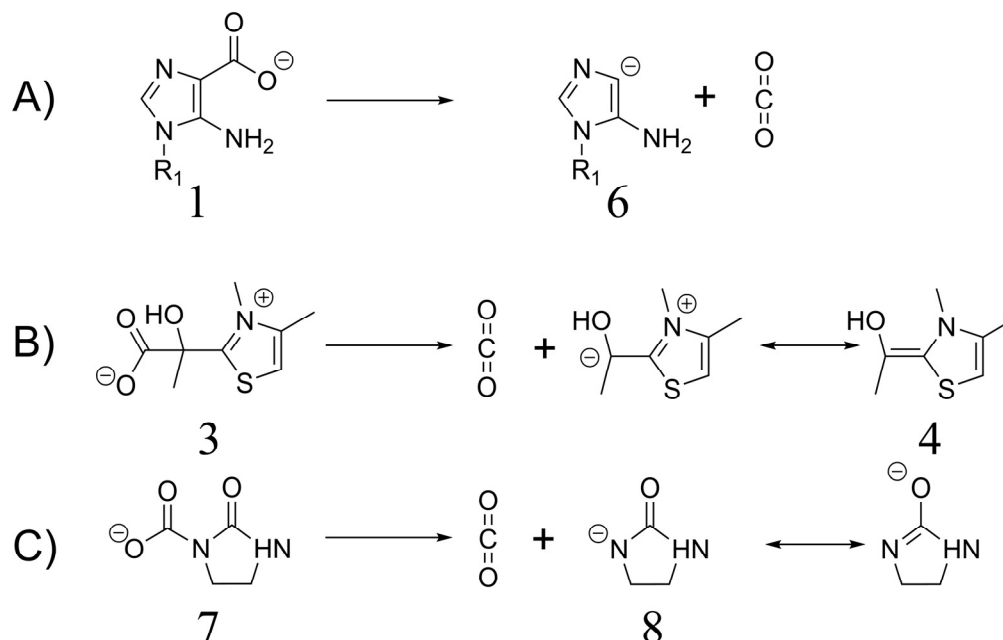
catalyze decarboxylation reactions, including thiamine-dependent pyruvate decarboxylase and biotin-dependent acyl-CoA carboxylases. In the decarboxylases, deprotonated thiamine attacks an  $\alpha$ -keto carboxylic acid. This species then decarboxylates to form a resonance-stabilized intermediate. In the acyl-CoA carboxylases, *N*(1′)-carboxybiotin is produced from biotin, bicarbonate, and ATP. The *N*(1′)-carboxybiotin intermediate is then transferred to another active site where  $\text{CO}_2$  is produced by decarboxylation.



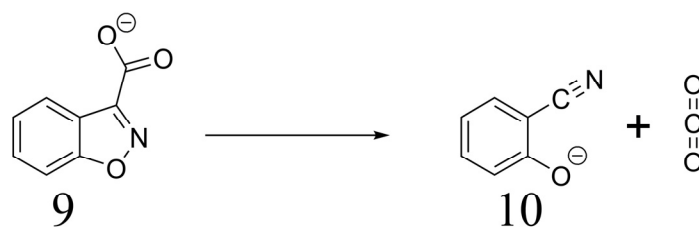
**Figure 2.2:** Nonenzymatic decarboxylation of CAIR (A) and 2-(1-carboxy-1-hydroxyethyl)-3,4,-dimethylthiazolium (B). A) Nonenzymatic decarboxylation of CAIR ( $\text{R}_1$ = ribose 5-phosphate) is studied in this work using CAIR and the corresponding ribonucleoside CAIR-s ( $\text{R}_1$  = ribose). B) Nonenzymatic decarboxylation of 2-(1-carboxy-1-hydroxyethyl)-3,4,-dimethylthiazolium has previously been studied (12, 13), as a model compound of an isolated intermediate from the thiamine containing enzyme pyruvate decarboxylase.

Several studies employing compounds resembling the carboxylated substrates of these enzymes indicate that these enzymes accelerate decarboxylation by sequestration of the carboxylate moiety in an apolar and aprotic environment (12, 14, 15). Included in these studies is the decarboxylation of 2-(1-carboxy-1-hydroxyethyl)-3,4,-dimethylthiazolium (3, Figure 2.2(B)), used as a model for 2-(1-carboxy-1-hydroxyethyl)-

3,4-thiamine pyrophosphate, an isolated intermediate of pyruvate decarboxylase. The rate of decarboxylation of this model compound was found to increase in acetate-buffered solutions of constant ionic strength by a factor of  $\sim 10^4$  as the percentage of ethanol was increased from 0 to 100%. Other experiments with a model for of *N*(1')-carboxybiotin also showed a large increase in the rate of decarboxylation of the anionic species of the model as the fraction of methanol or acetonitrile increases in mixed solutions (relative rates in 100% water, methanol, and acetonitrile were 1, 49, and 517) (15). The large increase in rate for these compounds is attributed to either ground state destabilization or stabilization of the transition state. In the case of thiamine or biotin adducts, the negative charge that results from decarboxylation is resonance stabilized (Figure 2.3). For CAIR, direct decarboxylation would not benefit from such stabilization.

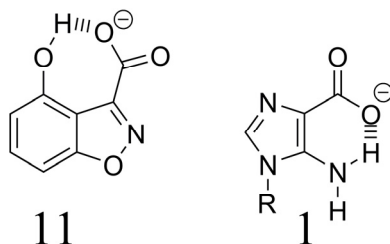


**Figure 2.3:** The direct decarboxylation of CAIR does not result in a resonance stabilized anion (A). Model compounds of substrates for enzymes which are proposed to use carboxylate destabilization result in resonance stabilized anions (B and C).



**Figure 2.4:** Decarboxylation of 3-carboxy-benzisoxazoles.

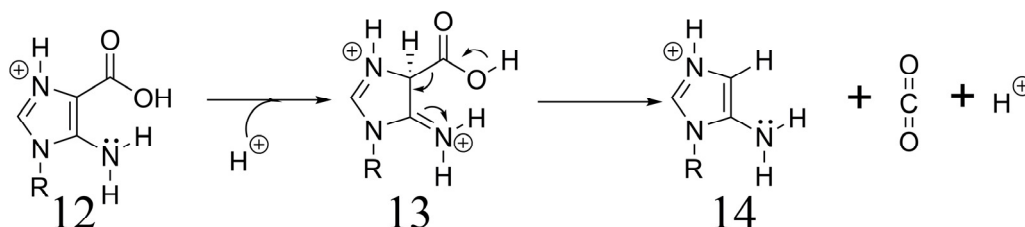
While decreasing solvent polarity has been shown to accelerate the decarboxylation of many compounds, a notable exception is found with 3-carboxy-4-hydroxybenzisoxazole (**11**). While the rate of decarboxylation of other 3-carboxy-benzisoxazoles increases with decreasing solvent polarity (Figure 2.4), **11** maintains a relatively solvent-independent rate of decarboxylation in water, acetonitrile, dimethylsulfoxide, dimethylformamide, and dimethylacetamide (**16**). Studies have suggested that the cause of this solvent-independent rate is an intramolecular hydrogen bond between the hydroxyl and carboxylate groups (**16**, **17**). CAIR (**1**) could form a similar hydrogen bond between the carboxylate and exocyclic amino group (Figure 2.5).



**Figure 2.5:** Intramolecular hydrogen bonds. *Left*, hydrogen bond in 3-carboxy-4-hydroxybenzisoxazoles proposed to cause solvent-independent decarboxylation. *Right*, possible similar hydrogen bond in CAIR.

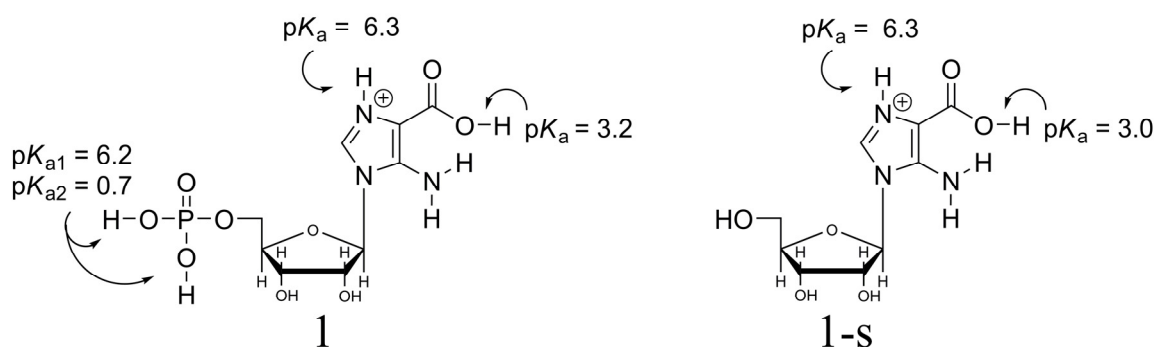
An alternate mechanism for decarboxylation involves ring protonation prior to decarboxylation. A direct ring protonation mechanism has been proposed for many compounds at low pH including anthranilic acid, pyrrole-2-carboxylic acids, and CAIR-s

(Figure 2.6) (18, 19). Protonation at the carbon adjacent to the carboxylate allows for decarboxylation. The rate of decarboxylation for such a mechanism is dependent on the concentration of hydronium ion.



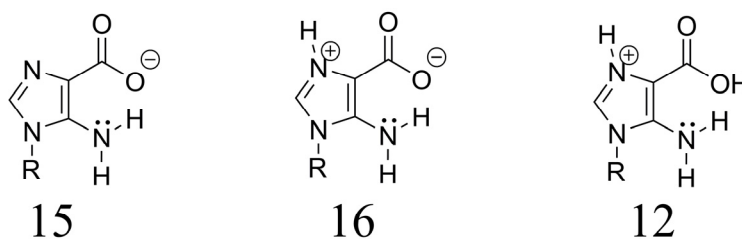
**Figure 2.6:** Ring protonation mechanism for CAIR-s at low pH.

The study of the nonenzymatic decarboxylation of CAIR must account for nucleobase ionizations (Figure 2.7). CAIR undergoes a total of four ionizations, two occurring from the 4-carboxy-5-aminoimidazole ring, and two occurring from the phosphate group of the ribonucleotide. In this study, both CAIR (**1**) and the corresponding ribonucleoside CAIR-s (**1-s**) were used. CAIR-s lacks the sugar phosphate group, and undergoes only two ionizations.



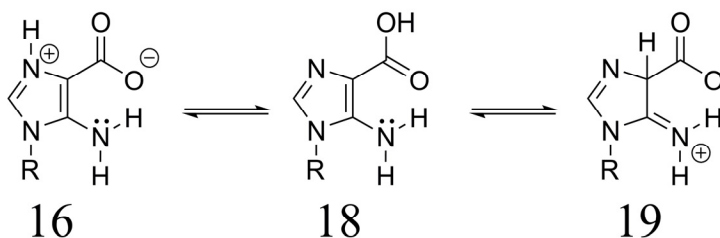
**Figure 2.7:** Ionizations of CAIR (**1**) and CAIR-s (**1-s**). Ionization constants for carboxylate and imidazole ring are values reported for CAIR and CAIR-s at 20°C (20). Ionization constants shown for phosphate group are those reported for the ribonucleotide uridine monophosphate at 25°C (21).

In this work the ionization states of the 4-carboxy-5-aminoimidazole ring will commonly be referred to as anionic (**15**), zwitterionic (**16**), and cationic (**12**) (Figure 2.8). For clarity, these same terms will also be used in reference to CAIR, and will refer specifically to a similar protonation state of the 4-carboxy-5-aminoimidazole ring.



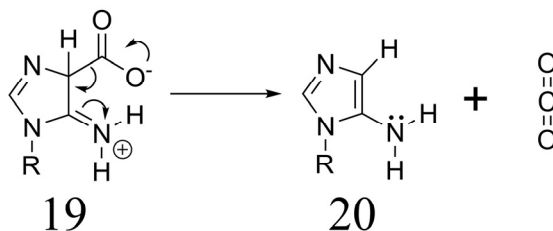
**Figure 2.8:** Illustration of ionization states of 4-carboxy-5-amino-imidazole ring. The terms anionic (**15**), zwitterionic (**16**) and cationic (**12**) will be used in the text to refer to these protonation states of the 4-carboxy-5-amino-imidazole ring (R= ribose 5-phosphate or ribose).

For several ionization states, many tautomers can be drawn. For instance, the zwitterion (**16**) may be in equilibrium with other species (Figure 2.9). Although species **19** may seem implausible, computational studies of the *N*<sup>1</sup>-methyl analogue of CAIR indicated that the C4-protonated form (**19** R = methyl) is only 3.1 kcal/mol higher in energy than the aromatic form (analogous to **16**) (2). Additionally, NMR studies of AIR



**Figure 2.9:** Possible alternate forms of CAIR-s in equilibrium with the zwitterion (**16**).

indicate that the C4 proton exchanges rapidly in D<sub>2</sub>O ( $k_{\text{exchange}} = 0.9 \text{ min}^{-1}$ ) (22). This suggests the intermediacy of **19**. Species **19** may be stabilized by an internal hydrogen bond with the iminium ion. This C4 protonated tautomer appears ideal for decarboxylation, and may provide alternate to direct decarboxylation.



**Figure 2.10:** Possible mechanism of decarboxylation for C4 protonated tautomer.

The previous study of the nonenzymatic decarboxylation of CAIR and CAIR-s was performed by Litchfield and Shaw (8, 9). In a pH rate profile of decarboxylation, they found that the zwitterionic (**16**) and cationic (**12**) forms of CAIR, and CAIR-s decarboxylate, but the anionic form of CAIR (**15**) does not. They also found that the rate of decarboxylation was higher for the zwitterionic form than for the cationic form. They also observed that the rate was accelerated by the presence of anions over the pH range where the zwitterionic form predominates. They also found evidence of catalysis by protons at very low pH, implicating a direct ring protonation at  $\text{pH} < 1$ . They also noted at a pH value where the anion of CAIR (**15**) would predominate, that decarboxylation is catalyzed by formaldehyde.

The last study (8) of the decarboxylation of CAIR and CAIR-s was done prior to several findings that are key to understanding the relevance to the enzymatic system.



These findings include the determination of pH-rate profile for the enzyme suggesting the involvement of an active site acid (Chapter 3) (2, 3) and the relatively energetically accessible C-4 tautomer (2). The previous study was also done at a higher temperature (50 °C) than subsequent enzymatic studies (30 °C). Because the temperature is too high, we performed experiments to explore at physiologically relevant temperature.

To test the theory that PurE may catalyze the decarboxylation of CAIR by sequestration of the carboxylate moiety into a hydrophobic pocket, the effects of solvent polarity on the decarboxylation of CAIR-s were assessed. The effect of deuterium oxide was also assessed because an observation of a solvent isotope effect would support a subset of plausible mechanisms.

Standard enzyme activity assays of PurE are measured in the reverse biosynthetic direction (CAIR  $\rightarrow$   $N^5$ -CAIR) (6), by following the change in absorbance of the reaction mixture using known difference in extinction coefficients ( $\Delta\epsilon_{260}$ ) for each nucleotide at pH 8, 7, or 6. The construction of a pH-rate profile for PurE (Chapter 3) necessitated the determination of  $\Delta\epsilon_{260}$  values for the conversion of CAIR  $\rightarrow$   $N^5$ -CAIR at many pHs. Burst assays were thus designed to measure the  $\Delta\epsilon_{260}$  at all pHs that were used in the construction of the pH rate profile.

This chapter reports the determination of  $\Delta\epsilon_{260}$  of CAIR  $\rightarrow$   $N^5$ -CAIR over a range of pH; these  $\Delta\epsilon_{260}$  will be used in Chapter 3 for the construction of pH-rate profiles of *A. acetii* PurE. Also reported in this chapter are the nonenzymatic rate of decarboxylation of CAIR at several pHs and the activation energy for decarboxylation of CAIR at pH 5.0 or

6.0 in phosphate buffer. The rate of change in absorbance of the ribonucleoside CAIR-s has been determined over a range of pH, and the rate of decarboxylation was estimated by using the  $\Delta\epsilon_{260}$  for the nucleotide when possible. The effect of solvent polarity on the rate of nonenzymatic decarboxylation of CAIR-s was also determined. The study of the decarboxylation of CAIR-s also includes the determination of solvent isotope effect at pH 6.0.

## 2.2 Materials and Methods

*Materials.* Unless otherwise noted, all chemicals were obtained from Sigma or Fisher in the highest purity available. CAIR-s and CAIR were prepared by H. Jiang from 5-aminoimidazole-4-carboxamide ribonucleoside (Toronto Research Chemicals) as described (1, 23). Spectroscopic measurements were recorded on a CARY-100 UV/Vis spectrometer with a thermostated 6X6 Peltier with a CARY temperature controller. Data were analyzed using KaleidaGraph 3.5 (Synergy Software).

*Determination of  $\Delta\epsilon_{260}$  values for CAIR  $\rightarrow$   $N^5$ -CAIR.* Extinction coefficients are known at 260 nm at pH 6, 7, and 8 for AIR and CAIR, which allows the computation of  $\Delta\epsilon_{260}$  values used in PurE activity assays (24). CAIR solutions were standardized by using endpoint assays at pH 8, in which a large amount of AaPurE (typically 66  $\mu$ g) was added to 100 mM Tris HCl pH 8.0 to which small aliquots of CAIR (<0.1 mM) had been added. The  $\Delta A_{260}$  of the resulting burst, corrected for the absorbance of the added enzyme, of product formation was used to compute the concentration of the stock solution. A standardized CAIR solution was then subjected to replicate endpoint assays at 30 °C and

pH values ranging from 4.5 – 9 to determine  $\Delta\epsilon_{260}$  values (Figure 2.1) in a triple buffer system containing 100 mM Tris-HCl, 50 mM MES, and 50 mM acetic acid (25).

*Nonenzymatic decarboxylation of CAIR-s.* The initial rate of change of [CAIR-s] was monitored by the loss of absorbance at 260 nm at 30°C. Experiments were performed in stoppered, masked, 1 cm path length quartz cuvettes containing a total final volume of 1.5 mL. Reactions contained 50  $\mu$ M CAIR-s in 100 mM potassium phosphate buffer. The reaction was initiated by the addition of CAIR-s from a concentrated stock solution to a cuvette containing all other components, the cuvettes had been equilibrated in the cell block for 5 min. The solution in the cuvette was mixed rapidly by inversion to start the reaction. The pH of the buffer was varied in 0.5 pH unit increments from pH 2.0 to 6.0, in 0.2 pH unit increments from pH 6.0 to 8.0, and in 0.5 pH unit increments from 8.0 to 9.0. The change in absorbance at 260 nm was monitored for 5 min. Relative rates in units of ( $\Delta A_{260}/\text{min}$ ) were determined from a linear least squares fit of the data. At pH values for which the  $\Delta\epsilon_{260}$  of  $\text{CAIR} \rightarrow \text{N}^5\text{-CAIR}$  is known, the rate constant for decarboxylation was computed.

*Solvent isotope effect on CAIR-s decarboxylation.* The decarboxylation of CAIR-s was monitored by the loss of absorbance at 260 nm at 30°C. Experiments were performed in stoppered, masked, 1 cm path length quartz cuvette containing a total final volume of 1.4 mL. Reactions contained 54  $\mu$ M CAIR-s and 100 mM potassium phosphate buffer pH 6.8 or 100 mM potassium phosphate buffer pD 6.8 (pH 6.4). The pD values were determined using a pH electrode previously calibrated in standard H<sub>2</sub>O-containing buffers, using the relationship  $\text{pD} = \text{pH} + 0.4$  (26). The estimated

decarboxylation rate was determined from a linear least squares fit of data collected over 5 min, and the  $\Delta\epsilon_{260}$  for CAIR  $\rightarrow$   $N^5$ -CAIR at pH 6.8 ( $7.91 \text{ mM}^{-1}$ ).

*Effect of mixed solvent on decarboxylation of CAIR-s.* Decarboxylation was monitored by the loss of absorbance at 260 nm over longer periods ( $\geq 30$  min) at 30 °C. Experiments were performed in stoppered, masked, 1 cm path length quartz cuvette containing a final total volume of 1.3 or 0.65 mL, depending on the cuvette size. Reactions contained 50  $\mu\text{M}$  CAIR-s, 25 mM sodium acetate, 10.5 mM acetic acid, and 0.94 M LiCl at 30 °C. The concentration of ethanol was varied from 0 to 90% by mixing volumes of 1 M LiCl (aq) and 1 M LiCl (ethanol). For calculation of the percentage ethanol, concentrated buffer (500 mM sodium acetate, 105 mM acetic acid, pH 4.95) and CAIR-s were considered part of the aqueous volume. To determine the rate of decarboxylation for each concentration of ethanol, the absorbance time-dependent data were fit to equation 1 below:  $OD_{260}$  = absorbance at 260 nm,  $OD_{f260}$  = final absorbance at 260 nm,  $\Delta OD_{260}$  = change in absorbance at 260 nm,  $k$  = rate of decarboxylation.

$$OD_{260} = OD_{f260} + \Delta OD_{260} \times e^{(-kt)} \quad \text{Equation (1)}$$

*Arrhenius plot for non-enzymatic CAIR decarboxylation.* The decarboxylation of CAIR was monitored at 260 nm at 30, 40, 50 or 60 °C. Experiments were performed in stoppered, masked, 1 cm path length quartz cuvette containing a total final volume of 1.3 or 0.65 mL, depending on the cuvette size. Reactions contained 50  $\mu\text{M}$  CAIR and either 50 mM potassium phosphate buffer (pH 6.0) or 50 mM potassium acetate buffer (pH 5.0). The ionic strength of each buffer was adjusted with potassium chloride to 0.1

M. Absorbance data were fit to a single-exponential expression. The activation energy for decarboxylation was determined by plotting the natural logarithms of the first order rate constants ( $F$ ) as a function of inverse temperature ( $x$ ) as governed by equation 2, with the variables:  $A$ , defined as the pre-exponential factor and  $E_a$ , defined as the activation energy.

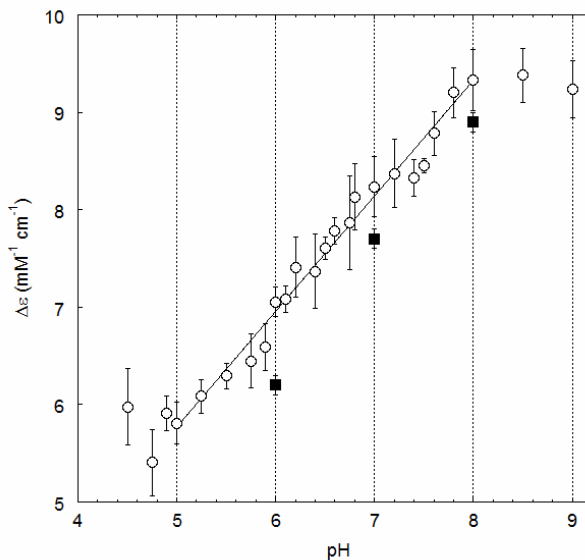
$$F = \ln(A) - \frac{E_a}{R} x \quad \text{Equation (2)}$$

*CAIR decarboxylation in a triple buffer system.* Decarboxylation was monitored by the loss of absorbance at 260 nm at 30 °C. Experiments were performed in stoppered, masked, 1 cm path length quartz cuvettes containing a total final volume of 1.3 or 0.65 mL, depending on the cuvette size. Reactions contained 50  $\mu$ M CAIR when using the 100 mM Tris-HCl, 50 mM MES, 50 mM acetic acid triple buffer system. These determinations were performed at pH 5.0, 6.0, 7.0, 7.5 or 8.0. At each pH, individual trials were also performed at 25  $\mu$ M CAIR to detect CAIR catalyzed decarboxylation or with the addition of 0.1 mM EDTA to detect trace metal-catalyzed decarboxylation. Several trials were run simultaneously, and a holmium standard was included in one cell to account for slight changes in the calibration of the instrument at 260 nm. Absorbance data were fit to a single exponential.

## 2.3 Results

*Determination of  $\Delta\epsilon_{260}$  for  $CAIR \rightarrow N^5\text{-CAIR}/AIR + CO_2$ .* Burst assays allowed the determination of the change in the absorbance of CAIR upon its conversion to  $N^5$ -CAIR and/or AIR (Figure 2.11).  $N^5$ -CAIR is known to decarboxylate rapidly under acidic

conditions to form AIR + CO<sub>2</sub> (estimated to be  $\sim 0.07 \text{ s}^{-1}$  at pH 4.0 and 30 °C) (3). In all cases, the bursts used for quantitation appeared to be monophasic and reached a stable endpoint within 2 min. Assuming some decarboxylation of N<sup>5</sup>-CAIR occurs during this period (especially at lower pH values), the flat endpoint implies there is no substantial difference in  $\epsilon_{260}$  values for N<sup>5</sup>-CAIR and AIR over the pH range used. Data collected from pH 8.0–5.0 follow a linear trend; the data for this range of pH were fit to a linear equation to allow for determination of interpolated extinction coefficients. Outside this range, extinction coefficients were determined at every pH used in the construction of the pH-rate profile of PurE (Chapter 3). The values determined by this method are slightly higher than the values based on published molar absorptivities for CAIR and N<sup>5</sup>-CAIR (6, 24), shown as filled squares in Figure 2.11 .



**Figure 2.11:** Determination of  $\Delta\epsilon_{260}$  for the conversion CAIR  $\rightarrow$  N<sup>5</sup>-CAIR as a function of pH. Mean values for each pH were determined in at least triplicate (open symbols) with the indicated standard deviation. The solid line is a fit of the data in the pH 5.0 - 8.0 interval, which yields the expression  $\Delta\epsilon_{260} = 1.19 \times \text{pH} - 0.15$ . Outside this range, individual  $\Delta\epsilon_{260}$  values were determined. Previously reported values for  $\Delta\epsilon_{260}$  and uncertainties are shown for reference (solid squares)(6, 24).

*Nonenzymatic decarboxylation of CAIR* – The rate of the decarboxylation of CAIR was determined at a series of pH values by monitoring the change in absorbance at 260 nm over time. Reactions were performed at 50  $\mu$ M CAIR. Data were fit to an exponential decay, and the rate constant was extracted (Table 2.1, second column).

The possibility of catalysis by trace metal ions was examined by the addition of 0.1 mM EDTA to one reaction at each pH (Table 2.1, third column). For most pH values, the rate constant for the trial containing EDTA was within 2% of the rates determined for reactions lacking EDTA. This indicates that catalysis by EDTA-chelatable trace metals does not play a significant role.

The possibility that CAIR could serve as a catalyst for decarboxylation was also examined by performing one reaction at a reduced concentration of CAIR (25  $\mu$ M, Table 2.1, fourth column). For all pH values, the rate constant determined at this reduced concentration is within 3% of the rate constant determined at higher concentration. This indicates that autocatalysis does not play a significant role.

pH	Rate constant ( $\text{min}^{-1}$ ) <sup>a</sup>	Rate ( $\text{min}^{-1}$ ) <sup>b</sup> 0.1 mM EDTA	Rate ( $\text{min}^{-1}$ ) <sup>b</sup> Reduced [CAIR]
8.0	0.00162 $\pm$ 0.00002	0.00161	0.00164
7.5	0.0038 $\pm$ 0.0001	0.00420	0.00366
7.0	0.00816 $\pm$ 0.00008	0.00820	0.00818
6.0	0.0210	0.0218	0.0215
5.0	0.0235 $\pm$ 0.0003	0.0239	0.0234

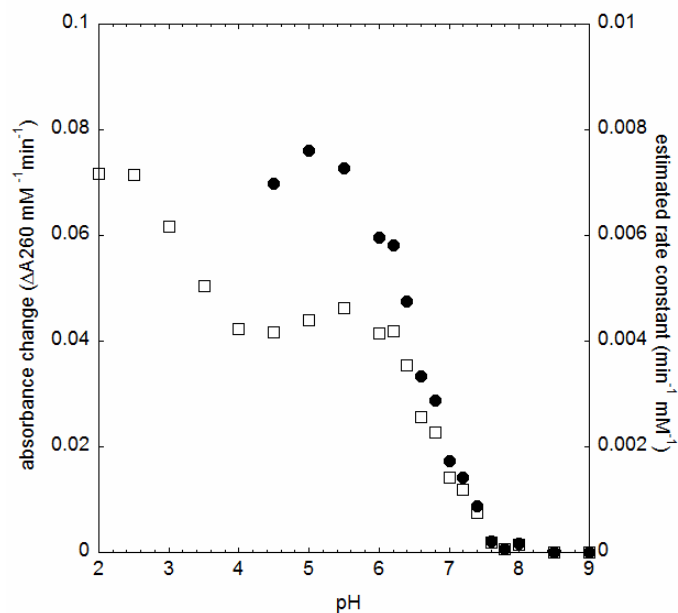
<sup>a</sup>The error shown is standard deviation of rate constants from at least three replicate trials. No error is reported at pH 6.0 as only two trials were performed. Other errors were about 1%. <sup>b</sup>Rate constant determined from a single trial.

*Activation energy for CAIR decarboxylation.* The rate constants for decarboxylation of CAIR at 30, 40, 50, or 60 °C in 50 mM potassium acetate buffer (pH 5.0) or 50 mM potassium phosphate buffer (pH 6.0) were determined. The activation energy at each pH was determined by fitting the natural logarithms of the rate constants to Eqn 2 (Table 2.2). The error reported for each value is based on the uncertainty of the slope at each pH. Values were similar to the previously reported activation energy for CAIR decarboxylation at pH 6.7 ( $20 \pm 2$  kcal/mol) (8).

pH	$E_a$ (kcal/mol)
5.0	$19.1 \pm 0.4$
6.0	$18.9 \pm 0.9$

*CAIR-s decarboxylation.* At basic pH, the rate of change in absorbance at 260 nm ( $\Delta A_{260} \text{ min}^{-1}$ ) for CAIR-s is vanishingly small. However,  $\Delta A_{260} \text{ min}^{-1}$  increases rapidly as the pH is lowered from 8.0 to 5.0 (open squares, Figure 2.12). A further increase in  $\Delta A_{260} \text{ min}^{-1}$  occurs as the pH is lowered from 4.0 to 2.0 (Figure 2.12). These increases are consistent with either a large increase in the rate of decarboxylation or a large change in the differential extinction coefficients for CAIR-s and AIR-s. Since the extinction coefficients for CAIR-s and  $N^5$ -CAIR-s are unknown, the  $\Delta \epsilon_{260}$  determined for the nucleotides in Figure 2.11 were used to estimate the rate constants in the pH range 4.5 to 9.0 (filled circles, Figure 2.12).





**Figure 2.12:** pH-profile of rate of absorbance change at 260 nm for CAIR-s. Initial relative rates of change ( $\Delta A_{260} \text{ min}^{-1} \text{ mM}^{-1}$ ) were determined by a linear fit of absorbance data recorded over 5 min (open squares) from pH 2.0-9.0. Estimated rates of decarboxylation (filled circles) were made using the known  $\Delta \epsilon_{260}$  for  $\text{CAIR} \rightarrow \text{N}^5\text{-CAIR/AIR} + \text{CO}_2$  (Figure 2.1). Reactions contained 100 mM potassium phosphate buffer and 54  $\mu\text{M}$  CAIR-s at 30°C.

*Deuterium isotope effect.* The effect of deuterium oxide solvent on the rate of decarboxylation of CAIR-s was examined (Table 2.3). CAIR-s was found to decarboxylate more slowly in deuterium oxide with a  $k_{\text{H}_2\text{O}}/k_{\text{D}_2\text{O}} = 1.30$ . This is consistent with the involvement of a solvent-exchangeable proton in the transition state for the rate limiting step in decarboxylation. The magnitude of this kinetic isotope effect suggests that a proton transfer is not the sole rate determining step. Alternatively, it is possible that the effect is due to a result of a change in the solvation of the anionic (**15**) and zwitterionic (**16**) forms of CAIR-s. The pH at which the solvent isotope effect has been determined (6.8) is near the  $\text{p}K_a$  of the N3-nitrogen of CAIR-s ( $\sim 6.2$ ). If the anionic form

is preferentially solvated in deuterium oxide, the change in solvation could cause the change in the observed rate because the anionic form is not known to decarboxylate (8).

Solvent	Rate (min <sup>-1</sup> ) <sup>a</sup>
H <sub>2</sub> O	0.00288
D <sub>2</sub> O	0.00221

<sup>a</sup>Determined in 100 mM Potassium phosphate buffer pH 6.8 (pD 6.4). Rate estimated using  $\Delta\epsilon_{260}$  for the reaction of CAIR  $\rightarrow$  N<sup>5</sup>-CAIR

*Mixed solvent experiments.* The effect of solvent polarity on the decarboxylation of CAIR-s was examined (Table 2.3). Increasing the amount of ethanol from 0 to 90% caused an approximately 2.5 fold decrease in the rate of CAIR-s decarboxylation. This is the opposite of the substantial rate increase observed for the model of thiamine-dependent  $\alpha$  keto acid decarboxylation (12, 13). This suggests that that CAIR-s decarboxylates by a different mechanism.

A reduction in rate suggests that the ground state of the decarboxylating species is stabilized in reference to the transition state. However, caution is advised in interpreting this small change due to the significant change in the solvent. The decrease in rate may be due to selective solvation of one or more CAIR-s forms present in solution. The pH of the buffer (4.95) is intermediate between the  $pK_a$  values for CAIR-s, which are 6.2 for the imidazole ring, and 3.0 for the carboxylic acid moiety at 30 °C (20). At pH 4.95, the zwitterion (**16**) is the majority species in aqueous solution; however, that may change as the amount of ethanol is increased and minor tautomers (**18** and **19**) may be favored.

% Ethanol (v/v)	$k_{\text{obs}}$ ( $\text{min}^{-1}$ )
0	0.122
0	0.112
10	0.113
10	0.112
20	$9.78 \times 10^{-2}$
30	$8.88 \times 10^{-2}$
40	$7.60 \times 10^{-2}$
50	$7.21 \times 10^{-2}$
50	$7.55 \times 10^{-2}$
70	$6.11 \times 10^{-2}$
70	$6.55 \times 10^{-2}$
80	$5.53 \times 10^{-2}$
90	$4.68 \times 10^{-2}$

<sup>a</sup>Reactions contained 50  $\mu\text{M}$  CAIR-s, 0.94 M LiCl, 25 mM acetic acid and 10.5 mM sodium acetate at 30°C.

## 2.4 Discussion

*CAIR-s decarboxylation.* The initial rates of change in absorbance at 260 nm for solutions of CAIR-s were monitored over a pH range of 2.0-9.0. This profile shows two increases in the rate of change in absorbance. This indicates either an increase in the rate of decarboxylation, or a change in the molar absorptivity of CAIR-s or AIR-s.

In previous studies of the nonenzymatic decarboxylation of CAIR-s over a range of pH, the rates were monitored by the change in absorbance, and were performed at longer time periods so the data could be fit to a single exponential, and the rate constant determined (8). Those data show an increase in the rate of decarboxylation from pH 7.6 to 5.0 at 50 °C, which is consistent with the increase observed here at 30 °C from pH 8.0 to 5.0 (Figure 2.12). The increased rate of decarboxylation is likely due to

protonation of the imidazole ring of CAIR-s, which has a reported  $pK_a$  of 6.34 at 20 °C (20).

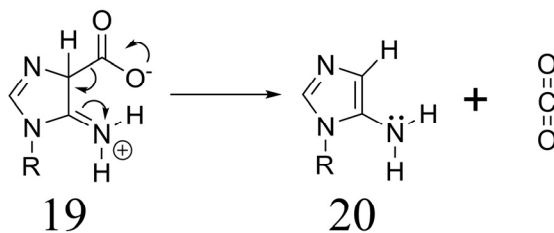
The previous study of CAIR-s decarboxylation at 50 °C shows that the rate constant for decarboxylation decreases as the pH is decreased from pH 5 to pH 2 (8). This indicates that the second increase in the  $\Delta A_{260} \text{ min}^{-1}$  seen at 30 °C in Figure 2.12 from pH 4 to 2 is due to a change in the extinction coefficient of CAIR-s, AIR-s, or both. UV spectra of CAIR and a model compound (4-carboxy 1-cyclohexyl 5-aminoimidazole) at pH  $\sim 0.4$  and  $\sim 3.5$  show an increase the absorbance of both compounds above 255 nm at the more acidic pH (20). The increase in absorbance suggests that the increase seen in from pH 5 to 2 in Figure 2.12 is due to an increase in the extinction coefficient of CAIR-s. The change in extinction coefficient is likely due to the protonation of the carboxylic acid, which has reported  $pK_a$  values of 3.00, 3.21 and 3.18 in CAIR-s, CAIR and a model compound of CAIR in which the ribose has been replaced with a cyclohexyl ring, respectively (20).

*Mixed solvent experiments.* The rate of decarboxylation of CAIR-s is reduced slightly as the amount of ethanol is increased. The change in rate may be due to the effect of increasing ethanol on the ionization state of CAIR-s, or may be due to the effect of increasing ethanol on the energy of the ground state versus that of the transition state of decarboxylation. Similar experiments have shown a large increase ( $\sim 10^4$ ) in the rate of decarboxylation of 2-(1-carboxy-1-hydroxyethyl)-3,4,-dimethylthiazolium (**3**) as the amount of ethanol is increased (12). It seems unlikely that the effect of increasing ethanol on the ionization constants of CAIR-s would be large enough to mask a large

increase in the rate of decarboxylation. The lack of a large increase suggests the decarboxylation of CAIR follows a different mechanism than that of 2-(1-carboxy-1-hydroxyethyl)-3,4,-dimethylthiazolium (12, 13).

*Possible mechanisms of decarboxylation.* As illustrated in Figure 2.12 and in the previous study of CAIR-s (8), the decarboxylation rate increases as the N3 ring nitrogen of the imidazole ring becomes protonated. This indicates that the most rapid decarboxylation is from the CAIR-s zwitterion (16), or a species in equilibrium with it (18 and 19). If a direct ring protonation mechanism were to occur, the rate would not plateau with the population of the zwitterion (16), but would continue to increase as the pH is lowered.

*C-4 protonation mechanism.* One possible mechanism involves the decarboxylation of species 19 (Figure 2.13). This is one of the several possible mechanisms proposed by Litchfield et al (8). In this species, the ring is protonated at C4 instead of N3, and this C4-protonated intermediate is the species that decarboxylates. Species 19 could either be in equilibrium with the zwitterion (16), or be formed from the zwitterion (16) or another species in equilibrium with the zwitterion (18). Regardless the rate of decarboxylation would be proportional to the population of the zwitterion. This mechanism is similar to the amine catalyzed decarboxylation of  $\beta$ -keto acids such as acetoacetic acid and oxaloacetic acid (27, 28).



**Figure 2.13:** C-4 protonation mechanism for CAIR zwitterion.

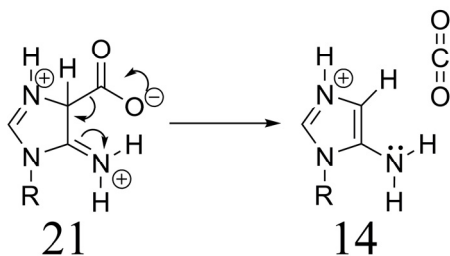
In this view, the observed solvent kinetic isotope effect could be due to partly rate limiting proton transfer to form **19**. In this case, a larger isotope effect might be expected. However, it is not possible to predict what effects the transfer to deuterium oxide would have on the ionization constants of CAIR. A larger isotope effect may be obscured if the anion (**15**) is not as well solvated as the zwitterion (**16**).

If the formation of species **19** is rate limiting, the transfer to less polar solvent would have an uncertain effect. In this mechanism, catalysis by anions could be explained by stabilization of the developing iminium ion promoting protonation at C4.

If the decarboxylation of species **19** is rate limiting, an increase in the rate of decarboxylation would not be expected upon the transfer to a less polar solvent. The transition state for decarboxylation is likely to be more polar than the ground state of the intermediate, as it requires disruption of the hydrogen bond between the iminium and carboxylate. In this mechanism, catalysis by anions could be explained by stabilization of the iminium ion.

At lower pH, the cation of CAIR (**12**) may decarboxylate by a similar C4-protonation mechanism (Figure 2.14). This would result in a species similar to species **19**

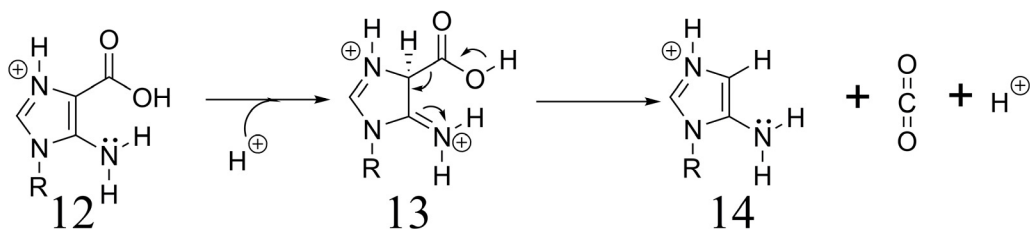
with a proton on either the ring nitrogen or the carboxylate (**20**). Owing to the additional proton, protonation at C4 should be more difficult. The rate of decarboxylation at the pH where **17** is the major species in solution, is lower than the rate seen when the zwitterion of CAIR (**16**) dominates (8).



**Figure 2.14:** C-4 protonation mechanism for decarboxylation of the CAIR cation.

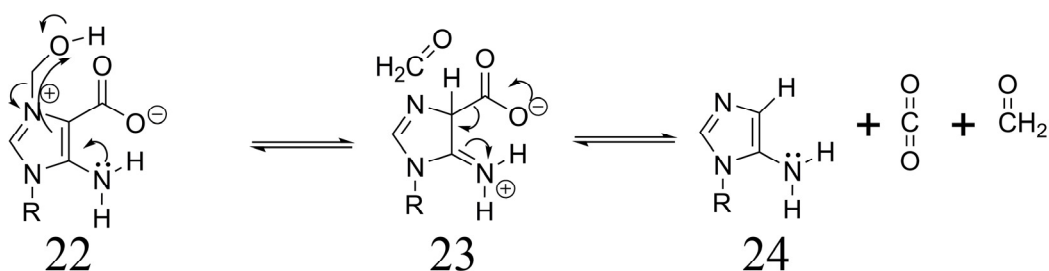
The above C-4 protonation mechanism can be used to explain the decarboxylation of the zwitterion (**16**) and the cation (**12**) of CAIR-s. The C-4 protonation mechanism is consistent with the observed catalysis by anions at moderate pH, the observed solvent isotope effect, and the decrease in the rate of decarboxylation in solutions of increasing ethanol. However, this mechanism cannot be used to explain the previous report of catalysis by protons at low pH (8).

Decarboxylation at very low pH has been proposed to proceed via a direct ring protonation mechanism (8) (Figure 2.15), and will not be discussed. This mechanism is similar to the direct ring protonation mechanism of anthralinic acid (**19**) and pyrrole-2-carboxylic acid (**18**).



**Figure 2.15:** Direct ring protonation mechanism at low pH.

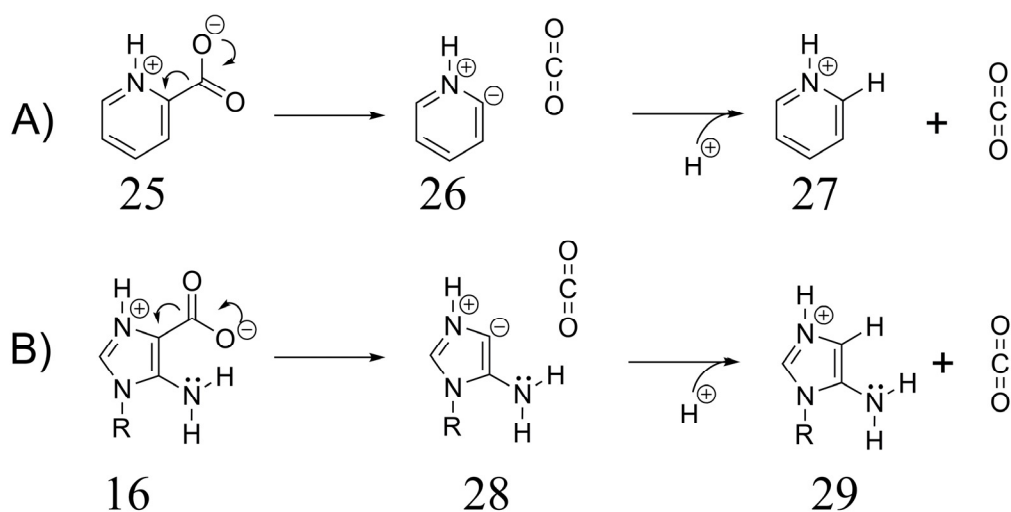
More recently the enzyme PurE was proposed to use a C4 protonation mechanism (2, 3). Of the nonenzymatic mechanisms presented the C4-protonation mechanism appears to be the most consistent with the data collected. The reported observation of catalysis by formaldehyde may also implicate such a mechanism (8). The addition of formaldehyde increases the rate of decarboxylation at pH 8.1 such that the sample completely decarboxylates within minutes at room temperature (8). However, the addition of formaldehyde at 40 °C and pH 4.5 only causes a 10 fold rate increase (8). At 30 °C and pH 8.0 the decarboxylation of CAIR is vanishingly small (Figure 2.12); the rate acceleration by formaldehyde is, therefore, extremely large. One of the *incredible* number of mechanisms that can be written, is for the formation of a carbinolamine that can protonate C4 (Figure 2.16). The authors noted a large elevation in pH upon addition of formaldehyde, which is consistent with the formation of a carbinolamine, among many other possibilities.



**Figure 2.16:** One of the *many* possible mechanisms for formaldehyde decarboxylation. Loss of formaldehyde shown as concerted with protonation at C4.



*Hammick type mechanism.* An alternate mechanism for decarboxylation of CAIR-s is analogous to the Hammick mechanism for the decarboxylation of picolinic acid (Figure 2.17) (29-31). In this mechanism, decarboxylation of the neutral or zwitterionic species (**25**) of picolinic acid produces a nitrogen ylide (**26**). (Note that the ionizations of the carboxylate and ring nitrogen are reversed relative to CAIR. A proton transfer from the carboxylic acid to the ring nitrogen precedes or is concerted with decarboxylation.)

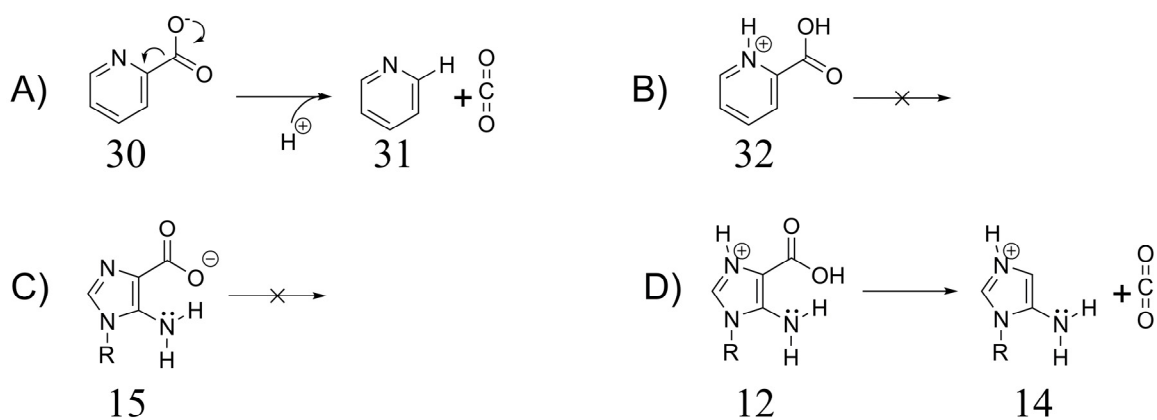


**Figure 2.17:** Hammick type mechanism. A) Hammick mechanism shown for picolinic acid; protonation of the ring nitrogen stabilizes the carbanion that results from decarboxylation. B) Protonation of the N3 ring nitrogen stabilizes the carbanion formed by decarboxylation. The resulting intermediate is then protonated at C4, shown here as the addition of a proton from the solvent.

The maximal rate for a mechanism involving a zwitterion occurs when the population of the zwitterion is highest. This occurs at a pH value that is the average of the two  $pK_a$  values (31). This is seen for the decarboxylation of the CAIR zwitterion (**16**). However, catalysis by anions has been noted in previous studies of CAIR and CAIR-s over the pH range at which the zwitterion dominates (8). In the Hammick-type mechanism,

there is no clear role that an anion or catalytic base could play. This suggests that the zwitterion of CAIR (**16**) decarboxylates by a different mechanism.

Additionally, in the Hammick mechanism, decarboxylation is inhibited when both the carboxylate and ring nitrogen are protonated. Because the cation of CAIR (**12**) is known to decarboxylate (**8**); this suggests that the Hammick mechanism could not apply to the cation of CAIR (**12**).

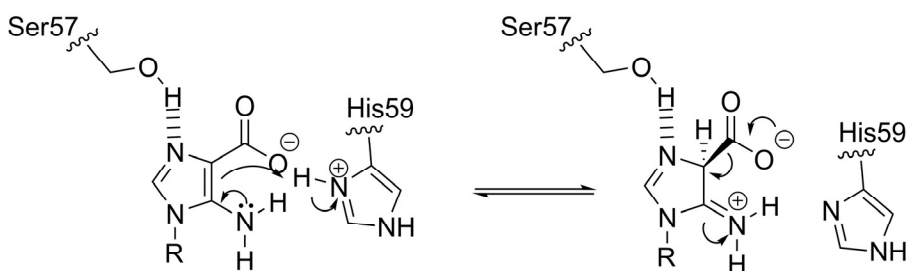


**Figure 2.18:** Hammick type mechanism, part 2. A) The anion of picolinic acid has been shown to decarboxylate (**31**). B) The cation of picolinic acid does not decarboxylate (**31**). C) The anion of CAIR is not known to decarboxylate (**8**). D) The cation CAIR is known to decarboxylate (**8**).

Furthermore, it was also shown that the anion of picolinic acid (**30**) decarboxylates (**31**). The anionic species of CAIR (**15**) does not appear to decarboxylate. If the zwitterion of CAIR-s (**16**) decarboxylates by a Hammick-type mechanism, protonation of the ring nitrogen should be necessary, because the anion of CAIR-s (**15**) does not decarboxylate appreciably at 30 °C. Because the cationic species of CAIR (**12**) is known to decarboxylate (**8**), and the anionic species (**15**) does not, the Hammick mechanism could only apply to the CAIR zwitterion.

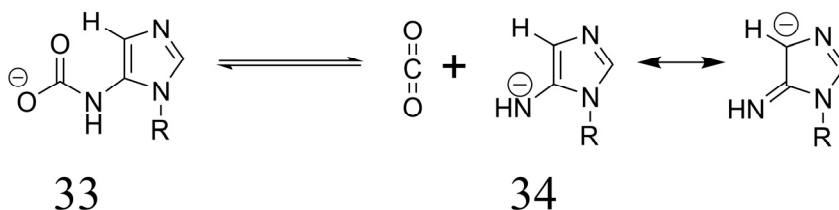
*Implications for mechanism of PurEs.* It has been proposed that Class I PurEs may catalyze the decarboxylation of CAIR and  $N^5$ -CAIR by sequestering the carboxylate in a hydrophobic pocket in the enzyme (2-4, 32). This is similar in manner to the decarboxylations of  $N(1')$ -carboxybiotin and the thiamine-catalyzed decarboxylation of  $\alpha$  keto carboxylic acids. The lack of an increase in the rate of decarboxylation with the transfer to ethanol indicates that a decrease in polarity is not sufficient to accelerate the nonenzymatic decarboxylation of CAIR. This suggests that PurE catalyzes the decarboxylation of CAIR via a different route.

PurE may increase the proportion of a species **19** by directing protonation to C4 (Figure 2.19). Crystal structures of PurE with bound aminoimidazole nucleotides indicate that the active site residue Ser57 (*A. acetii* numbering) forms a hydrogen bond with the N3 of the imidazole ring. Such an interaction would prevent protonation at this position, directing protonation at C4. Crystal structures of PurE with bound nucleotide indicate that the nucleobase is isolated in the active site of the enzyme. We therefore expect there to be an active acid that allows for protonation of C4, and propose that the conserved active site residue His59 as a likely candidate.



**Figure 2.19:** Possible role of PurE Ser57 and His59.

The carboxylate of  $N^5$ -CAIR (**33**), the actual substrate for Class I PurE, lacks the internal hydrogen bond that could prevent direct decarboxylation. Direct decarboxylation would also result in a resonance stabilized anion (**34**) (Figure 2.20). The decarboxylation of  $N^5$ -CAIR (**33**) would therefore be expected to be accelerated simply by sequestration into the hydrophobic pocket. PurE may therefore catalyze the conversion of  $N^5$ -CAIR to CAIR by destabilizing the substrate  $N^5$ -CAIR to yield AIR-azaenolate (**34**) +  $\text{CO}_2$  in the enzyme active site, while providing an environment that lowers the activation energy for the formation of CAIR.



**Figure 2.20:** Decarboxylation of  $N^5$ -CAIR.

In this manner,  $N^5$ -CAIR can be regarded as a  $\text{CO}_2$  carrier similar to  $N(1')$ -carboxybiotin. For both  $N^5$ -CAIR and  $N(1')$ -carboxybiotin, an amine is carboxylated using ATP and bicarbonate, likely through a fleeting carboxy-phosphate intermediate. In a separate active site, both decarboxylate to provide  $\text{CO}_2$ . The enzyme bound  $\text{CO}_2$  is attacked by a nucleophilic carbon in both enzymes. The difference in these two  $\text{CO}_2$  carriers is that biotin remains tethered to the carboxylase multienzyme complex by covalent linkage, while  $N^5$ -CAIR is able to diffuse to the next enzyme, PurE. The possibility of direct  $N^5$ -CAIR transfer, by metabolite channeling, is considered in Chapter 4.

## 2.5 Future Directions

By far the most interesting prospect is to study the formaldehyde catalyzed decarboxylation reaction (8).

The modest solvent kinetic isotope effect (KIE) seen at pH 6.8 is suggestive of a proton transfer or a change in hydrogen bonding during the rate limiting step of CAIR decarboxylation. The possibility that the rate decrease might be due to a change in solvation cannot be ruled out until KIEs are determined at multiple pH values. The magnitude of the KIE will help distinguish the rate limiting step for decarboxylation. For such a study, CAIR-s would likely be the best model as the data will not be complicated by the ionizations of the phosphate group of CAIR.

## 2.6 References

1. Firestine, S. M., and Davisson, V. J. (1994) Carboxylases in de novo purine biosynthesis. Characterization of the *Gallus gallus* bifunctional enzyme, *Biochemistry* 33, 11917-11926.
2. Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, I. I., Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E., and Stubbe, J. (2007)  $N^5$ -CAIR mutase: Role of a CO<sub>2</sub> binding site and substrate movement in catalysis, *Biochemistry* 46, 2842-2855.
3. Constantine, C. Z., Starks, C. M., Mill, C. P., Ransome, A. E., Karpowicz, S. J., Francois, J. A., Goodman, R. A., and Kappock, T. J. (2006) Biochemical and structural studies of  $N^5$ -carboxyaminoimidazole ribonucleotide mutase from the acidophilic bacterium *Acetobacter aceti*, *Biochemistry* 45, 8193-8208.
4. Mathews, I. I., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1999) Crystal structure of *Escherichia coli* PurE, an unusual mutase in the purine biosynthetic pathway, *Structure* 7, 1395-1406.
5. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) Evidence for the direct transfer of the carboxylate of  $N^5$ -carboxyaminoimidazole ribonucleotide ( $N^5$ -CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by *Escherichia coli* PurE, an  $N^5$ -CAIR Mutase, *Biochemistry* 38, 3012-3018.
6. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994)  $N^5$ -carboxyaminoimidazole ribonucleotide: Evidence for a new intermediate and two new enzymic activities in the de novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry* 33, 2269-2278.
7. Firestine, S. M., Misialek, S., Toffaletti, D. L., Klem, T. J., Perfect, J. R., and Davisson, V. J. (1998) Biochemical role of the *Cryptococcus neoformans* ADE2 protein in fungal de novo purine biosynthesis, *Archives of Biochemistry and Biophysics* 351, 123-134.
8. Litchfield, G. J., and Shaw, G. (1971) Part XXXVIII. A kinetic study of the decarboxylation of 5-amino-1- $\beta$ -D-ribofuranosyl imidazole-4-carboxylic acid 5'-phosphate and related compounds., *Journal of the Chemical Society (B)*, 1474-1484.
9. Litchfield, G. J., and Shaw, G. (1965) The mechanism of decarboxylation of some 5-aminoimidazole-4-carboxylic acids and the influence of transition metals, *Chemical Communications (London)* 22, 563-565.
10. Cusack, N. J., Shaw, G. and G. J. Litchfield (1971) Purines, Pyrimidines, and Imidazoles. Part XXXVI. Carboxylation of some 5-aminoimidazoles and related compounds, including nucleosides, with potassium hydrogen carbonate in aqueous solution, *Journal of the Chemical Society (C)*, 1501-1507.
11. Alenin, V. V., Kostikova, T. R., and Domkin, V. D. (1987) Detection of products of  $N$ -carboxylation of  $N^1$ -substituted 5-aminoimidazoles in aqueous of potassium bicarbonate, *Zhurnal organicheskoi khimii* 57, 692-701.
12. Crosby, J., and Lienhard, G. E. (1970) Mechanisms of thiamine-catalyzed reactions. Kinetic analysis of the decarboxylation of pyruvate by 3,4-dimethylthiazolium ion in water and ethanol, *Journal of the American Chemical Society* 92, 5707-5716.
13. Crosby, J., Stone, R., and Lienhard, G. E. (1970) Mechanisms of thiamine-catalyzed reactions. Decarboxylation of 2-(1-carboxy-l-hydroxyethyl)-3,4-dimethylthiazolium chloride, *Journal of the American Chemical Society* 92, 2891-2900.

14. Acevedo, O., and Jorgensen, W. L. (2006) Medium effects on the decarboxylation of a biotin model in pure and mixed solvents from QM/MM simulations, *Journal of Organic Chemistry* 71, 4896-4902.
15. Rahil, J., You, S., and Kluger, R. (1996) Solvent-accelerated decarboxylation of *N*-carboxy-2-imidazolidinone. Implications for stability of intermediates in biotin-dependent carboxylations, *Journal of the American Chemical Society* 118, 12495-12498.
16. Kemp, D. S., Cox, D. D., and Paul, K. G. (1975) Physical organic chemistry of benzisoxazoles. IV. Origins and catalytic nature of the solvent rate acceleration for the decarboxylation of 3-carboxybenzisoxazoles, *Journal of the American Chemical Society* 97, 7312-7318.
17. Acevedo, O., and Jorgensen, W. L. (2005) Influence of inter- and intramolecular hydrogen bonding on Kemp decarboxylations from QM/MM simulations, *Journal of the American Chemical Society* 127, 8829-8834.
18. Dunn, G. E., and Lee, G. K. J. (1971) Kinetics and mechanism of the decarboxylation of pyrrole-2-carboxylic acid in aqueous solution, *Canadian Journal of Chemistry* 49, 1032-1035.
19. Dunn, G. E., and Dayal, S. K. (1970) Mechanism of decarboxylation of substituted anthranilic acids at high acidity, *Canadian Journal of Chemistry* 48, 3349-3353.
20. Litchfield, G. J., and Shaw, G. (1971) Purines, pyrimidines and imidazoles. Part XXXV. Potentiometric and spectroscopic studies of some imidazoles related to intermediates in the biosynthesis *de novo* of purine nucleotides, *Journal of the Chemical Society (C)*, 817-820.
21. Massoud, S. S., and Sigel, H. (1988) Metal ion coordinating properties of pyrimidine-nucleoside 5'-monophosphates (CMP, UMP, TMP) and of simple phosphate monoesters, including D-ribose 5'-monophosphate. Establishment of relations between complex stability and phosphate basicity, *Inorganic Chemistry* 27, 1447-1453.
22. Schendel, F. J. (1986) Ph.D. Thesis, Department of Biochemistry, University of Wisconsin, Madison, WI.
23. Srivastava, P., Mancuso, R., Rousseau, R., and Robins, R. (1974) Nucleoside Peptides. 6. Synthesis of Certain N-[5-Amino-1-( $\beta$ -D-ribofuranosyl)imidazole-4-carbonyl] amino acids related to naturally occurring intermediates in the purine biosynthetic pathway, *Journal of Medicinal Chemistry* 17, 1207-1211.
24. Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J., and Smith, J. M. (1992) Purification and characterization of the *purE*, *purK*, and *purC* gene products: identification of a previously unrecognized energy requirement in the purine biosynthetic pathway, *Biochemistry* 31, 5022-5032.
25. Ellis, K. J., Morrison, J. F., and Daniel, L. P. (1982) [23] Buffers of constant ionic strength for studying pH-dependent processes, in *Methods in Enzymology*, pp 405-426, Academic Press.
26. Lumry, R., Smith, E. L., and Glantz, R. R. (1951) Kinetics of Carboxypeptidase Action. I. Effect of Various Extrinsic Factors on Kinetic Parameters, *Journal of the American Chemical Society* 73, 4330-4340.
27. Pedersen, K. J. (1929) The ketonic decomposition of beta-keto carboxylic acids, *Journal of the American Chemical Society* 51, 2098-2107.
28. Guthrie, J. P., and Jordan, F. (1972) Amine-catalyzed decarboxylation of acetoacetic acid. The rate constant for decarboxylation of a  $\beta$ -imino acid, *Journal of the American Chemical Society* 94, 9136-9141.

29. Dyson, P., and Hammick, D. Ll., (1937) Experiments on the mechanism of decarboxylation. Part I. Decomposition of quinaldinic and isoquinaldinic acids in the presence of compounds containing carbonyl groups, *Journal of the Chemical Society*, 1724-1725.
30. Ashworth, M. R. F., Daffern, R. P., and Hammick, D. Ll. (1939) The mechanism of decarboxylation. Part II. The production of cyanide-like ions from  $\alpha$ -picolinic, quinaldinic, and isoquinaldinic acids, *Journal of the Chemical Society*, 809-812.
31. Dunn, G. E., Lee, G. K. J., and Thimm, H. (1972) Kinetics and mechanism of decarboxylation of some pyridinecarboxylic acids in aqueous solution, *Canadian Journal of Chemistry* 50, 3017-3027.
32. Kappock, T. J., Ealick, S. E., and Stubbe, J. (2000) Modular evolution of the purine biosynthetic pathway, *Current Opinion in Chemical Biology* 4, 567-572.

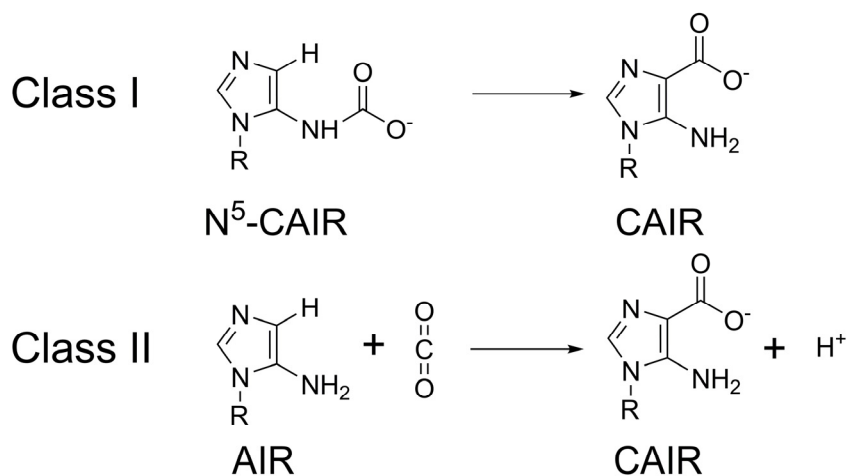


## **Chapter 3**

$N^5$ -Carboxyaminoimidazole Ribonucleotide Mutase from  
*Acetobacter Aceti*

### 3.1 Introduction

Class I and Class II PurEs show a great level of sequence and structural similarity to each other and contain ‘universally’ conserved residues among the classes (1). Each class, however, has ‘class specific’ residues that are conserved and different in each class. Histidine 59 (*A. acetii* numbering) is a universally conserved residue in the “forties loop” of both classes of PurE (“forties loop” refers to *EcPurE* numbering) and has been suggested to serve as an active site acid/base. Histidine 89, however, is a Class I-specific conserved residue. Histidine 89 is part of a four histidine chain that connects the two active sites on opposite faces of the PurE octamer (His 59, His 89, His89’, His59’ where the ‘ denotes a separate subunit). Owing to the proximity of histidine 89 to histidine 59 and the different reactions catalyzed by Class I and Class II PurEs (Scheme 3.1), it has been suggested that His89 could serve as a “proton wire’ between active sites (2).



**Scheme 3.1:** Reactions catalyzed by Class I and Class II PurEs. The substrate of class I PurE contains 1 less proton than the substrated of Class II PurEs.

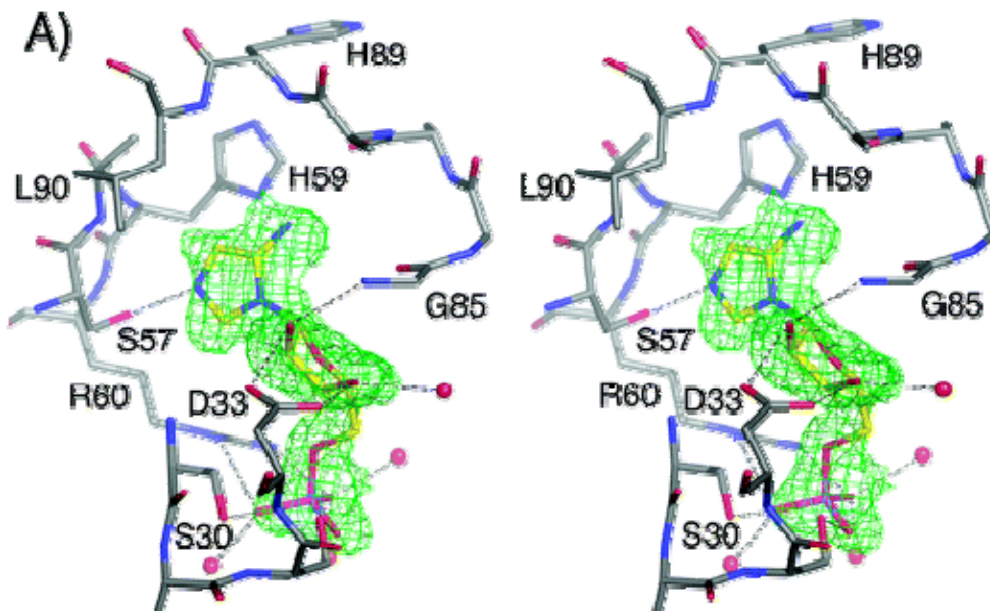
In the E1 component of the pyruvate dehydrogenase complex, a string of polar and ionizable residues has been proposed to serve as a proton wire that allows for the

introduction of a proton into the active site (3). His89 serving as a proton wire could allow for the introduction of a proton into the active site of Class I PurE, joining the reactions catalyzed by both classes. To explore the roles of His59 and His89, a series of mutants was made at these positions, and the mutants were characterized. Crystal structures were determined for many of these mutants by Courtney Starks of this lab.

This lab previously hypothesized that cytoplasmic proteins from *A. acetii* have been subject to evolutionary selection for acid resistance by the acidic cytoplasm routinely experienced by the organism (4-6). This selection could lead to cytoplasmic proteins that are specifically suited for functioning at low pH. Thus, our goal has been to determine if enzymes from *A. acetii* are intrinsically resistant to acid-mediated inactivation, and if they have been catalytically adapted to function optimally at lower pH.

The first reported crystal structure of *AaPurE* (2) had citrate bound in the active site (citrate was present in the crystallization solution). Comparisons to other crystal structures of PurE indicated that citrate interacted with several residues that are involved in nucleotide binding (Figure . The extensive interactions between citrate and the active site indicated that citrate might serve as an active site probe. The residues involved in binding citrate are highlighted below, and the roles the residues play in nucleotide binding are shown in parentheses. The citrate C1 carboxylate (*pro-R* arm) interacts with *AaPurE* Ser30 (bound to ribose-5-phosphate (R5P) moiety phosphate oxygen of AIR), Asp33 (bound to the 2' and 3' hydroxyls of AIR) and Gly85 NH amide

(bound to the 3'-hydroxyl of AIR); the central C6 carboxylate interacts with His59 N<sup>δ</sup>, which has been identified as the active site acid involved in catalysis.



**Figure 3.1:** Crystal structure of *AaPurE* with AIR bound. Several active site residues used in nucleotide binding are also used in binding citrate. Citrate interacts with Ser30, Asp33, Gly85 and His59. Crystal structure determined by Courtney Starks (5).

This chapter reports the biophysical and biochemical characterization of PurE from *A. aceti* and 11 mutants made at residue 59 and 89. These enzymes were characterized kinetically and through active site probes to determine specific roles for residues His59 and His89. A biophysical examination of the stability of *EcPurE* compared to *AaPurE* at a number of pHs was also undertaken. pH-rate profiles constructed for *AaPurE*-wt and an interesting, active mutant *AaPurE*-H59D allowed confirmation that His59 serves as an active site base. Also reported is a proposal for how proteins from an acidophile like *A. aceti* may be optimized to survive at low pH, based on the analysis of several crystal structures of *AaPurE*.

### 3.2 Methods

*Materials and methods* – All materials were from Sigma Aldrich or Fisher Scientific and of the highest purity unless otherwise noted. Hydroxyapatite HTP and Bradford reagent were obtained from Bio-Rad. Restriction endonucleases, Vent DNA polymerase, and T4 DNA ligase were from New England Biolabs. Mutagenesis was performed using Quikchange Mutagenesis kits from Stratagene. Oligodeoxynucleotides (ODNs) were from Integrated DNA Technologies. PD-10 disposable columns were obtained from GE Health Care (formerly Amersham Biosciences). Centrifugation steps were performed using a Beckman Avanti J-20 centrifuge with a JLA-10.5 or JA-20 rotor. Routine activity assays were performed on an Agilent 8453 diode array UV/vis spectrophotometer. Kinetic characterizations including minimal activity assays were performed on a CARY-100 UV/Vis spectrophotometer. Cell disruption by sonication was performed using a Virsonic 100 sonicator. Proteins were routinely concentrated using Amicon ultrafiltration devices with a 10 kDa molecular weight cutoff (YM-10). Aminoimidazole carboxamide ribonucleoside (AICAR) was obtained from Toronto Research Chemicals and used to synthesize CAIR by Hong Jiang of this lab by a published method (7). Protein concentrations were determined by the method of Bradford (8), using bovine serum albumin as a standard. Plasmids were sequenced by Sanger sequencing at the Protein and Nucleic Acid Chemistry Laboratory at Washington University in St. Louis. Protein masses were determined by ESI-MS on a Micromass Q-TOF Ultima quadrupole-TOF mass spectrometer by Washington University Center for Biomedical and Bioorganic Mass Spectrometry.

*Bacterial Strains.* Mutagenesis steps were performed using *E. coli* strains DH5 $\alpha$  and XL-10 Gold. For protein expression, strain BL21(DE3) was employed.

*Construction of plasmids for AaPurE expression.* Plasmid pJK175 was used for the expression and purification of AaPurE, and was previously made by Christopher Mill of this lab (5). Mutant AaPurE constructs for expression were made using Quikchange mutagenesis kits with plasmid pJK175 as template, and the appropriate ODNs listed in Table 2.1 to create plasmids pJK278 (AaPurE-H89N), pJK279 (AaPurE-H89D), pJK280 (AaPurE-H89F), pJK281 (AaPurE-H59F), pJK282 (AaPurE-H59D), pJK283 (AaPurE-H59N),

**Table 3.1** Oligodeoxynucleotides (ODNs) used in this Chapter<sup>a</sup>

ODN	Sequence 5'→3'	Mutation coded(cut site)
1002	GTGGAGCTGCAaATtTGCCGGGTATG	5'-H89N (ApoI)
1003	CATACCCGGCAaATtTGCAGCTCCAC	3'-H89N
1011	GTGGAGCTGCAgATCTGCCGGGTATG	5'-H89D
1012	CATACCCGGCAGATcTGCAGCTCCAC	3'-H89D
1013	GTGGAGCTGCAttTCTGCCGGGTATG	5'-H89F
1014	CATACCCGGCAGAaaTGCAGCTCCAC	3'-H89F
1015	CATTGTTTCAGCAttTCGTACGCCAGAC	5'-H59F
1016	GTCTGGCGTACGAaaTGCTGAAACAATG	3'-H59F
1017	CATTGTTTCtGCAgATCGTACGCCAGAC	5'-H59D (PstI)
1018	GTCTGGCGTACGATcTGCaGAAACAATG	3'-H59D
1019	CATTGTTTCAGCAaATCGTACGCCAGAC	5'-H59N
1020	GTCTGGCGTACGATtTGCTGAAACAATG	3'-H59N
1062	GTGGAGCTGCAGTTCTGCCGGGTATG	5'-H89V
1063	CATACCCGGCAGAACTGCAGCTCCAC	3'-H89V
1064	GTGGAGCTGCAGGTCTGCCGGGTATG	5'-H89G
1065	CATACCCGGCAGACCTGCAGCTCCAC	3'-H89G
1066	CATTGTTTCAGCACAaCGTACGCCAGAC	5'-H59Q
1067	GTCTGGCGTACGtTGTGCTGAAACAATG	3'-H59Q
1068	CATTGTTTCAGCAgcTCGTACGCCAGAC	5'-H59A
1069	GTCTGGCGTACGAgcTGCTGAAACAATG	3'-H59A
1070	CATTGTTTCAGCAagTCGTACGCCAGAC	5'-H59S
1071	GTCTGGCGTACGActTGCTGAAACAATG	3'-H59S

<sup>a</sup>Letters shown in lower case code for mutagenesis. In some cases, a silent mutation was also included to create a restriction site for screening plasmids.

pJK314 (*AaPurE*-H89V), pJK315 (*AaPurE*-H89G), pJK318 (*AaPurE*-H59Q), pJK319 (*AaPurE*-H59A), and pJK320 (*AaPurE*-H59S).

*Overexpression and purification of AaPurE.* Glycerol stocks of BL21(DE3) cells transformed with pJK 175 were streaked on LB plates supplemented with 100 µg/mL ampicillin and were grown overnight at 37 °C. A single colony was used to inoculate a starter culture of LB broth, supplemented with ampicillin as above, that was grown to saturation. Starter cultures were diluted 1:50 into fresh media and the cultures grown at 37 °C to an  $A_{600} = 0.6$ , and induced by the addition of IPTG to 0.4 mM. After 4 h of induction, cells were harvested by centrifugation (10000g, 20 min). Cell pellets were then frozen at -80 °C until used. All subsequent steps were performed at 4 °C.

Cell pellets were resuspended in 50 mM Tris-HCL, pH 8.0 (6 mL/g of cells). Cells were disrupted by sonication, and debris was removed by centrifugation (37000 g, 20 min). Streptomycin was added to the lysate to 1% (w/v) from a 10% (w/v) stock and the lysate gently mixed for 15 min on ice. Solids were removed by centrifugation (37000g, 60 min). The resulting supernatant was then adjusted to 35% saturation by the slow addition of solid ammonium sulfate over 30 min. Following an additional 30 min, solids were collected by centrifugation (37000 g, 10 min).

The 35% ammonium sulfate pellet was redissolved in a minimal volume of 10 mM potassium phosphate buffer, pH 8.0. The solution was then desalted in 2.5 mL portions using a PD10 column equilibrated in 10 mM potassium phosphate buffer, and applied to a hydroxyapatite column (2.5 X 1.5 cm) equilibrated in 10 mM potassium phosphate buffer. The column was washed with the same buffer, and *AaPurE* was

found in the flow through fractions. These fractions were pooled and concentrated to >5 mg/mL using Amicon ultrafiltration devices with a YM10 membrane. The concentrated protein solution was then stored in small single use aliquots at -80 °C.

*Overexpression and purification of mutant forms of AaPurE.* AaPurE mutants were isolated by using the same procedure as described above for wild type AaPurE. To avoid cross contamination, each mutant was purified using its own new column material and ultrafiltration devices. Mutants were concentrated to >5 mg/mL using Amicon ultrafiltration devices with a YM10 membrane. The concentrated protein solution was then stored in small single use aliquots at -80°C.

*Thermal denaturation of E. coli PurE monitored by circular dichroism (CD).* Data were obtained on a Jasco J-500A spectropolarimeter. A jacketed cylindrical quartz cell with a 0.1 cm path length was heated from 24 °C to >75 °C in 1° or 2 °C steps (smaller steps were taken near the melting transition), allowing 200 s for equilibration at each temperature. The loss of structure of EcPurE was followed by monitoring the ellipticity of the protein at 222 nm at 0.23 mg/mL EcPurE, 50 mM potassium phosphate buffer and 100 mM KCl at pH 4.0, 5.0, and 7.0. A trial at pH 3.5 indicated the protein unfolded at the starting temperature (24 °C). The melting temperature ( $T_m$ ) for each pH was determined by the inflection point of each progress curve.

*Carboxylic titrations of AaPurE.* Spectra were obtained on a Spex FluoroMax fluorimeter equipped with a thermostated stirred cell holder using a 1 cm fluorimeter cell at 25°C (the temperature inside cuvette was measured using a thermocouple). Titrations were performed in a starting solution of 2.0 mL 50 mM potassium acetate, pH



5.2, 100 mM KCl and 3  $\mu$ M AaPurE. The emission spectra was recorded from 300 to 420 nm (5 nm slit width, 0.5 s integration, 1 nm intervals), exciting at 295 nm (1 nm slit width). The intensity of  $\lambda_{\max}$  of the unliganded protein was then measured after small additions of a 100 mM stock solution of carboxylic acids listed in Table 2.4. At the end of each titration the pH of the solution was determined. Data were corrected for dilution and inner-filter effect when needed. Fluorescence data ( $F$ ) were fit as a function of  $C$ = [titrant] to three parameters using equation 1 below where:  $K_d$  = dissociation constant,  $F_0 = F$  prior to addition of titrant, and  $\Delta F$  = maximal change in  $F$ .

$$F = F_0 - \Delta F \frac{C}{C + K_d} \quad (1)$$

*Citrate pH-affinity profile.* Titrations with citrate were performed as above at a number of pHs from pH 3.7 – 6.2 in 50 mM potassium acetate buffer, 100 mM KCl, and 3  $\mu$ M AaPurE, and the value of  $K_{eq}$  was determined, ( $K_{eq} = 1 / K_d$ ). These values were plotted as a function of pH and fit to equation 2.

$$\log (K_{eq}) = \log \left( \frac{K_{eq}^{max}}{\left[ \left( \frac{10^{-pH}}{10^{-pK1}} \right) + \left( \frac{10^{-pK2}}{10^{-pH}} \right) + 1 \right]} \right) \quad (2)$$

*Enzyme activity assays.* All active forms of AaPurE were assayed in the reverse direction (CAIR  $\rightarrow$   $N^5$ CAIR) by using the published method at 100  $\mu$ M CAIR(9). Activities for all active forms (see below) of AaPurE (except H59S, the activity of which was too low for characterization) were determined in 50 mM Tris-HCl pH 8.0, in a total volume

of 0.5 mL, in a stoppered 1 cm pathlength cuvette. For Michaelis-Menten plots, initial rates were determined at varying CAIR concentration. At higher concentrations of CAIR, a 0.5 cm pathlength cuvette was used. Reactions were initiated by the addition of enzyme after a background rate of decarboxylation had been determined. The initial rate of CAIR conversion (<10% conversion of CAIR) was determined at 260 nm, and plotted as a function of CAIR concentration. Kinetic constants were obtained by non-linear least squares fitting to the Michaelis-Menten equation by using KaleidaGraph 3.6 (Synergy software).

*Determination of maximal activities for inactive AaPurE mutants.* Standard activity assays were altered to contain up to 0.9 mg/mL of protein in an attempt to detect any activity in mutants that showed no activity during the purification steps. For AaPurE-H59N, this procedure was also attempted in the three buffer system at pH 5.0, 6.0, 7.0 and 8.0. The maximal possible activity was calculated as twice the uncertainty of the slope in  $A_{260}$  over time.

*pH-rate profiles.* The initial rates of reaction for AaPurE-wt and AaPurE-H59D were determined as above except that a triple buffer system was used at many pHs, using the values of  $\Delta\epsilon_{260}$  of CAIR  $\rightarrow$  N<sup>5</sup>-CAIR determined for each pH as described in Chapter 2. Reactions contained 100 mM Tris-HCl, 50 mM MES, 50 mM acetic acid, variable amounts of CAIR (0.4-400  $\mu$ M) and either 0.54-4.3 nM AaPurE-wt or 36.5-365 nM AaPurE-H59D. Data at each pH were plotted as a function of [CAIR], and the data were fit to the Michaelis-Menten equation by nonlinear least squares fit. Values of  $k_{cat}$  and  $k_{cat}/K_m$  were plotted as a function of pH and fitted by nonlinear least squares fit to

equation 3, or equation 4 in the case of  $A\alpha$ PurE-wt  $k_{\text{cat}}/K_m$ , which accounts for three ionizations, two of which are indistinguishable.

$$\log y = \log\left[\frac{c}{1 + \frac{H}{K_1} + \frac{K_2}{H}}\right] \quad (3)$$

$$\log y = \log\left[\frac{c}{1 + \left(\frac{H}{K_1}\right)^2 + \frac{K_2}{H}}\right] \quad (4)$$

*CAIR titration of inactive mutants.* The ligand-binding affinity of  $A\alpha$ PurE mutants H59N, H59F and H59A were monitored in titrations detecting changes in intrinsic tryptophan fluorescence. Titrations initially contained 50 mM Tris-HCl, pH 8.0 and 3  $\mu$ M  $A\alpha$ PurE. The emission spectrum was recorded from 300 to 420 nm (5 nm slit width, 0.75 s integration at every nanometer), exciting at 295 nm (1 nm slitwidth). Spectra were recorded following small additions from a titrant solution of 100  $\mu$ M CAIR in 50 mM Tris-HCl pH 8.0. After corrections for dilution, fluorescence intensity data ( $F$ ) were fit as a function of the known values  $C = [\text{CAIR}]$ , and  $E = 0.5$  [subunit] to three parameters using equation 5 below where:  $K_d$  = dissociation constant,  $F_0 = F$  prior to addition of titrant, and  $\Delta F$  = maximal change in  $F$ . Half sites binding was used because it yielded a superior fit to data.

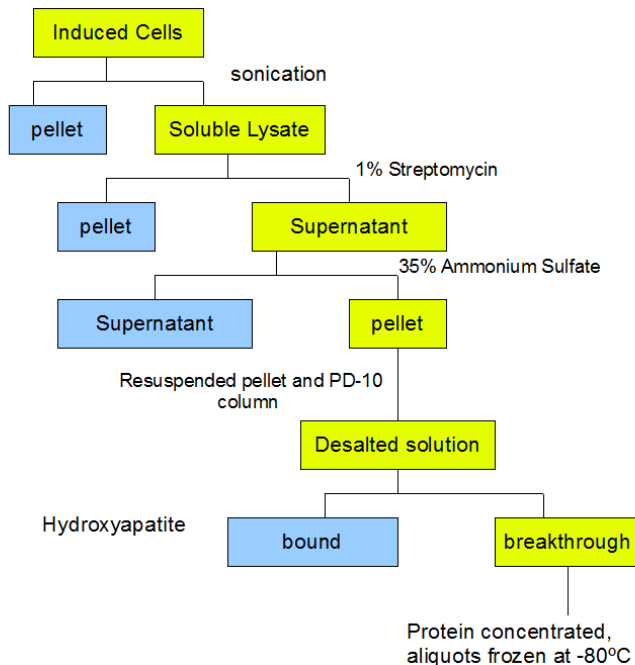
$$F = F_0 - \frac{\Delta F}{2} \left[ (C + E + K_d) - \sqrt{(C + E + K_d)^2 - 4EC} \right] \quad (5)$$

*Acid Stability Assays.* In an adaptation of a previously described method for measuring the relative acid stability of *A. acetii* proteins were compared to mesophile proteins (4); *AaPurE* and *EcPurE* were incubated at 30 µg/mL in the three buffer system at 30 °C at pH 3.0 and 3.5, as well as in 25 mM potassium phosphate buffer with KCl added to an ionic strength of 0.1 M at pH 2.5. Aliquots were removed and assayed for activity in a standard reverse direction activity assay.

### **3.3 Results**

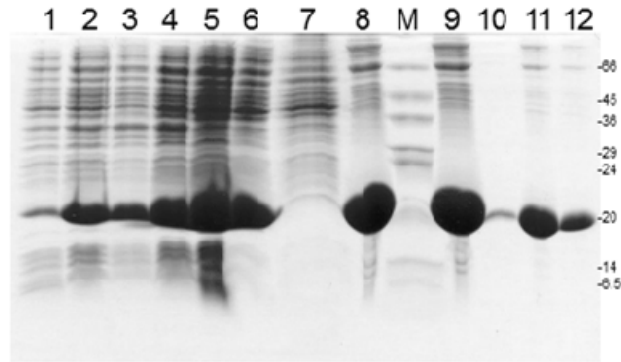
*Creation of AaPurE mutants.* Mutants of *AaPurE* were made to investigate the roles of universally conserved residues His59 and Class I conserved residue His89. Quikchange mutagenesis allowed for the rapid construction of plasmids for expressing these mutants. Sequencing indicated that all of the intended mutants were obtained, usually in the first attempt. The successful mutations were obtained without the need to screen the plasmids using the silent restriction sites that had been added in some cases.

*Purification of Recombinant AaPurE.* The gene coding for *AaPurE* was originally isolated in plasmid pJK173 by Christopher Mill of this lab, by functional complementation of the *E. coli purE* auxotroph PC0135. This indicated that *AaPurE* is expressed as an active protein in *E. coli*. For purification, *AaPurE* was expressed in BL21(DE3) cells using plasmid pJK175. The purification of *AaPurE* (Scheme 3.2), followed standard steps for the

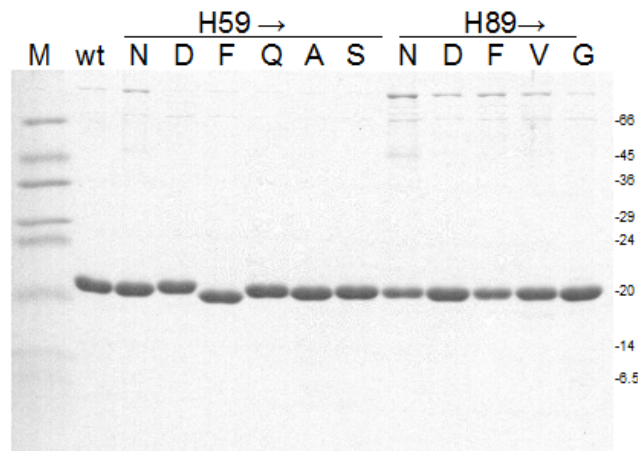


**Scheme 3.2:** Purification scheme for *AαPurE*

disruption of cells (sonication) and removal of polynucleotides by streptomycin precipitation. Ammonium sulfate (AS) fractionation at 35% AS saturation provided a quick means to enrich *AαPurE* as evidenced by the increase in specific activity (Table 3.2). After removing salts and small molecules using a PD-10 size exclusion column, several contaminating proteins were removed by applying the protein solute to a hydroxyapatite column. Flowthrough fractions containing protein, including *AαPurE*, were pooled and concentrated using an Amicon ultracentrifugation device (YM-10). The same purification scheme was successfully used for all mutants. A representative gel analysis of one purification is shown in Figure 3.2. A gel analysis of all concentrated final products is shown in Figure 3.3.



**Figure 3.2:** SDS-PAGE analysis (15% polyacrylamide) of an *AaPurE* purification. Shown above is the purification of mutant *AaPurE*-H89G. *Lane 1*, Uninduced cells; *Lane 2*, Cells after 4 h induction at 400  $\mu$ M IPTG; *Lane 3*, crude lysate; *Lane 4*, insoluble lysate; *Lane 5*, soluble lysate; *Lane 6*, 1% streptomycin supernatant; *Lane 7*, 35% Ammonium sulfate supernatant; *Lane 8*, 35% ammonium sulfate pellet; *Lane 9*, PD10 elutant; *Lanes 10-12*; flowthrough fractions from hydroxyapatite column. Flow through fractions from hydroxyapatite column were pooled and concentrated as described.



**Figure 3.3:** SDS-PAGE analysis of purified *AaPurEs*. (15% polyacrylamide, 3  $\mu$ g of each protein). Sigma low range marker is shown in the first lane, and molecular weights associated with each are shown on the right in kDa. *AaPurE*-wt is denoted by the abbreviation wt. All mutants are denoted by the corresponding one letter code above each lane.

The purification of wild type and mutants yielded final products of similar purity (Figure 3.3). The activity of purifications (when activity was detected during

purification) was followed, and the specific activity of each step determined (Table 3.2).

The identities of proteins were further confirmed by ESI-MS (Table 3.3).

**Table 3.2:** Purification chart for mutants with detectable activity throughout purification.

	Wt	H59D	H89N	H89D	H89F	H89V	H89G
Soluble	355 <sup>a</sup>	109	68.8	76.8	72.6	62.9	66.8
Lysate	(16.4) <sup>b</sup>	(0.14)	(0.022)	(0.61)	(0.75)	(1.5)	(0.72)
Strepto- mycin	268 (22.4)	79.4 (0.22)	45.8 (0.031)	57.9 (0.87)	46.0 (1.4)	39.5 (2.3)	47.3 (0.94)
35% AS pellet	132 (35.4)	50.0 (0.31)	12.8 (0.091)	19.8 (3.2)	10.8 (4.3)	15.3 (4.9)	27.5 (1.6)
PD10 column	107 (38.6)	50.1 (0.30)	11.0 (0.099)	13.6 (1.9)	7.2 (5.8)	14.0 (5.4)	24.9 (1.7)
Hydroxy- apatite	86 (38.8)	39.1 (0.33)	9.1 (0.11)	14.5 (1.7)	6.7 (5.8)	11.3 (5.7)	19.5 (1.6)

<sup>a</sup>Amount of protein in milligrams <sup>b</sup>specific activity of protein determined Units/mg.

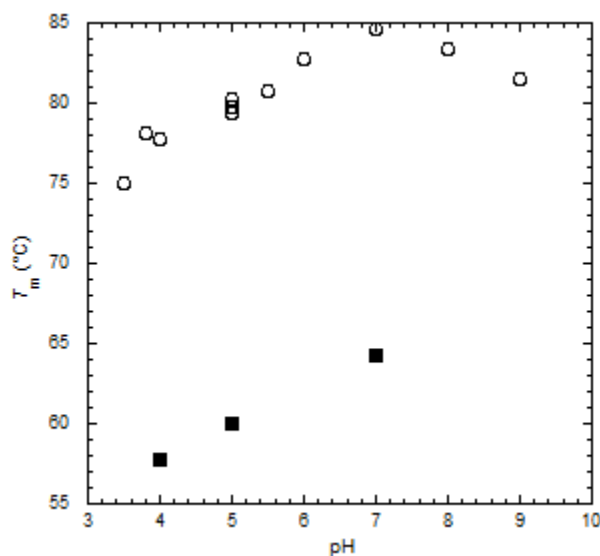
**Table 3.3:** Results of ESI-MS of purified wild type and mutant A $\alpha$ PurEs

	MW (Da)	Expected (Da)
wild type	18865 $\pm$ 2	18865.8
H59N	18842 $\pm$ 2	18842.7
H59D	18844 $\pm$ 2	18843.7
H59F	18875 $\pm$ 2	18875.8
H59Q	18856 $\pm$ 2	18856.8
H59A	18799 $\pm$ 2	18799.7
H59S	18815 $\pm$ 2	18815.7
H89N	18841 $\pm$ 2	18842.7
H89D	18844 $\pm$ 2	18843.7
H89F	18875 $\pm$ 2	18875.8
H89V	18828 $\pm$ 2	18827.8
H89G	18785 $\pm$ 2	18785.7

Samples were diluted in water, bound to an inline C18 guard column, and desalted with water. After a solvent change via a 1 min gradient to 1% acetonitrile 0.1% trifluoroacetic acid and the protein eluted in a gradient to 60% acetonitrile 0.1% fluoroacetic acid on a Micromass Q-ToF Ultima with Waters CapLC. Spectra were decharged to give a MW by using MaxEnt.

*Acid stability studies.* *A. aceti* is an acidophilic bacterium that tolerates an acidic cytoplasm (10), and should therefore contain proteins suited to functioning at low pH.

Previously, the acid stability of proteins from *A. aceti* was characterized in reference to a similar protein from a mesophile (4, 6). The thermal denaturation of *EcPurE* under the same conditions as *AaPurE* indicated that *EcPurE* is far less thermostable at all assayed pHs (Figure 3.4). Attempts to compare the acid stability of *AaPurE* and *EcPurE* by monitoring the loss of enzyme activity when the enzymes were incubated under acid conditions failed. Upon transfer to standard activity assays (pH 8.0), the activities for both enzymes were found to increase while being assayed. This indicated that both unfolded proteins may partially refold in the assay buffer. This prevented the determination of a rate of inactivation for either PurE.

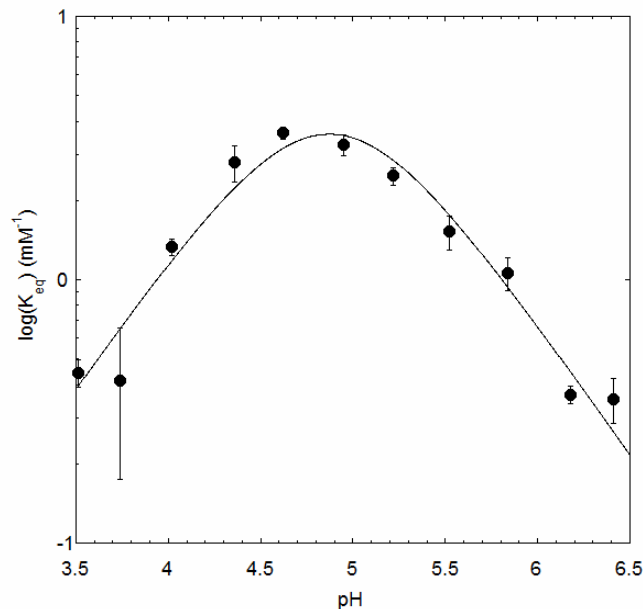


**Figure 3.4:** Comparison of *E. coli* and *A. aceti* PurE  $T_m$  values. Melting temperatures for proteins ( $T_m$ ) were determined following circular dichroism at 222 nm to monitor thermal denaturation of protein.  $T_m$ s were determined as a function of pH, *EcPurE* (filled squares) was found to be far less thermostable than *AaPurE* (open circles) at all pHs assayed. *EcPurE* unfolded at the starting temperature of 24°C at pH 3.5 (not shown).

*Carboxylic acid titrations of AaPurE.* Owing to the extensive interactions between citrate and the *AaPurE* active site, fluorescence titrations were performed to determine



if a citrate complex might form in vivo. The binding of citrate leads to a decrease in the intrinsic fluorescence of *AaPurE*. This decrease was used to determine the binding affinity of citrate for *AaPurE* by following the decrease in fluorescence emission intensity as a function of citrate concentration. The resulting data were then fit to equation 1. The binding of citrate to *AaPurE* was determined over a broad range of pH. The pH was measured at the start and at the end of each titration. In no case did this value differ from the starting value by more than 0.11 pH units. A pH-affinity profile was created by plotting the citrate affinities as a function of pH (Figure 3.5). This yields a bell-shaped profile, with an optimum near pH 4.6. Both the basic and acidic arms of the profile have a slope of 0.9, indicating that a single protonation is involved in determining each arm of the profile. This is consistent with the binding of the citrate mono- or di- anion ( $pK_{a1}= 3.13$ ;  $pK_{a2}= 4.76$ ;  $pK_{a3}= 6.80$ )(11) assuming that one of the ionizations of the profile derives from citrate. The data collected at each pH were fit to equation 2 to determine apparent  $pK_a$  values of the profile.



**Figure 3.5:** pH dependence of wt AaPurE citrate affinity ( $K_{eq}$ ). Binding equilibrium constants ( $K_{eq}$ ) were determined at 0.3 pH unit increments from pH 3.4 to 6.4. Fluorescence titrations were performed at 25 °C using AaPurE-wt. The solid line represents a fit of the  $\log(K_{eq})$  to equation 2, with  $pK_1 = 5.1 \pm 0.4$  and  $pK_2 = 4.7 \pm 0.4$  and a maximum pH independent  $K_{eq} = 20 \pm 10 \text{ mM}^{-1}$ . The strongest affinity was recorded at pH 4.6 =  $3.6 \pm 0.2 \text{ mM}^{-1}$  ( $K_d = 0.28 \pm 0.02 \text{ mM}$ ).

The fluorescence quenching assay was used to screen several other common citric acid cycle intermediates and amino acids for binding at pH 5.4. For compounds that appeared to bind, the  $K_d$  was determined in the same manner as for citrate (Table 3.4). None of 13 metabolites screened bound AaPurE more strongly than citrate, although isocitrate, DL-2-fluorocitrate, *cis*-aconitate, *trans*-aconitate and malic acid showed measurable  $K_d$ s. For DL-2-fluorocitrate, titrations were performed at 3 pHs. None of these binding affinities is strong enough to suspect that AaPurE is regulated in vivo by the compounds tested.

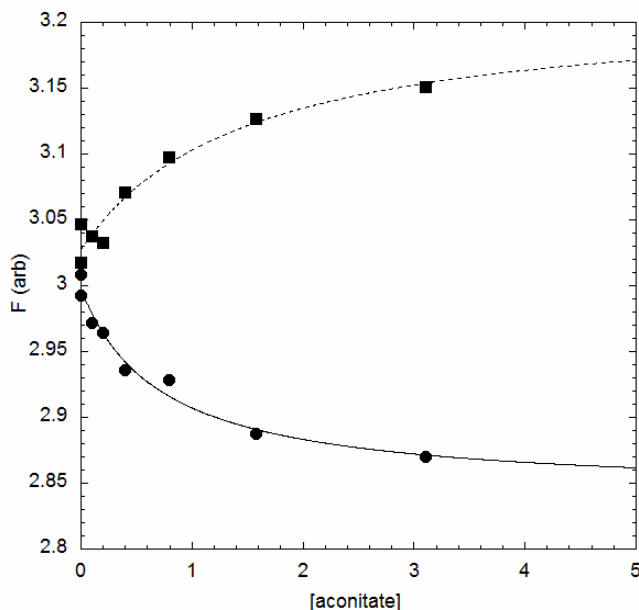
**Table 3.4:** Carboxylic acids screened as possible AaPurE ligands

Compound	Counterion <sup>c</sup>	Final pH	K <sub>d</sub> (mM) <sup>d</sup>	ΔF
Citrate <sup>a</sup>	Sodium	5.5	0.19 ± 0.02	Decrease
DL-isocitrate <sup>a</sup>	Sodium	5.6	0.8 ± 0.2	Decrease
OAA <sup>a</sup>	-	5.41	>3	Decrease
fumaric acid <sup>a</sup>	-	4.9	>5	Decrease
Succinic acid <sup>a</sup>	-	4.98	>10	None
α-ketoglutarate <sup>a</sup>	sodium	5.16	>5	None
L-glutamate <sup>a</sup>	sodium	5.43	>30	None
L-lactate <sup>a</sup>	-	5.43	>20	None
<i>cis</i> -aconitate <sup>a</sup>	-	5.45	~1	Increase
<i>trans</i> -aconitate <sup>a</sup>	-	5.44	~1	Decrease
DL-apartate <sup>a</sup>	-	5.5	>1.5	Decrease
DL-2-fluorocitrate <sup>b</sup>	potassium	4.16	0.53 ± 0.04	Decrease
		4.78	0.70 ± 0.06	Decrease
		5.39	1.0 ± 0.1	Decrease
DL-malate <sup>b</sup>	-	5.23	2.8 ± 1.0	Decrease
L-malate <sup>b</sup>	-	5.23	3.8 ± 0.6	Decrease

<sup>a</sup>Buffer for titration was 50 mM ammonium acetate, 100 mM potassium chloride,  
<sup>b</sup>Buffer was 50 mM potassium acetate, 100 mM potassium chloride, <sup>c</sup>The commercially available or prepared salt (12) was used to make titrant solution, a (-) indicates the free was acid used, <sup>d</sup>From fluorescence emission titration

Among the compounds surveyed, an interesting difference was seen in the between *cis*-aconitate and *trans*-aconitate (Figure 3.6). The binding of *cis*-aconitate caused an increase in intrinsic fluorescence of AaPurE, while the binding of *trans*-aconitate caused a decrease. The data at higher concentrations of *cis*- and *trans*-aconitate both show a sharp increase in the intrinsic fluorescence of AaPurE. This made the data difficult to fit. To obtain approximate K<sub>d</sub> values, only data collected below 5 mM were fit to equation 1. This yielded values of 0.7 ± 0.2 mM and 1.5 ± 0.8 mM for *trans*- and *cis*- aconitate respectively. Because only a portion of the data was used, the K<sub>d</sub> for *cis*- and *trans*- aconitate has been shown as ~1 mM in Table 3.4. In both titrations, data were corrected for the large inner filter effect of the respective titrants. Therefore,

it is unknown if the increase in intrinsic fluorescence seen at higher concentrations of titrant is real or an artifact of the correction.



**Figure 3.6:** Fluorescence titration of AaPurE with *cis*- and *trans*- aconitate. Fluorescence emission intensity at 337 nm plotted as a function of [aconitate] for *trans* aconitate (filled circles) and for *cis* aconitate (filled squares). *Cis*-aconitate causes an increase in the intrinsic fluorescence of AaPurE while *trans* aconitate causes an initial decrease in the intrinsic fluorescence of AaPurE. At higher concentrations, both *cis* and *trans* aconitate cause an increase in the intrinsic fluorescence of AaPurE. Only data collected under 5 mM (filled tokens) for both *cis* and *trans* aconitate were fit to Equation 1 to yield an estimate of  $K_d$ . The solid line represents a fit of the data for *trans* aconitate with  $K_d = 0.7 \pm 0.2$  mM;  $F_0 = 2.998 \pm 0.006$ ; and  $\Delta F = -0.16 \pm 0.02$ ; dashed line represents fit of the data for *cis* aconitate with  $K_d = 1.5 \pm 0.8$  mM;  $F_0 = 3.027 \pm 0.008$ ; and  $\Delta F = 0.19 \pm 0.04$ . Data collected at higher concentrations of aconitate are not included in the fits.

*Kinetic characterization of wt AaPurE and mutants at pH 8.0.* The kinetic constants  $k_{cat}$  and  $K_m$  were determined for wild type and all of the active mutants, except H59S, at pH 8.0 (Table 3.5). Although the H59S mutant had a barely observable activity (see following section) that scaled with the amount of enzyme added, this activity was too small for reliable determination of kinetic constants.

**Table 3.5: Properties of wt and mutant AaPurEs**

	Relative $k_{cat}/K_m^a$ (units/mg) <sup>b</sup>	$k_{cat}$ (s <sup>-1</sup> ) <sup>a</sup>	$K_m$ CAIR (μM) <sup>a</sup>	$K_d$ CAIR (μM) <sup>a</sup>	$K_d$ citrate (mM)
wild type	1 <sup>c</sup> (39)	12.2 ± 0.6	8.0 ± 1.2		0.40 ± 0.04
H59N	0 (<0.0005)	nd	nd	0.78 ± 0.07	1.2 ± 0.1
H59D	0.0016 (0.33)	0.49 ± 0.03	200 ± 30		7.1 ± 1.0
H59F	0 (<0.0005)	nd	nd	0.46 ± 0.09	0.8 ± 0.3
H59Q	0.000014 (0.00082)	0.0006 ± 0.0001	30 ± 20		2.8 ± 0.7
H59A	0 (<0.0001)	nd	nd	0.08 ± 0.02	>10
H59S	nd <sup>d</sup> (0.0001) <sup>n</sup>	nd	nd		>10
H89N	0.029 (0.11)	0.063 ± 0.002	1.4 ± 0.2		2.6 ± 0.4
H89D	0.044 (1.7)	0.40 ± 0.02	6 ± 1		2.7 ± 0.4
H89F	0.14 (5.8)	1.97 ± 0.09	9 ± 1		2.1 ± 0.2
H89V	0.17 (5.7)	2.10 ± 0.09	8 ± 1		0.49 ± 0.04
H89G	0.91 (1.6)	0.97 ± 0.06	0.7 ± 0.5		2.5 ± 0.3

<sup>a</sup> Determined at pH 8.0 and 30 °C in 50 mM Tris pH 8.0. <sup>b</sup> Determined at 0.1 mM CAIR. <sup>c</sup> wt  $k_{cat}/K_m = 1.5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ . nd= not determined.

Only three of the His59 mutants had detectable activity, indicating an important role for this residue. The lack of a detectable activity for several mutants indicates that contamination by host *EcPurE* does not occur. Of the mutants made at this position, only the H59D, H59Q, and H59S mutants had any activity. Of these, only the H59D and H59Q mutants had characterizable kinetics. All of the His89 mutants retained some catalytic activity. Surprisingly several of the His89 mutants approach the  $k_{cat}/K_m$  of the wt enzyme, most notably the H89G mutant. Mild substrate inhibition was noted for the *AaPurE*-H89F mutant. Because this mutation results in a larger residue at this position, the protein might be “overpacked” in the region between adjacent active sites. This could cause negative cooperativity in substrate binding, either when both adjacent active sites are occupied or during the conversion of occupied to vacant sites.

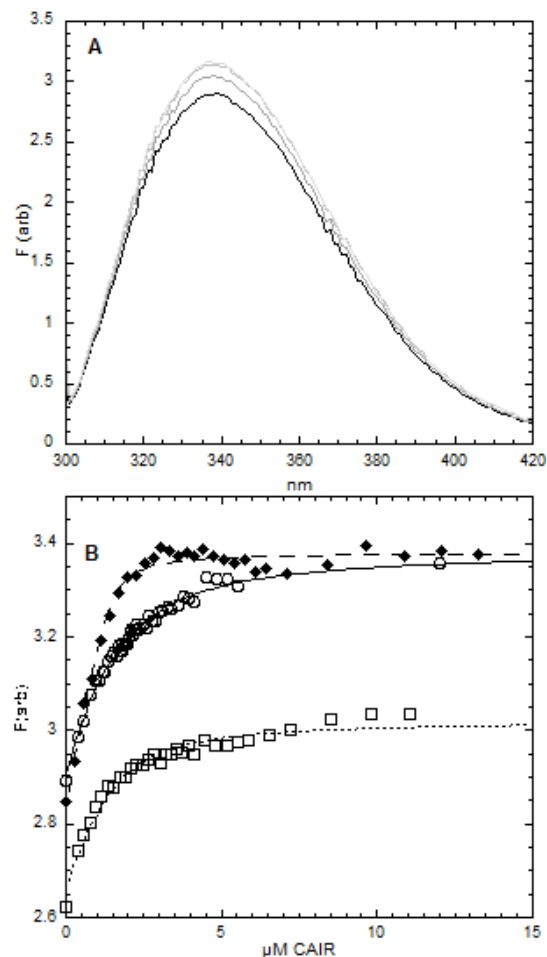
*Minimally active mutant and inactive mutants.* Activity was not detected during the purification of several mutants (H59N, H59F, H59Q, H59A and H59S). To determine

if these mutants lacked detectable activity, the standard activity assay was performed using a large amount of the purified protein (up to 0.9 mg/mL final concentration in reaction). These activity assays indicated that the mutants H59Q and H59S had a very low specific activity that scaled with the amount of enzyme added (Table 3.5). For the other mutants, no activity was detected at pH 8. The maximal residual activities for these mutants were determined from detection thresholds; maximal residual activities ranged from <0.0003 to <0.0008 units/mg. The term “inactive mutant” will be applied to these mutants and is defined as not having activity that can be detected. Given that these values are greater than the activity found for the barely active H59S mutant, they can be taken as an upper limit of residual activity. For the H59N mutant, this minimal activity assay was also repeated in the three buffer system at pH 5, 6, 7 and 8; again no measurable activity was detected (Table 3.6). At lower pH, the detection threshold is larger owing to the nonenzymatic decarboxylation of CAIR.

<b>Table 3.6:</b> Activity detection threshold for residual activity of <i>Aα</i> PurE-H59N by pH	
pH	Possible residual activity (units/mg)
8.0	<0.0002
7.0	<0.0002
6.0	<0.0003
5.0	<0.0004

<sup>a</sup> Determined at 100 μM CAIR at 30°C in triple buffer system as described in methods and materials.

*CAIR titrations of inactive mutants.* For inactive mutants H59N, H59A and H59F, fluorimetric titrations were performed using CAIR, at pH 8.0 in Tris-HCl buffer. While the binding of citrate to *AaPurE* caused a decrease in intrinsic fluorescence, the binding of CAIR caused an increase (Figure 3.7). This is an interesting distinction from *EcPurE* mutants, which show a decrease in intrinsic fluorescence upon the binding of CAIR (13). The data collected for each titration were fit to equation 5 with half of the sites binding. Fitting to a half-sites model always gave a superior fit than full sites and will be discussed later.



**Figure 3.7:** Fluorescence titration of inactive AaPurE mutants with CAIR. (A) Fluorescence emission spectra for AaPurE-H59N mutant at 0 (black line), 0.78  $\mu\text{M}$  CAIR (dark grey line), 2.2  $\mu\text{M}$  CAIR (grey line) and 19  $\mu\text{M}$  CAIR (light grey line), for 2.7  $\mu\text{M}$  protein subunit, recorded at 25  $^{\circ}\text{C}$  with excitation at 295 nm, final pH 7.99. (B) Fluorescence emission intensity at 338 nm as function of CAIR concentration, for AaPurE-H59N (open circles), AaPurE-H59F (open squares), and AaPurE-H59A (filled diamonds). Titrations were performed at 2.7, 2.9 and 2.9  $\mu\text{M}$  protein [subunit] respectively. Data were fit to equation 5, which yielded values of  $K_{d,\text{CAIR}}$  summarized in Table 3.5.

*pH rate profiles of AaPurE-wt and AaPurE-H59D.* Due to the nature of the reaction catalyzed, and the ability for the product  $N^5\text{-CAIR}$  to undergo rapid degradation to AIR and  $\text{CO}_2$ , special care was taken in designing how the pH-rate profile would be compiled. Simple buffer systems such as Tris-HCl or acetate buffer have a narrow

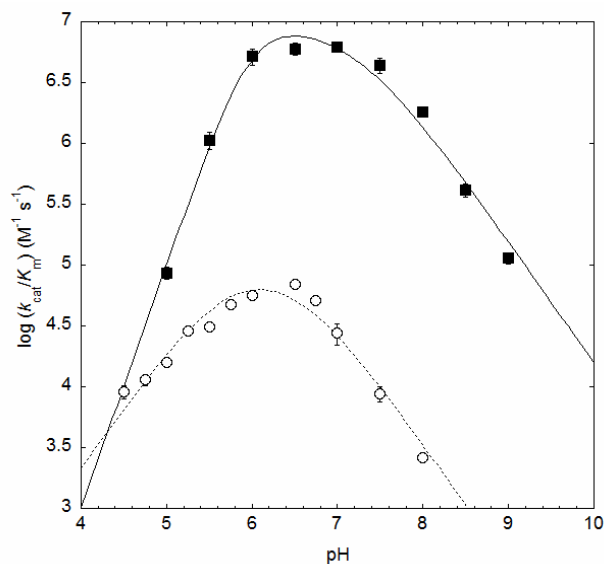


buffering capacity, and to set the pH of a buffer, acid or base is added. This changes the total ionic strength of the solution, requiring salt to be added to normalize the ionic strength of each buffer at each pH. The use of a triple buffer system can avoid this problem with the choice of the proper components by fixing the ionic strength of the buffer at all pHs within the buffering capacity of the system (14).

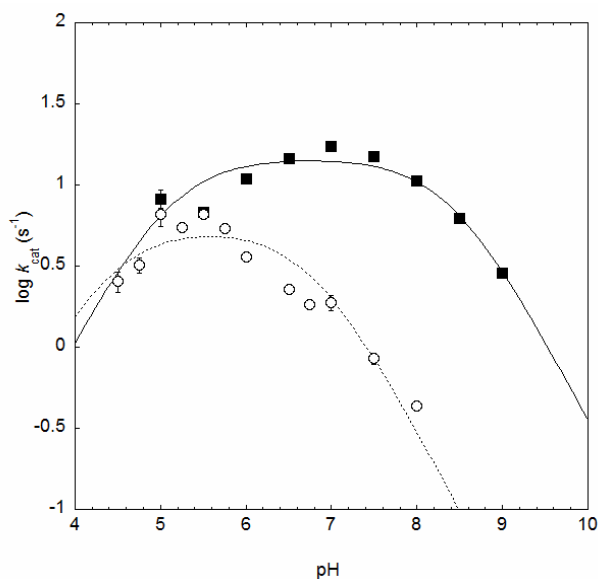
Michaelis-Menten plots were determined for AaPurE-wt at 0.5 unit intervals from pH 5.0 to 9.0, and for AaPurE-H59D from pH 4.5 to 8.0 using the  $\Delta\epsilon_{260}$  for each pH (Chapter 2). The data collected at each pH were fit to equation 3 to determine the values of  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_m$ . A semi-logarithmic plot of these values versus pH gives bell-shaped profiles consistent with two or more  $pK_a$ s (Figures 3.8 and 3.9).

*$k_{\text{cat}}/K_m$  pH rate profile.* The  $k_{\text{cat}}/K_m$  pH rate profiles of AaPurE-wt and AaPurE-H59D yield bell shaped profiles that are slightly different in shape (Figure 3.8). The AaPurE-H59D profile fit well to equation 4 (AaPurE-H59D  $pK_1 = 5.7 \pm 0.2$ ,  $pK_2 = 6.5 \pm 0.2$ ). However, the steeper acidic arm slope of the AaPurE-wt profile required fitting with equation 5, which assumes that two ionizations occur on the acidic arm ( $pK_1^*$ ) of the profile (AaPurE-wt  $pK_1^* = 6.0 \pm 0.1$ ,  $pK_2 = 7.2 \pm 0.1$ ).

*$k_{\text{cat}}$  pH rate profile.* The  $\log k_{\text{cat}}$  versus pH profiles for AaPurE-wt and AaPurE-H59D yielded bell shaped curves that were fit to equation 4 (Figure 3.9, AaPurE-wt  $pK_1 = 5.1 \pm 0.2$ ,  $pK_2 = 8.4 \pm 0.1$ ; AaPurE-H59D  $pK_1 = 4.4 \pm 0.4$ ,  $pK_2 = 6.8 \pm 0.2$ ).



**Figure 3.8:**  $k_{cat}/K_m$  pH-rate profiles.  $\log(k_{cat}/K_m)$  as a function of pH for AaPurE-wt (filled squares) and AaPurE-H59D (open circles). Values of  $k_{cat}$  and errors shown were determined from initial velocity data obtained varying amounts of CAIR at each pH indicated, fit to the Michaelis Menten equation. For most points, the error was smaller than tokens used. Lines shown are the fit of H59D data to eq 3, and the fit of wt data to eq 4, which accounts for a second protonation (second  $pK_1$ ) occurring on the acidic arm of the profile.



**Figure 3.9:**  $k_{cat}$  pH-rate profiles.  $\log(k_{cat})$  as a function of pH for AaPurE-wt (filled squares) and AaPurE-H59D (open circles). Values of  $k_{cat}$  and errors shown were determined from initial velocity data obtained varying amounts of CAIR at each pH indicated, fit to the Michaelis Menten equation. For most points, the error was smaller than tokens used. Lines shown are fits of data to eq 3.

### 3.4 Discussion

The foci of the study of *AaPurE* were two-fold: (1) to investigate the role of Class I PurE conserved residues His59 and His89, (2) to characterize *AaPurE* over a range of pHs to investigate how it might be adapted to function at low pH. There are three non-exclusive ways one can imagine for an enzyme to be adapted to function at low pH (4-6): (A) an enzyme may be catalytically optimized to function at low pH; (B) a protein may have improved stability at low pH; (C) a protein may work with the aid of chaperone proteins at low pH. Only B is supported by our data, through a comparison of *AaPurE* and *EcPurE*, which was not expected to have a particular stability at low pH. In this work, it has been shown that *AaPurE* is intrinsically stable to acid-mediated denaturation over the range of pH experienced in the cytoplasm of *A. acetii*, which suggests that an acid chaperone is not required for the functioning of *AaPurE*.

#### (1) Investigation into the roles of conserved residues His59 and His89

*Purification of AaPurE.* The purification scheme outlined in Scheme 3.2 provided a rapid method of purifying *AaPurE* and mutants to >90% purity. The ammonium sulfate fractionation procedure removes most *EcPurE*, which remains soluble at a much higher level of ammonium sulfate saturation (>60% AS) (15). *AaPurE* is also distinguished from *EcPurE* (and *Cryptococcus neoformans Ade2*) in its inability to bind hydroxyapatite at pH 8.0 and low phosphate conditions (15, 16). These two differences allowed *AaPurE* and mutants to be separated from any host *EcPurE*, as is evidenced by

the complete lack of a detectable activity for the H59A, H59F and H59N mutants.

Although the purifications of most of these inactive mutants have only been performed once, the lack of activity in these mutants is supported by results from a mutant *EcPurE* with equivalent to *AaPurE*-H59N (13).

*Carboxylic acid titrations.* The first reported crystal structure of *AaPurE* indicated that citrate was bound in the active site. This suggested that citrate might be used as an active-site probe. The ability of citrate to serve as an active site probe was also used to study the active site of *AaPurE* using REDOR with <sup>13</sup>C labeled citrate (17). These studies indicated that His59, under the conditions used, was found in a mixture of two protonation states (17). Binding of citrate caused a diminution of fluorescence in *AaPurE*, which was used to determine the  $K_d$  for citrate. To investigate whether citrate or any common carboxylic acid metabolites had the potential to regulate *AaPurE* in vivo, 13 different compounds were titrated with *AaPurE*. Citrate was found to bind most strongly but with a modest affinity ( $K_d = 0.19 \pm 0.02$  mM). None of the compounds appeared to bind strongly enough to suggest they regulate *AaPurE* in vivo. It is interesting to note that some carboxylic acids had different effects on the fluorescence of *AaPurE*: binding of *cis*-aconitate led to an increase in fluorescence, while the binding of *trans*-aconitate led to a decrease. Of all of the compounds that bound, and those that did not, no clear structure-affinity relationship could be discerned.

The affinity of *AaPurE* for citrate allowed citrate to be used as an active site probe in most cases. Only two of the mutants failed to show binding of citrate by the fluorescence titrations. Given that the two mutants that do not appear to bind citrate

were shown to either bind CAIR (H59A) or to have a detectable activity (H59S), the active sites of each mutant does appear to be intact. The failure to bind citrate detectably could be due to a couple of reasons: (1) these mutants may bind citrate, but the binding causes no measurable change in their intrinsic fluorescence. (2) different portions of *AaPurE* are used to bind citrate and CAIR; namely His59, which is not required for the binding of CAIR but does have an interaction with citrate.

*Citrate pH-binding profile.* To further investigate the binding of citrate to *AaPurE*, titrations were performed at a number of pHs (0.3 pH unit intervals) from pH 3.5 to 6.4, and used to construct a pH binding profile, which was fit to equation 2 (Figure 3.5). This analysis was consistent with a single protonation on each arm of the profile, and yielded  $pK_a$  values ( $pK_1 = 4.7 \pm 0.2$  and  $pK_2 = 5.1 \pm 0.2$ ) that were very close to each other and to the known  $pK_2$  of citrate (4.76). The value of  $pK_1$  is consistent with *AaPurE* binding to the mono-anionic or di-anionic of citrate. The value of  $pK_2$  may correspond to Asp 33.

*Binding CAIR.* Thus far there is no direct evidence to indicate that more than half of the active sites of the *AaPurE* octamer are occupied by CAIR at any given time. All of the crystal structures of *AaPurE* with a bound nucleotide have only half of the active sites occupied. However, recent crystal structures of *EcPurE* with nucleotide bound, showed all eight sites of the octamer occupied (13). Fitting the data collected for *AaPurE* mutants with half-sites binding yielded superior fits because the majority of fluorescence change occurred at a concentration of CAIR that was lower than the concentration of enzyme. Whether this difference between these enzymes is real is unknown. If the determined value for the concentration of *AaPurE* is actually higher,

this could result in the requirement of fitting the data to half sites binding. When fit under the half-sites requirement, all three mutants bind CAIR with sub-micromolar affinities (Table 3.5). If full-sites binding does occur and the determined concentrations of *AaPurE* are too high, this would likely mean that the affinity for CAIR is even stronger. The fact that all of the mutants bind CAIR or have an activity indicates that the active site for all forms of *AaPurE* studied remains largely intact.

*Role of conserved Histidine 59.* Mutagenesis studies reported herein suggest an acid/base role for this residue, as only the mutant with an ionizable residue at this position (H59D) is active. All other mutants at this position are inactive or nearly so. The pH-rate profiles that will be discussed in the following section suggest that residue 59 serves as an active-site acid in the reverse direction. The consistently sub-micromolar  $K_{d,CAIR}$  for the inactive His59 mutants along with the slight steric clash seen in the wt-*AaPurE*-AIR structure (2fwj) indicate that His59 does not contribute to the high affinity binding. In the *AaPurE*-wt-AIR structure, the aminoimidazole ring of AIR is wedged between the edge of His59, a hydrophobic pocket, and two backbone amides. In the barely active mutants (H59S, H59Q), water may serve in the active site as an acid, as either of these residues serving as an acid seems unlikely.

*Role of conserved residue His89.* Owing to the difference in reactions catalyzed by Class I and Class II PurEs, it was hypothesized that His89 might serve as part of a “proton wire” that could move protons into and out of an occupied active site (2), similar to the role of His64 in carbonic anhydrase (18) or the extended proton wire in the E1 component of the pyruvate dehydrogenase complex (3). Mutagenesis studies

reported herein show that His89 is not catalytically essential. The replacement of this residue with many other amino acids, some of which would preclude the ability to form a proton shuttle such as the H89F mutant, still results in active enzymes.

All of the His 89 mutants have a decreased activity in comparison to wt *AαPurE*; however, the most efficient substitution, with nearly wt  $k_{cat}/K_m$ , is *AαPurE*-H89G. In this mutant the  $K_m$  for CAIR is ~10 fold lower than that of the wild type (wt) enzyme and approaches the values of  $K_{d,CAIR}$  determined for the inactive His59 mutants. This lower affinity may indicate that *AαPurE*-wt exerts some ground-state destabilization on the bound nucleotide. The replacement of His89 with the much smaller glycine would alleviate this destabilization. Although substitution of His89 with other amino acids led to some changes in the value of  $K_{d,citrate}$ , all mutants bind citrate.

If His89 serves as a proton shuttle, it is clearly not needed for the functioning of *AαPurE*. It is interesting to compare the two different reactions catalyzed by Class I and Class II PurEs (Scheme 3.3). While in the forward direction His59 is proposed to serve as an active site acid in Class I PurEs, the equivalent histidine is proposed to serve as an active site base in Class II PurEs. The conserved role of His89 in Class I PurEs may be to stabilize protonated His59, as is indicated by the relatively high  $pK_a$  of His59 ( $7.2 \pm 0.1$ ). Owing to the different reactions, Class I PurEs require a protonated histidine, while the opposite is true for class II PurEs. The equivalent residue for His89 in animals is a conserved glycine, which would not be expected to stabilize a protonated His59.

## (2-A) Investigation into the catalytic properties of AaPurE and AaPurEH59D

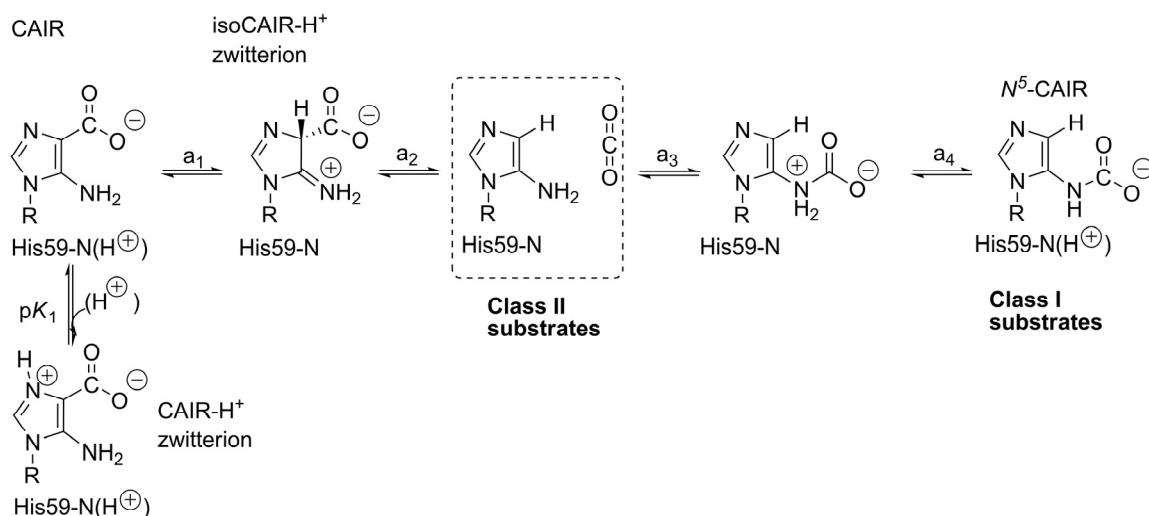
*k<sub>cat</sub>/K<sub>m</sub> profile* - The value of  $K_m$  is the concentration of substrate at which the enzyme functions at half of its maximal activity ( $k_{cat}$ ). The value of  $K_m$  is therefore dependent on the affinity of the free enzyme for the free substrate and is dependent on the enzyme-substrate complex. Thus the pH profile of  $K_m$  for enzymes can be difficult to interpret. However, the value of  $k_{cat}/K_m$  is only affected by the affinity of free enzyme for free substrate. Changes in this value as a function of pH are indicative of ionizations of moieties involved in enzyme-substrate binding.

The value of  $pK_1$  for both the AaPurE-wt and AaPurE-H59D pH profiles is similar ( $6.0 \pm 0.1$  and  $5.7 \pm 0.2$  respectively), suggesting some common feature is responsible. If the substrate does not ionize, these similar values might be associated with an enzyme side chain that must be unprotonated for substrate binding, such as the side chain of Asp33, which is involved in binding of the ribose moiety (Figure 3.1). If the substrate does ionize, a plausible candidate is the N3 nitrogen of CAIR which has a  $pK_a$  of 6.11 at 30 °C (19). The N3 nitrogen of CAIR interacts with Ser57 (Figure 3.1). Serine can serve as a hydrogen bond donor or acceptor; therefore,  $pK_1$  may be due to CAIR N3 protonation if Ser57 functions as a hydrogen bond donor. Alternatively,  $pK_1$  may be due to the protonation of the substrate ribose 5-phosphate (R5P) group ( $pK_{a2}$  for R5P = 6.3 (20)).

The value of  $pK_2$  in the pH profiles of AaPurE-wt and AaPurE-H59D shows a large shift,  $7.2 \pm 0.1$  and  $6.5 \pm 0.2$  respectively, that is dependent on the identity of residue 59. This assigns  $pK_2$  to His59. The shift that occurs from His59  $\rightarrow$  Asp59 is due to the



lower  $pK_a$  of the Asp side chain. Alternatively  $pK_2$  may correspond to the protonation state of His89, if its interaction with residue 59 plays a role in *AaPurE* being able to bind CAIR productively.



**Scheme 3.3:** Possible mechanism of *AaPurE*. Individual steps shown may be concerted, such as steps a<sub>1</sub> and a<sub>2</sub> or a<sub>3</sub> and a<sub>4</sub>

*k<sub>cat</sub> profile.*  $k_{cat}$  is the value of the maximal rate of catalysis for an enzyme at saturating substrate concentration. Thus, the effects of substrate and enzyme binding steps are eliminated, and all of the enzyme is taken to be in the form of the enzyme-substrate complex. Changes in this value, as a function of pH, are indicative of protonations of key catalytic moieties in the enzyme-substrate complex (21).

The large difference in the value of  $pK_2$  for the *AaPurE*-wt and *AaPurE*-H59D mutant,  $8.4 \pm 0.1$  and  $6.8 \pm 0.2$  respectively, indicates that a protonated residue 59 has a key role in catalysis. These values are higher than the  $pK_2$  determined for the  $k_{cat}/K_m$  profiles, which indicates that the  $pK_a$  of residue 59 is higher in the enzyme-substrate complex.

Compared to the values of  $pK_2$ , the values of  $pK_1$  for both the *AaPurE*-wt and *AaPurE*-H59D pH profiles are similar in value ( $5.1 \pm 0.2$  and  $4.4 \pm 0.4$  respectively), indicating that residue 59 may not affect the acidic arm of the profile. The hydrogen bond that forms between Ser 57 and the N3 of CAIR would likely make CAIR more difficult to protonate, and would cause the  $pK_a$  of the N3 of CAIR to be lower than that of the free nucleotide ( $pK_a=6.11$ ) as is observed.

*Assignments of  $pK_2$  in the  $k_{cat}$  versus pH profile.* It has been previously postulated that Class I PurEs function in the reverse direction by catalyzing the protonation of CAIR C4 (22). While protonation of this carbon might seem like a candidate for the  $pK_2$  (basic arm) of the  $k_{cat}$  pH-rate profile, the shift noted in the profiles of His  $\rightarrow$  Asp indicates the basic arm in the pH-rate profile is due to residue 59 in the enzyme-substrate complex. A structure determined from a crystal of the inactive *AaPurE*-H59N mutant soaked in CAIR (structure 2fwp), exhibits what is likely a mixture of AIR + CO<sub>2</sub> and the tetrahedral intermediate isoCAIR in the active site (5, 13). This species is likely formed due to the introduction of a proton into the active site of the inactive mutant at the pH which the crystal structure was determined (pH 5.4). This indicates that a residue capable of acid/base chemistry is required to catalyze the conversion of CAIR to  $N^5$ -CAIR.

A possibility for  $pK_2$  is that a protonated His59 is required in the first step of the reaction (step  $a_1$ ), in which the C4 position of CAIR is protonated to form the isoCAIR zwitterion. The isoCAIR zwitterion could decarboxylate to form AIR and CO<sub>2</sub> in the active site (step  $a_2$ , Scheme 3.3). The enzyme bound AIR and CO<sub>2</sub> would seem equally

capable of forming  $N^5$ -CAIR or CAIR, depending on whether the  $\text{CO}_2$  molecule is attacked by the AIR exocyclic amino group or C4, respectively. The product formed from the enzyme-bound  $\text{CO}_2$  intermediate may be determined by movement of the aminoimidazole moiety of AIR (13).

The active H59D mutant shows that Asp is an acceptable substitute to catalyze the reaction. One possible problem with the hypothesis that residue 59 protonates CAIR is the difference in charge between protonated histidine and a protonated aspartic acid. However, crystal structures of *AaPurE*-H59D indicate that Asp59 forms a hydrogen bond with the neighboring His89. This interaction has no counterpart in the wild type structure. The protonated His89 may facilitate protonation or binding of the substrate. With such an interaction, requiring both His89 and Asp59 to be protonated, the pH-rate profile would be determined by the second protonation, that of Asp59. This is consistent with the lower  $\text{p}K_2$  seen with the Asp59 mutant. In contrast, protonation of His89 in the wild type enzyme may lead to an inactive form of the enzyme, which is what we hypothesize is part of the reason a two proton  $\text{p}K_1$  seen in the wt  $k_{\text{cat}}/K_m$  pH rate profile.

*Comparison of kinetic constants for AaPurE-wt and AaPurE-H59D.* The seemingly non-conservative Asp mutant is a surprisingly efficient enzyme, with a maximal  $k_{\text{cat}}$  equal to ~40% of wt maximal  $k_{\text{cat}}$ . Since  $\text{p}K_1$  and  $\text{p}K_2$  are closer together for *AaPurE*-H59D, the actual pH-independent maximal mutant  $k_{\text{cat}}$  is likely closer to 55% of wt. At all pH values tested, the  $K_m$  for *AaPurE*-H59D is much higher than that for *AaPurE*-wt, so the  $k_{\text{cat}}/K_m$  for *AaPurE*-H59D is significantly lower than wt. This is consistent with the proposal that the *AaPurE*-H59D mutant does not bind CAIR as effectively as *AaPurE*-wt.

This would be expected if the mutant enzyme electrostatically repels the anionic substrate. The anionic active site probe citrate shows a similarly lower affinity for the mutant enzyme, with an 18 fold increase in the  $K_d$  from wt  $\rightarrow$ H59D.

*Comparison of pH rate profiles of EcPurE and AaPurE.* To date, the only other PurE for which a pH-rate profile has been constructed is wt *EcPurE* (13). The data reported by those authors yielded a bell shaped profile for both the  $k_{cat}$  and  $k_{cat}/K_m$ . They determined values of  $pK_1 = 5.9 \pm 0.4$  and  $pK_2 = 8.6 \pm 0.4$  from the *EcPurE*  $k_{cat}$  pH-profile. The value determined for  $pK_2$  is very similar to the value determined herein for *AaPurE* ( $pK_2 = 8.4 \pm 0.1$ ), while the value of  $pK_1$  is slightly higher than the value for *AaPurE* ( $pK_1 = 5.1 \pm 0.2$ ). They determined values of  $pK_1 = 6.7 \pm 1.6$  and  $pK_2 = 7.5 \pm 1.5$  for the *EcPurE*  $k_{cat}/K_m$  pH profile. The value determined for  $pK_2$  is very similar to the value determined herein for *AaPurE* ( $pK_2 = 7.2 \pm 0.1$ ), while the value of  $pK_1$  is slightly higher than the value for *AaPurE* ( $pK_1 = 6.0 \pm 0.1$ ). These authors claimed that equation 4 was sufficient to fit their data, and that there was no need to fit the data to an equation with three ionizations, as we found to be necessary for *AaPurE*-wt (two ionization on the acidic arm). Regardless of the interpretation, much caution is warranted for the interpretation of data on the acidic arm of the pH-rate profiles, because of the high rate of nonenzymatic CAIR decarboxylation, and the lack of extinction coefficients for  $N^5$ -CAIR below  $pH \leq 6.0$ .

While the absolute magnitude of  $k_{cat}$  for *AaPurE* does not approach that of *EcPurE* (*AaPurE* =  $17 \text{ s}^{-1}$  at pH 7.0, maximum rate from *EcPurE* plot  $\sim 26 \text{ s}^{-1}$  at pH 7.3), *AaPurE* has been shown to have a consistently lower  $K_m$ , indicating that *AaPurE* may

bind CAIR more tightly. A higher affinity is also seen in the  $K_{d,CAIR}$  determined for equivalent mutants, *AaPurE*-H59N and *EcPurE*-H45N ( $0.78 \pm 0.06$  and  $21 \pm 2 \mu\text{M}$  respectively). This higher affinity in *AaPurE* results in an enzyme which reaches a higher  $k_{cat}/K_m$ ,  $6.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$  (pH 7.0) than *EcPurE*,  $\sim 3.6 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$  (pH 6.6).

## **(2B) – Structural analysis**

It would seem that proteins that experience acidic conditions for prolonged amounts of time would necessarily have adapted to survive under those conditions. It would generally be expected that proteins from an acidophile would have increased stability under acidic conditions. This is evidently the case for *AaPurE*. *AaPurE* appears to be well suited for survival at acidic pHs as is evidenced by the thermal melt profile constructed (Figure 3.5). Even at pH 3.5, *AaPurE* retains its structure while *EcPurE* unfolds at room temperature. Similarly, increased stability under acidic conditions has been seen in the cases of citrate synthase (4) and alanine racemase (6) from *A. acetii*, in comparison to equivalent enzymes from the neutralophiles, pig and *B. stearothermophilus*.

*Stability of AaPurE compared to EcPurE.* In the report of the initial crystal structure for *AaPurE* (2), comparisons were made between it and its homologs from *E. coli*, a mesophilic neutralophile, and *T. maritima*, a thermophilic neutralophile. While one might expect that the crystal structure of *AaPurE* would share the most similarity in structure to *EcPurE*, as both are mesophilic organisms and share 52% sequence identity, the surface charge decoration and hydrogen bonding patterns were more similar to

those of *TmPurE*. Interestingly, *AaPurE* was found not only to be more resistant to acidic denaturation than *EcPurE* (*EcPurE* unfolded immediately at pH 3.5, while *AaPurE* did not), but it was also far more resistant to thermal denaturation than *EcPurE* at all pHs tested. The high  $T_m$ s observed for *AaPurE* may be consistent with its structural similarity to *TmPurE*. Similar experiments performed on *A. acetii* citrate synthase and alanine racemase have shown that they are both acid-resistant and relatively thermostable. This indicates that there may be general similarities in the development of proteins that are stable under unusual conditions, as will be discussed in greater detail below (4, 6).

*Crystal structures of AapurE and mutants.* To date, a total of 11 crystal structures of *AaPurE*-wt and mutants have been determined over a range of pH from 5.4-8.5 (Table 3.7). Notably, there are only a few subtle differences among the structures; in most cases these differences can be explained as a result of ligand binding or as a result of a mutation. Some mutant crystal structures show a slight rotation of residue Tyr154 that results in the residue forming a hydrogen bond of shorter length with residue Asn78 when compared to *AaPurE*-wt structures. This rotation appears not to occur as a consequence of a pH change. Other proteins, for which crystal structures have been determined over a range of pH, show changes ranging from the reorientation of a few side chains (23, 24), to large scale alterations of the entire protein structure (25).

The cause of changes in the structure of any protein at different pHs will likely be due to a change in the ionization state of one or more side chains. The lack of change

noted in crystal structures of *AaPurE* may indicate that no changes in protonation state have occurred. An examination of what changes might be expected over the range of pH for which crystal structures have been determined highlights a possible way that proteins from acidophiles may have evolved to survive an acidic environment. Proteins from acidophiles may discourage changes in protonation state that could destabilize the protein over the range of pH in which they grow.

**Table 3.7:** Summary of AaPurE crystal structures

PDB code	1u11	2fwj	2fw1	2fwi	2fwp	2fw6	2fw7	2fw9	2fwb	2fw8	2fwa
AaPurE	wt	wt	wt	H59D	H59N	H59N	H59N	H59F	H89F	H89G	H89N
pH	pH 5.4	pH 7.0	pH 8.5	pH 7.0	pH 5.4	pH 5.4	pH 8.0	pH 8.0	pH 8.0	pH 8.0	pH 7.0
Ligand <sup>a</sup>	citrate	AIR	-	AIR	AIR + CO <sub>2</sub>	citrate	-	-	-	-	-
Tyr154 <sup>b</sup> rotation	-	N	N	Y	Y	Y	N	Y	N	Y	N

<sup>a</sup>symbol (') denotes different subunit. <sup>b</sup>Made in reference to orientation found in crystal structure 1u11. In crystal structure 1u11, the  $\chi_2$  dihedral in reference to C $\delta$ 1 is 141.1°; In crystal structure 2fw6 this dihedral is 102.0°



From pH 8.5 to pH 5.4, one might expect the following ionizations to occur: the more acidic lysines in the protein become protonated (usually caused by being buried in hydrophobic region), histidine becomes protonated, and the more basic aspartates and glutamates become protonated. The above is only a broad guideline to the ionizations expected, as the protonation state of any residue is dependent on the environment in which it is found.

*AaPurE* contains a total of four lysines (Table 3.8). In the crystal structures, two are in portions of the protein that are unstructured, so observations cannot be made. The remaining two are exposed to solvent, and would thus be expected to be protonated at all the pHs for which crystal structures have been determined. *AaPurE* contains a total of 4 histidines, 2 of which are the conserved residues 59 and 89; the  $pK_a$  of His59 appears to be 7.2 as determined from the pH rate profile and is also consistent with REDOR of the *AaPurE*-citrate complex (17). The other two histidines are on the surface of the protein. His39 is near Arg38 (3.8 Å) and may have a depressed  $pK_a$  owing to electrostatic considerations and would not be expected to become protonated over the pH range examined. His51 interacts with Asp41 and may be protonated over the pH range examined (the N<sup>ε</sup> to carboxylate oxygen distance is 2.6 Å in 1u11, pH 5.4, and 2.7 Å in 2fw1 ,pH 8.5).

The remainder of residues that might ionize over the pH range examined are the carboxylic acids aspartate and glutamate. *AaPurE* contains a total of 17 of these residues, 4 which are always disordered, and 1 (Glu180) that is the last ordered residue in a few structures, but disordered in most crystal structures. Any change in the

unstructured residues cannot be observed. Of the remaining 12 carboxylic residues that may ionize in the structure, 10 are involved in salt bridges with arginines (2), and an eleventh salt bridge may exist between His 51 and Asp 41.

The interaction between two oppositely charged residues stabilizes the ionized state of each, making the acidic residue more acidic and the basic residue more basic, in comparison to the residues when lacking this electrostatic interaction. Experimentally this stabilization results in a  $\Delta pK_a$  of 0.5 to 3.0 units for both residues (26-28). It has been noted in the crystal structures of *AaPurE* (2) that the geometry of the salt bridges appears ideal. Thus it seems reasonable to expect that the resulting shifts may be large. Since almost all of the structured aspartate and glutamate side chains are involved in salt bridges, it seems unlikely that they would ionize over the range of pHs for which crystal structures have been determined. Further, these salt bridges may prevent changes in the protonation state of *AaPurE* over the range of pH experienced in vivo.

**Table 3.8:** Comparison of *A. acetii* and *E. coli* crystal structures

	<i>AaPurE</i> (1u11)	<i>EcPurE</i> (1d7a)	<i>AaCS</i> ·OAA·CMX <sup>b</sup> (2h12)	<i>EcCS</i> ·NADH <sup>b</sup> (1nxg)
Structured His	4	6	9	15
Structured Lys	2	6	22	23
Structured Arg	9	6	21	23
Structured Asp+Glu	13 <sup>a</sup>	16	38	51
Total Salt Bridges	11 <sup>a</sup>	10	13	9
Asp+Glu not in Salt Bridge	2	6	25	44

<sup>a</sup>Glu180 is not structured in all crystal structures, but is included in the total for 1u11. <sup>2</sup>Includes possible salt bridge between Asp41 and His51. <sup>b</sup>Summarized from (4).

*AaCS*, another protein from *A. acetii* for which a crystal structure has been determined, also shows an increased number of salt bridges in comparison to its

counterpart from the mesophile *E. coli* (Table 3.8). *AaCS* also shows a decreased number of the residues that would be expected to ionize over the range of pH experienced in the cytoplasm of *A. acetii*. Thioredoxin, the only other protein from *A. acetii*, for which a crystal structure has been determined (29), does not show the marked differences that *AaPurE* and *AaCS* show to their *E. coli* counterparts. However, it has been noted *E. coli* thioredoxin is an acid-stable protein (30).

*Extension to salt bridges in thermophiles.* The  $pK_a$  values of nitrogenous bases have been shown to be temperature-dependent (31-34), becoming weaker bases at higher temperatures. While a vast body of work has been compiled on salt bridges in thermophiles, there are few examples of investigating the temperature dependence of protein side chain  $pK_a$ s (35). Over a 34 °C range, the  $pK_a$ s of six histidines in myoglobin were found to decrease by 0.35-0.72 units. The lower values ( $\sim 0.01$   $pK_a$  unit per °C) are similar to that seen in the temperature dependence of the second ionization of the free amino acid histidine in aqueous solution (33), while the larger values ( $\sim 0.02$ - $0.025$   $pK_a$  unit per °C) are similar to the change seen for imidazole (32), the side chain of lysine (33), and alkanamines (34).

At the higher temperatures in which thermophiles grow, one role that some salt bridges could serve is to hinder changes in the protonation state of the basic residues involved. While even at high temperature, lysine will predominately be protonated, the increase in population of unprotonated lysine may lead to a form of the thermophilic protein that denatures. This is analogous to the proposed role that some salt bridges

may serve in an acidophile such as *A. aceti* with salt bridges hindering changes in protonation of carboxylic residues.

It is interesting to note that there is a lower fraction of histidine residues in thermophile proteins (36-38). This preference has not been explained (36-39). One of the causes for the reduction of histidine may be that at the higher temperatures of thermophiles, the properties of histidine are quite different based on the temperature dependences of the ionization of the imidazole ring.

### **3.5 Future Directions**

It has been proposed in this chapter that the mutant *AαPurE*-H59D mutant may require a protonated His89 in order to function. The characterization of a double mutant such as *AαPurE*-H59D-H89F, should yield an answer to this proposal. If this mutant has an activity near that of the *AαPurE*-H59D mutant, this would indicate that a protonated His89 is not required. However, if very low or no activity is noted, and the active site of the double mutant remains intact, this would indicate that a protonated His89 is required.

### 3.6 References

1. Mathews, I. I., Kappock, T. J., Stubbe, J., and Ealick, S. E. (1999) Crystal structure of *Escherichia coli* PurE, an unusual mutase in the purine biosynthetic pathway, *Structure* 7, 1395-1406.
2. Settembre, E. C., Chittuluru, J. R., Mill, C. P., Kappock, T. J., and Ealick, S. E. (2004) Acidophilic adaptations in the structure of *Acetobacter aceti* N<sup>5</sup>-carboxyaminoimidazole ribonucleotide mutase (PurE), *Acta Crystallographica Section D: Biological Crystallography D60*, 1753-1760.
3. Frank, R. A. W., Titman, C. M., Pratap, J. V., Luisi, B. F., and Perham, R. N. (2004) A molecular switch and proton wire synchronize the active sites in thiamine enzymes, *Science* 306, 872-876.
4. Francois, J. A., Starks, C. M., Sivanuntakorn, S., Jiang, H., Ransome, A. E., Nam, J.-W., Constantine, C. Z., and Kappock, T. J. (2006) Structure of a NADH-insensitive hexameric citrate synthase that resists acid inactivation, *Biochemistry* 45, 13487-13499.
5. Constantine, C. Z., Starks, C. M., Mill, C. P., Ransome, A. E., Karpowicz, S. J., Francois, J. A., Goodman, R. A., and Kappock, T. J. (2006) Biochemical and structural studies of N<sup>5</sup>-carboxyaminoimidazole ribonucleotide mutase from the acidophilic bacterium *Acetobacter aceti*, *Biochemistry* 45, 8193-8208.
6. Francois, J. A., and Kappock, T. J. (2007) Alanine racemase from the acidophile *Acetobacter aceti*, *Protein Expression and Purification* 51, 39-48.
7. Srivastava, P., Mancuso, R., Rousseau, R., and Robins, R. (1974) Nucleoside Peptides. 6. Synthesis of Certain N-[5-Amino-1-(β-D-ribofuranosyl)imidazole-4-carbonyl] amino acids related to naturally occurring intermediates in the purine biosynthetic pathway, *Journal of Medicinal Chemistry* 17, 1207-1211.
8. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry* 72, 248-254.
9. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994) N<sup>5</sup>-carboxyaminoimidazole ribonucleotide: Evidence for a new intermediate and two new enzymic activities in the de novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry* 33, 2269-2278.
10. Menzel, U., and Gottschalk, G. (1985) The internal pH of *Acetobacterium wieringae* and *Acetobacter aceti* during growth and production of acetic acid, *Archives of Microbiology* 143, 47-51.
11. Bates, R., and Pinching, G. (1949) Additions and corrections-Resolution of the dissociation constants of citric acid at 0 to 50 degrees and determination of certain related thermodynamic functions, *Journal of the American Chemical Society* 71, 4165-4165.
12. Erlichman, J. S., Li, A., and Nattie, E. E. (1998) Ventilatory effects of glial dysfunction in a rat brain stem chemoreceptor region, *Journal of Applied Physiology* 85, 1599-1604.
13. Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, I. I., Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E., and Stubbe, J. (2007) N<sup>5</sup>-CAIR mutase: Role of a CO<sub>2</sub> binding site and substrate movement in catalysis, *Biochemistry* 46, 2842-2855.
14. Ellis, K. J., Morrison, J. F., and Daniel, L. P. (1982) [23] Buffers of constant ionic strength for studying pH-dependent processes, in *Methods in Enzymology*, pp 405-426, Academic Press.

15. Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J., and Smith, J. M. (1992) Purification and characterization of the *purE*, *purK*, and *purC* gene products: identification of a previously unrecognized energy requirement in the purine biosynthetic pathway, *Biochemistry* 31, 5022-5032.
16. Firestine, S. M., Misialek, S., Toffaletti, D. L., Klem, T. J., Perfect, J. R., and Davisson, V. J. (1998) Biochemical role of the *Cryptococcus neoformans* ADE2 protein in fungal de novo purine biosynthesis, *Archives of Biochemistry and Biophysics* 351, 123-134.
17. Schaefer, J., Jiang, H., Ransome, A. E., and Kappock, T. J. (2007) Multiple active site histidine protonation states in *Acetobacter acetii* N<sup>5</sup>-carboxyaminoimidazole ribonucleotide mutase detected by REDOR NMR, *Biochemistry* 46, 9507-9512.
18. Silverman, D. N., and Lindskog, S. (1988) The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water, *Accounts of Chemical Research* 21, 30-36.
19. Litchfield, G. J., and Shaw, G. (1971) Part XXXVIII. A kinetic study of the decarboxylation of 5-amino-1-β-D-ribofuranosyl imidazole-4-carboxylic acid 5'-phosphate and related compounds., *Journal of the Chemical Society (B)*, 1474-1484.
20. Smithers, G., and O'Sullivan, W. (1982) <sup>31</sup>P nuclear magnetic resonance study of phosphoribosyldiphosphate and its interaction with magnesium ions, *Journal of Biological Chemistry* 257, 6164-6170.
21. Fersht, A. (2000) *Structure and mechanism in protein science*, W. H. Freeman and Co., New York.
22. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) Evidence for the direct transfer of the carboxylate of N<sup>5</sup>-carboxyaminoimidazole ribonucleotide (N<sup>5</sup>-CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by *Escherichia coli* PurE, an N<sup>5</sup>-CAIR Mutase, *Biochemistry* 38, 3012-3018.
23. Carpena, X., Wiseman, B., Deemagarn, T., Herguedas, B., Ivancich, A., Singh, R., Loewen, P. C., and Fita, I. (2006) Roles for Arg426 and Trp111 in the modulation of NADH oxidase activity of the catalase-peroxidase KatG from *Burkholderia pseudomallei* inferred from pH-induced structural changes, *Biochemistry* 45, 5171-5179.
24. Suzuki, N., Fujimoto, Z., Morita, T., Fukamizu, A., and Mizuno, H. (2005) pH-Dependent structural changes at Ca<sup>2+</sup>-binding sites of coagulation factor IX-binding protein, *Journal of Molecular Biology* 353, 80-87.
25. Kimberly A. Stieglitz, J. X., Evan R. Kantrowitz,. (2009) The first high pH structure of *Escherichia coli* aspartate transcarbamoylase, *Proteins: Structure, Function, and Bioinformatics* 74, 318-327.
26. Baran, K. L., Chimenti, M. S., Schlessman, J. L., Fitch, C. A., Herbst, K. J., and Garcia-Moreno, B. E. (2008) Electrostatic effects in a network of polar and ionizable groups in *Staphylococcal* nuclease, *Journal of Molecular Biology* 379, 1045-1062.
27. Fersht, A. R. (1972) Conformational equilibria in α- and δ-chymotrypsin : The energetics and importance of the salt bridge, *Journal of Molecular Biology* 64, 497-509.
28. Marti, D. N., and Rudolf Bosshard, H. (2003) Electrostatic interactions in leucine zippers: Thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges, *Journal of Molecular Biology* 330, 621-637.
29. Starks, C. M., Francois, J. A., MacArthur, K. M., Heard, B. Z., and Kappock, T. J. (2007) Atomic-resolution crystal structure of thioredoxin from the acidophilic bacterium *Acetobacter acetii*, *Protein Science* 16, 92-98.

30. Hiraoki, T., Brown, S. B., Stevenson, K. J., and Vogel, H. J. (1988) Structural comparison between oxidized and reduced *Escherichia coli* thioredoxin. Proton NMR and CD studies, *Biochemistry* 27, 5000-5008.
31. Hall, N. F., and Sprinkle, M. R. (1932) Relations between the structure and strength of certain organic bases in aqueous solution, *Journal of the American Chemical Society* 54, 3469-3485.
32. Perrins, D. D. (1964) The effect of temperature on the pK values of organic bases, *Australian Journal of Chemistry* 17, 484-488.
33. Nagai, H., Kuwabara, K., and Carta, G. (2008) Temperature dependence of the dissociation constants of several amino acids, *Journal of Chemical and Engineering Data* 53, 619-627.
34. Oscarson, J. L., Wu, G., Faux, P. W., Izatt, R. M., and Christensen, J. J. (1989) Thermodynamics of protonation of alkanolamines in aqueous-solutions to 325°C *Thermochimica Acta* 115, 119-127.
35. Bhattacharya, S., and Lecomte, J. T. (1997) Temperature dependence of histidine ionization constants in myoglobin, *Biophysical Journal* 73, 3241-3256.
36. Singer, G. A. C., and Hickey, D. A. (2003) Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content, *Gene* 317, 39-47.
37. Kreil, D. P., and Ouzounis, C. A. (2001) Identification of thermophilic species by the amino acid compositions deduced from their genomes, *Nucleic Acids Research* 29, 1608-1615.
38. Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R., and Olsen, G. J. (1999) Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species, *Proceedings of the National Academy of Sciences of the United States of America* 96, 3578-3583.
39. Jaenicke, R., Böhm, G., Michael, W. W. A., and Robert, M. K. (2001) [33] Thermostability of proteins from *Thermotoga maritima*, in *Methods in Enzymology*, pp 438-469, Academic Press.

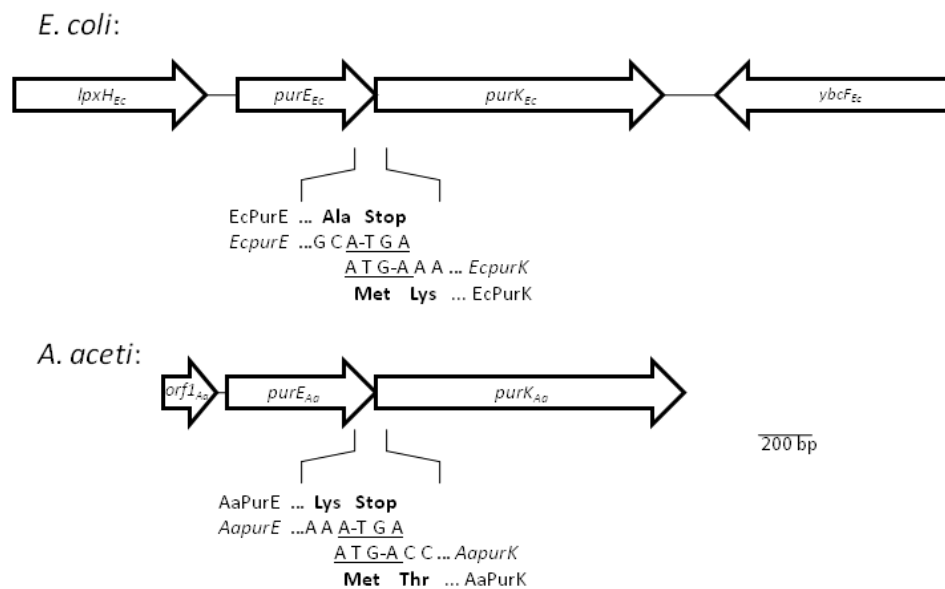
## **Chapter 4**

The *purE-purK* operon of *A. aceti*



## 4.1 Introduction

The genes encoding *purE* and *purK* are usually adjacent in bacteria and appear to form an operon. In *E. coli* and *A. aceti*, the *purE* and *purK* genes overlap (Scheme 4.1), with the final four nucleotides of the *purE* gene comprising the first four nucleotides of the *purK* gene. The *purE* start codons are different: GTG in *A. aceti* and ATG in *E. coli*. As will be discussed in this chapter, there is an open reading frame (*orf1*) upstream of the *A. aceti* *purE* gene that codes for a 68 amino acid protein of unknown function (Scheme 4.1).



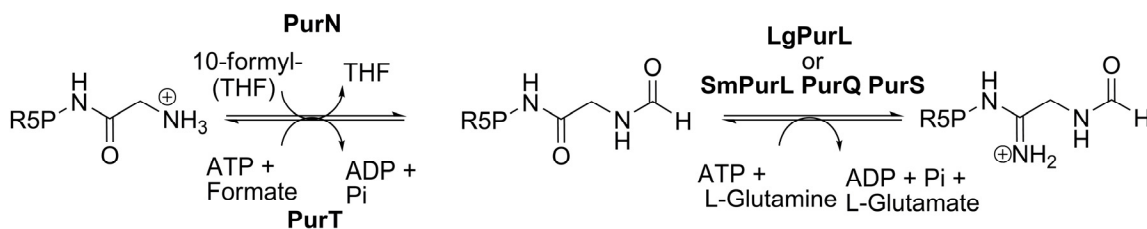
**Scheme 4.1:** Organization of *purE-purK* operon in *A. aceti* and *E. coli*. Arrows represent genes for *purE*, *purK* and *orf1*. The final two codons of *purE* and the first two codons of *purK* are indicated. The bold text represents what the codons code for in their respective proteins. The overlapping nucleotide sequence is underlined.

Overlapping start and stop codons in operons are often indicative of translational coupling. In translational coupling, the translation of the downstream gene (or genes) is dependent on the translation of the preceding gene (1, 2). One typical

effect of translational coupling is a 1:1 stoichiometry for the translationally coupled proteins (3, 4). A second feature of translational coupling is that mutations that cause early termination, or frame-shift mutations, can have a “polar effect” on the translation of downstream genes (1, 5). This polar effect is indicated by a reduction in the translation of downstream genes. Even though these downstream genes are unaltered, the proteins they encode are produced in diminished quantities. A polar effect is evidence for translational coupling. Translational coupling and polar effects have been seen in many operons, including the *E. coli* operons for tryptophan and histidine biosynthesis (1, 5).

The biosynthesis of a compound in different organisms may contain pathway alternates. Pathway alternates may use either altered chemistry or a different synthetic route to obtain the same product. The divergent Class II PurE and Class I PurE/PurK systems are the best known examples of pathway alternates in the biosynthesis of purines; however, there are several other examples. There are two routes for the conversion of 5′phosphoribosyl glycinamide to 5′phosphoribosyl-N-formylglycinamide by the enzymes PurN or PurT (Scheme 4.2) (6). These enzymes are folate-dependent and –independent, respectively. Another divergence is seen in the final two steps of purine biosynthesis between archaea and other bacteria and eukaryotes. Archaea utilize the enzymes PurP and PurO, while other bacteria and eukaryotes use the bifunctional PurH enzyme. Another divergence involves FGAR-AT, the fourth step of purine biosynthesis, in which a single protein (PurL) or a group of three proteins (PurL, PurQ and PurS) performs the same enzymatic reaction.

FGAR-AT converts 5'-phosphoribosyl-N-formylglycinamide to 5'-phosphoribosyl-N-formylglycinamide (Scheme 4.2). In the multiple-protein form, PurL is significantly smaller, and it requires PurS, a small (~80 amino acid) protein that is homologous to the N-terminal domain of larger PurL forms (7). Biochemical and structural studies determined that the complex contains PurL:PurQ:PurS in a ratio of 1:1:2 (8, 9). Crystal structures indicate that PurS interacts with portions of both PurQ and PurL (9).



**Scheme 4.2:** Other divergences in purine biosynthesis.

PurS was originally identified as an open reading frame amid the eleven functionally characterized genes of the single purine biosynthesis operon in *B. subtilis*, and was postulated to be a regulatory protein (10). The involvement of PurS in the multi-protein FGAR-AT activity was identified recently using a single copy genomic insertion strain (7). Simultaneous expression of all three *B. subtilis* genes in *E. coli* using three different promoters proved problematic (8).

The study of strains of bacteria is an integral part of elucidating biosynthetic pathways. The wild form of a bacterium is commonly referred to as “prototrophic.” A prototrophic strain is able to grow when provided with essential minimal nutrients. In the case of *E. coli*, these essential nutrients are a carbon source (e.g., glucose), a source of nitrogen (e.g., ammonium salt), a source of phosphorous (e.g., phosphate), a source of sulfur (e.g., sulfate), the cations potassium, calcium, sodium and magnesium, and

smaller amounts of several transition metals. Provided with these essential nutrients, *E. coli* is able to synthesize all of the compounds required for growth and division. Strains that have an additional nutritional requirement are referred to as “auxotrophic.”

Depending on the strain, the auxotrophic phenotype may only appear under certain conditions. These conditions may be environmental (temperature or atmosphere) or nutritional (in the presence of a specific nutrient(s)). In the simplest case, auxotrophy is caused by a mutation or deletion in a gene that eliminates the activity of the enzyme encoded by that gene. In most cases, the introduction of a gene that encodes for an enzyme with the same activity, or an activity that bypasses the activity of the adulterated gene can lead to “functional complementation” of the auxotrophic phenotype. Complementation of auxotrophy, or the restoration of protrophy, means that the strain can now grow on minimal media. In some cases, the functional complementation of a strain may occur under certain environmental or nutritional conditions.

Purine auxotrophs that result from alteration of the *purK* gene provide a unique example of many of the features described above. Under normal atmosphere, a *purK* auxotroph requires a source of purines for growth; however, when a *purK* auxotroph is grown in the environment of a “candle jar,” the strain is conditionally complemented. A candle jar is a CO<sub>2</sub> rich environment created by placing a lit candle into a sealable container. Once the container is sealed, the candle is allowed to burn until it extinguishes itself. It is proposed that *purK* auxotrophs are able to grow under such conditions owing to the non-enzymatic production of *N*<sup>5</sup>-CAIR from AIR and the

introduced CO<sub>2</sub> (11) . This phenotype led early investigators to conclude that PurK functions as a 'CO<sub>2</sub> carrier' akin to the biotin-dependent components of fatty acid synthesis (12). The introduction of a plasmid bearing a *purK* gene or a Class II *purE* gene leads to complementation of the strain. The activity lost in the auxotroph is restored by the *purK* gene, or is bypassed by the introduction of a Class II *purE*, which does not require *purK*.

Reported herein are initial functional complementation studies of the *E. coli purE* auxotroph PC0135, which were performed to assess the qualitative in vivo activity of *AaPurE* mutants H59D and H59N. Although wild type and the enzymatically active H59D mutant complemented the auxotrophy as expected, so did the apparently inactive H59N mutant. This puzzling observation required the construction of an auxotrophic strain with a less leaky phenotype.

This chapter also reports evidence that *A. acetii purE* is only able to unconditionally complement purine auxotrophy if it is expressed in the *orf1-purE-purK* gene context. This suggested the possibility that *AaPurE* does not function with *EcPurK* as a source of N<sup>5</sup>-CAIR. The possibility that *AaPurE* may require *AaPurK* for activity prompted the construction of new strains to explore the operons of *A. acetii* and *E. coli*. Here we use a combination of stable deletion strains, and strains harboring precise gene replacements to address this question.

Many functional complementation studies focus on either binary growth/no growth determinations or on measuring the time required for the number of cells in cultures of each strain to double (doubling time). Initial complementation studies

showed that many strains had a prolonged “lag phase,” or delay of rapid growth, following inoculation into minimal media. This lag, or acclimation phase, is unusual and is the focus of these studies. The presence of *orf1<sub>Aa</sub>* profoundly affects strains which express *purE<sub>Aa</sub>* or *purE<sub>Aa</sub>purK<sub>Aa</sub>*, and its effects appear dependent on the identity of *purE* and *purK*.

## 4.2 Methods and Materials

*Materials.* All materials were from Sigma Aldrich or Fisher Scientific and of the highest purity unless otherwise noted. Vent DNA polymerase (New England Biolabs) was used in all PCRs for cloning. Restriction endonucleases and T4 DNA ligase were from New England Biolabs. Mutagenesis was performed using Quikchange Mutagenesis kits from Stratagene (Pfu polymerase). Oligodeoxynucleotides (ODNs) were from Integrated DNA Technologies (Tables 4.1 and 4.2). GoTaq polymerase (Promega) was used in PCRs to confirm modifications to genomic DNA and to create products for sequencing. DNazol (Molecular Research Center) was used for isolation of genomic DNA as noted. Plasmid Miniprep kits were from either Qiagen or Sigma. Large quantities of plasmids were isolated using Qiagen Midikits. Other DNA purification kits were from Qiagen. Chemically competent XL-10 Gold cells were from Stratagene. Vectors pET23a and pUC118 were from Novagen. Vector pKOV was from George Church’s lab (13, 14). Centrifugation steps were performed on a tabletop Eppendorf 5810R centrifuge; centrifugation of small aliquots was routinely performed in Eppendorf 5415D centrifuge. Cell disruption by sonication was performed using a Virsonic 100

sonicator. DNA was sequenced by Sanger sequencing at the Protein and Nucleic Acid Chemistry Laboratory at Washington University in St. Louis.

**Table 4.1** Oligodeoxynucleotides (ODNs) used in the construction of deletion strains

ODN	Sequence (5'→3') <sup>a</sup>
1362	TAAAGCCGAGAGTTGTGCACCACAGGAGTTTTAAGACGCATGTCTTCCCGCAATAATC CGgttaggctggagctgcttc
1363	CGTCGGTCTGGGCTTTCGCCAGTCATTCAGACGCTGGTGCAGTTCTTTATCATGAGTC Gattccggggatccgtcgacc
1364	CGTGCTATTCTCTGTGCCCTCTAAAGCCGAGAGTTGTGCACCACAGGAGTTTTAAGACG Cgttaggctggagctgcttc
1365	TACCGGATCGGTAGGCCGATAAGGCGTTTACGCCGCATCCGGCAAGAATAGAGCAC CAGattccggggatccgtcgacc
1374	GGCGCAAAGCCCAGACCGACGAAGTGCTGGAAAACCCGGACCCGCGAGGTGCGGCAT GAAgttaggctggagctgcttc
1375	GTACCGGATCGGTAGGCCGATAAGGCGTTTACGCCGCATCCGGCAAGAATAGAGCA CCAattccggggatccgtcgacc

<sup>a</sup>Sequence shown in upper case is homologous to target sites in the *E. coli* genome. Sequence in lowercase is homologous to pKD13 and was used to amplify the Kan cassette.

**Table 4.2:** ODNs used in the construction of plasmids and sequencing

ODNs	Sequence (5'→3') <sup>a,b,c</sup>
601	CGACGTTGTA AACGACGGCCAGTG
617	CACACAGGAAACAGCTATGACATG
1017	CATTGTTTcTGCAGATCGTACGCCAGAC
1018	GTCTGGCGTACGATcTGCaGAAACAATG
1019	CATTGTTTCAGCAaATCGTACGCCAGAC
1020	GTCTGGCGTACGATtTGCTGAAACAATG
1078	CATTGCCTGAATTCCTCCGTTACAGAC
1079	GCATTGGGCTTAAGCTTAGAATGTAAG
1161	CGGATGACGTTGAGCTGATTCATTTTCC
1162	GAAGCCGTGGTAAACCCGCGTGGATTGG
1163	GATCATCACCGACGGATTATTCAC
1164	CAGCTCACCTAAACGATCAAACAC
1165	CAGCTTTACCAATCGCCAGCGTAC
1166	CATGAAACAGGTTTGCCTCCTC
1167	CTATGTGGGCGTGATGGCGATG
1269	AAAATAGCGGCCGCGTGGCGACTCTTTATTG
1270	GGCTTTGCCAGTCATTCGCCGATTATTGCGGGAAG

1271	GGCGGTTTCGCTCACGCGTCTTAAACTCCTG
1273	CGACTCTAGAGGATCCTTAACCGAACTTACTCTGCGCCC
1274	CCCATTACCGAAGATAAATGAAACAGGTTTGCGTCCTCG
1275	CAGGAGTTTTAAGACGCGTGAGCGAAACCGCC
1276	CGAGGACGCAAACCTGTTTCATTTATCTTCGGTAATGGG
1301	CGACTCTAGAGGATCCTTAACCGAACTTACTC
1302	CTTCCCACAATAATCCGGCGAATGACTGGCGCAAAG
1303	CGGTTTCGCTCACGCGTCTTAAACTCC
1305	CAGGAGTTTTAAGACGCGTGAGCGAAACCG
1370	CTGCCCTCTGCTTCCAGTACATGGCGCGCATTGCAGACG
1371	CGTCTGCAATGCGCGCCATGTACTGGAAGCAGAGGGCAG
1372	CACAGGAGTTTTAAGACGCaTGAGCGAAACCGCC
1373	GGCGGTTTCGCTCAtGCGTCTTAAACTCCTGTG
1386	TATTATTGGATCCATGAAAACACTGGTTGTGGCTC
1388	GAAGTGGTTTCTGCTaACCGCACCCCGATAAAC
1389	GTTTATCGGGGGTGCGGTtAGCAGAAACCACTTC
1390	GATATACATATGTCTTCCCGCAATAATCC
1391	CAAGCTTGCATGCCTGCAGTTAACCGAACTTACTC
1394	CAGGAGTAAGCTTACGCGTGAGCGAAACCGC
1395	CAAGAATAGAGCTCCAGTTAACTGGCAGATTC
1396	GTAACTGG <b>AG</b> CTCTATTCTTGCCG
1397	GACCCGCGAGGTGCGGG <b>GT</b> CGACCTCCTTACATTC
1398	GAATGTAAGGAGGT <b>CG</b> ACCCGCACCTCGCGG
1399	GACCCGCGAGGTGCGG <b>cat</b> GACCTCCTTACATTCTC
1400	GAGAATGTAAGGAGGT <b>catg</b> CCGCACCTCGCGGGTC
1401	GAATCTGCCAGTTAACTGG <b>t</b> GCTCTATTCTTGCCG
1402	CGGCAAGAATAGAG <b>Ca</b> CCAGTTAACTGGCAGATTC
1403	GATTGATTCGCTGCTG

<sup>a</sup>Letters shown in lower case code for mutagenesis. <sup>b</sup>Sequence shown in bold codes for “introduced” cut sites used in cloning. <sup>c</sup>Sequence shown in lowercase bold codes for mutagenesis used to revert introduced restriction sites back to wild type sequence.

*Method of naming strains.* As reported in the following sections, a series of unique strains of *E. coli* were made. The following strains will have the designation CC followed by 4 digits. The first digit will denote the type of strain. A deletion strain will be denoted by 0, and a precise gene replacement strain will be denoted by a 1. The second digit will correspond to the method used to create the strain. Strains created by



the method of Datsenko and Wanner (15) will be denoted by a 1, while strains created using the pKOV precise gene replacement method (14) will be denoted by a 2. The final two digits are used to differentiate strains.

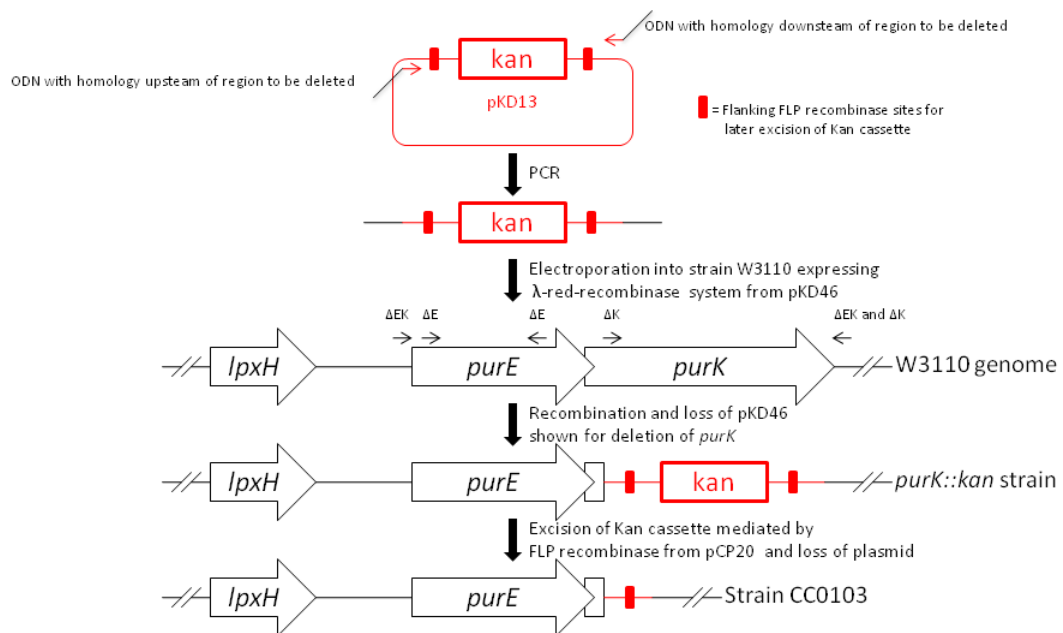
**Table 4.3:** Description of strains

Strain	Description <sup>a</sup>	Reference or Source
PCO135	<i>lacY1 purE55 glnV44(AS) gal-3 rpsL117(str<sup>R</sup>) malT1 (Lam<sup>R</sup>) xlyA7 mtIA2 thi-1</i>	<i>E. coli</i> genetic Stock Center
DH5α	<i>supE44 ΔlacU169 (φ80 lacZΔM15) hsdR recA1 endA1 gyrA96 thi-1 relA1</i>	
XL-1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI<sup>q</sup> ZΔM15 Tn10 (Tet<sup>r</sup>)]</i>	Stratagene
XL-10 Gold	<i>Tet<sup>r</sup> Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI<sup>q</sup> ZΔM15 Tn10 (Tet<sup>r</sup>) Amy Cam<sup>r</sup>]</i>	Stratagene
XL-10 Gold KAN	<i>Tet<sup>r</sup> Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI<sup>q</sup> ZΔM15 Tn10 (Tet<sup>r</sup>) Tn5 (Kan<sup>r</sup>) Amy]</i>	Stratagene
W3110	<i>LAM- In(rrnD-rrnE)1 rph-1</i>	
CC0101 <sup>b</sup>	W3110 $\Delta purE_{Ec}$	This study
CC0102 <sup>c</sup>	W3110 $\Delta purE_{Ec} \Delta purK_{Ec}$	This study
CC0103 <sup>d</sup>	W3110 $\Delta purK_{Ec}$	This study
CC1201 <sup>e</sup>	W3110 $purE_{Ec} \langle \rangle purE_{Aa}$	This study
CC1202 <sup>f</sup>	W3110 $purE_{Ec} \langle \rangle purE_{Aa(ATG)}$	This study
CC1203 <sup>g</sup>	W3110 $purK_{Ec} \langle \rangle purK_{Aa}$	This study
CC1204 <sup>h</sup>	W3110 $purE_{Ec} purK_{Ec} \langle \rangle purE_{Aa} purK_{Aa}$	This study
CC1205 <sup>i</sup>	W3110 $purE_{Ec} purK_{Ec} \langle \rangle purE_{Aa(ATG)} purK_{Aa}$	This study

<sup>a</sup>  $\langle \rangle$  symbol denotes precise gene replacement; subscript indicates origin of gene (<sub>Aa</sub>) refers to *A. acetii*, (<sub>Ec</sub>) refers to *E. coli*. The *purE* gene of *A. acetii* begins with a GTG start codon, subscript (<sub>ATG</sub>) denotes strains in which the start codon for *purE*<sub>Aa</sub> has been changed to ATG. <sup>b-i</sup> Alternate strain names: <sup>b</sup>CC0101 a.k.a. 1611, <sup>c</sup>CC0102 a.k.a. 2C311, <sup>d</sup>CC0103 a.k.a. DK111, <sup>e</sup>CC1201 a.k.a. Church 31, <sup>f</sup>CC1202 a.k.a. Church ATG 61, <sup>g</sup>CC1203 a.k.a. Church K2121, <sup>h</sup>CC1204 a.k.a. Church EK 1111, <sup>i</sup>CC1205 a.k.a. Church ATG EK 1131.

*Method for the creation of deletion strains.* Deletion strains were made using the  $\lambda$ -red recombinase system of Datsenko and Wanner (15) with guidance from Peter Chivers, Department of Biochemistry, Washington University in St. Louis. The general method is described below and illustrated in Scheme 4.3. A specific description for the creation of each strain is given in the following sections.

The kanamycin resistance gene (Kan cassette) of plasmid pKD13 was amplified by PCR using Vent DNA polymerase and the appropriate ODNs as described in the following sections. ODNs were designed to amplify the Kan cassette with flanking 60-bp sequences derived from regions flanking the genomic region being deleted. PCRs contained 4 units of Vent, 0.4 pmol of both ODNs, 0.8 nmol of



**Scheme 4.3:** Method of gene deletion. Deletion strains created using  $\lambda$ -Red recombinase system of Datsenko and Wanner (15), illustrated for deletion of *purK*. Arrows above W3110 genome denote regions of homology used in ODNs for amplification of Kan cassette for the creation of each strain as noted. Representations are not to scale.

each dNTP, at a total volume of 400  $\mu\text{L}$  (for four 100  $\mu\text{L}$  reactions). The PCR was performed with an annealing temperature of 45  $^{\circ}\text{C}$  and melting, annealing, and extension times of 30, 60 and 120 s for 30 cycles. Following PCR, 80 units of *DpnI* were added, and the reaction was digested for 4 h at 37  $^{\circ}\text{C}$ . The entire reaction was run on a 1% agarose gel and the PCR product was extracted from the gel. The product was eluted in a minimal volume of elution buffer. The purified PCR product (2-4  $\mu\text{L}$ ) was electroporated into freshly prepared, arabinose induced, W3110/pKD46 cells. The pKD46 plasmid contains the  $\lambda$  red recombinase system, which facilitates the homologous recombination of the linear PCR product into the genome at the gene targeted for deletion.

The electroporated cells were recovered in SOC broth at 37  $^{\circ}\text{C}$  for 3 h and spread on LB plates supplemented with 50  $\mu\text{g}/\text{mL}$  kanamycin (Kan). The plates were placed in a 37  $^{\circ}\text{C}$  incubator overnight. To select for genomic incorporation of the Kan cassette and loss of plasmid pKD46, colonies were suspended in a small volume of LB broth (20  $\mu\text{L}$ ), and 2  $\mu\text{L}$  of the suspension was spotted on a LB plate supplemented with 100  $\mu\text{g}/\text{mL}$  Amp and an LB plate supplemented with 50  $\mu\text{g}/\text{mL}$  Kan, and the plates were placed in a 42  $^{\circ}\text{C}$  incubator. Colonies that were Kan resistant and Amp sensitive were selected. The spot grown on the Kan plate was used to inoculate 5 mL LB/Kan that was grown at 37  $^{\circ}\text{C}$ . Whole colony PCR was performed to confirm the insertion of the Kan cassette into the genome. Chemically competent cells of a culture containing the Kan cassette were transformed with the plasmid pCP20. Cells were recovered in SOC broth and spread on LB agar supplemented with either 20  $\mu\text{g}/\text{mL}$  chloramphenicol or 100  $\mu\text{g}/\text{mL}$  Amp and

grown at 30 °C for 18 h to select for transformants containing plasmid pCP20. Several colonies were selected and used to inoculate 5 mL LB cultures that were grown at 42 °C for at least 3 h. Dilutions of each culture were made in LB, and a portion of dilutions were placed on LB plates and LB/Kan plates. Plates were placed at 42 °C. From dilutions that gave few to no colonies on LB/Kan/agar, single colonies from the corresponding LB plate were used to inoculate 5 mL LB. Cultures were grown to turbidity, and glycerol stocks were made. Whole colony PCR was performed on the turbid cultures to confirm excision of the Kan cassette, and the PCR product was sent for sequencing.

*Creation of E. coli  $\Delta purE_{Ec}$  strain* - ODNs were designed to delete the majority of the *purE<sub>Ec</sub>* gene, leaving the first 21 bp and the last 98 bp of the gene separated by the 82 bp scar left after excision of the Kan cassette. Primers were designed to leave the *purK<sub>Ec</sub>* gene in the same reading frame in reference to the start codon of *purK<sub>Ec</sub>*. The Kan cassette was amplified using ODNs 1362 and 1363, giving a product of 1.5 kb. Insertion of the Kan cassette into the genome of W3110 was confirmed by whole colony PCR, using ODNs 1161 and 1164 and Taq DNA polymerase, with 58 °C annealing temperature and melting, annealing and extension times of 30, 60, and 120 s for 30 cycles, giving a product of 2.0 kb. Finally, excision of the Kan cassette was confirmed by whole colony PCR, using the same conditions, and giving a product of 0.7 kb. The process yielded the  $\Delta purE_{Ec}$  strain CC0101.

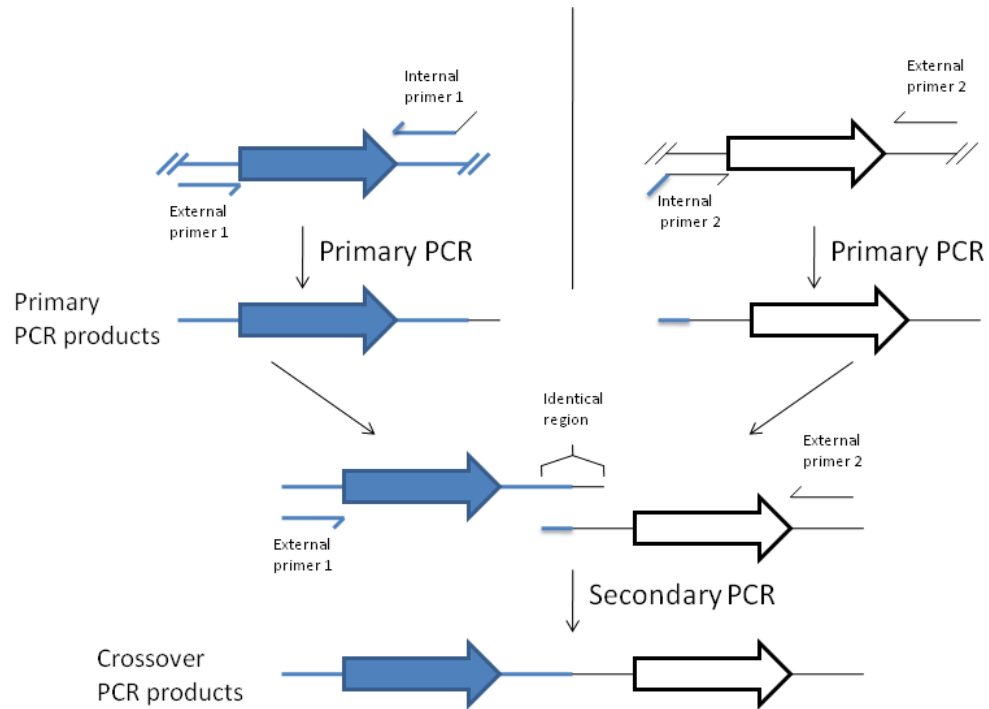
*Creation of E. coli  $\Delta purE_{Ec}\Delta purK_{Ec}$  strain.* ODNs were designed to delete all of the *purE<sub>Ec</sub>purK<sub>Ec</sub>* coding region in W3110, leaving only the 82 bp scar after excision of

the Kan cassette. The Kan cassette was amplified by PCR using ODNs 1364 and 1365, giving a product of 1.5 kb. Insertion of the Kan cassette into the genome of W3110 was confirmed by whole colony PCR using ODNs 1161, 1162 and Taq DNA polymerase, with a 58 °C annealing temperature and melting, annealing and extension times of 30, 60, and 120 s for 30 cycles, giving a product of 2.0 kb. Finally, excision of the Kan cassette was confirmed by whole colony PCR using the same conditions, giving a product of 0.5 kb. The process yielded the  $\Delta purE_{Ec}\Delta purK_{Ec}$  strain CC0102.

*Creation of E. coli  $\Delta purK_{Ec}$  deletion strain.* ODNs were designed to delete the majority the  $purK_{Ec}$  gene, leaving the first 6 bp of the gene upstream of the 82 bp scar remaining after excision of the Kan cassette. The Kan cassette was amplified using ODNs 1374 and 1375, giving a product of 1.5 kb. Insertion of the Kan cassette into the genome of W3110 was confirmed by whole colony PCR using ODNs 1161, 1162 and Taq DNA polymerase, with a 58 °C annealing temperature and melting, annealing and extension times of 30, 60, and 180 s for 30 cycles, giving a product of 2.3 kb. Finally, excision of the Kan cassette was confirmed by whole colony PCR using the same conditions, giving a product of 0.8 kb. The process yielded the  $\Delta purK_{Ec}$  strain CC0103.

*Method for construction of plasmids used in modifying genomic DNA.* Two methods were employed to construct the plasmids used in modifying the genomic DNA of *E. coli*. As described in the following sections, the plasmid for the precise insertion of  $purE_{Aa}$  and a plasmid constructed for deletion of  $purE_{Ec}$  were obtained by the technique of cross-over PCR (16, 17). This technique allows for the joining of two separate pieces of DNA (Scheme 4.4). For this technique, internal ODNs for the amplification of the two

fragments of DNA were designed with at least 20 bp regions of identical sequence. Initial PCR products were combined in a second PCR containing external primers. The regions of identical sequence allowed for the joining of the two DNA fragments.



**Scheme 4.4:** Illustration of cross-over PCR.

The plasmids for the precise insertion of *purE<sub>Aa</sub>purK<sub>Aa</sub>* and *purK<sub>Aa</sub>* were constructed in parts by the use of existing restriction sites (*Mlu*I and *Nco*I) present in the expected sequence of the final product, and by the use of introduced restriction sites (*Sac*I, *Sal*I, and *Hind*III). The introduced sites were introduced by primers and were not present in the expected sequence of the final product. The introduced restriction sites, *Sac*I and *Sal*I, were then changed to the desired final sequence by Quikchange mutagenesis. The introduced *Hind*III site was lost during subcloning.

**Table 4.4** Plasmids used in this Chapter<sup>a</sup>

Plasmid	Description	Source
pNC2	<i>purE<sub>Ec</sub>purK<sub>Ec</sub></i> ligated into NdeI and HindIII sites in pET23a	(11)
pKOV	Gene replacement vector	(13, 14)
pJK173	<i>orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub></i> ligated into SphI site in pUC118	(18)
pJK324	<i>purE<sub>Aa</sub></i> ligated into EcoRI and HindIII sites in pUC118	This study
pJK340	<i>purEH59N<sub>Aa</sub></i> ligated into EcoRI and HindIII sites in pUC118	This study
pJK346	<i>purEH59D<sub>Aa</sub></i> ligated into EcoRI and HindIII sites in pUC118	This study
pJK347	<i>orf1<sub>Aa</sub>purEH59D<sub>Aa</sub>purK<sub>Aa</sub></i> ligated into SphI site in pUC118	This study
pJK348	<i>orf1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK<sub>Aa</sub></i> ligated into SphI site in pUC118	This study
pJK386	<i>purE<sub>Aa</sub>purK<sub>Ec</sub></i> ligated into BamHI and filled Sall site in pET23a	This study
pJK388	<i>lpxH<sub>Ec</sub>purE[30]<sub>Ec</sub>purK<sub>Ec</sub></i> ligated into NotI and BamHI sites in pET23a	This study
pJK389	<i>lpxH<sub>Ec</sub>purE<sub>Aa</sub>purK<sub>Ec</sub></i> ligated into NotI and BamHI sites in pET23a	This study
pJK390	<i>lpxH<sub>Ec</sub>purE[30]<sub>Ec</sub>purK</i> ligated into NotI and BamHI sites in pKOV	This study
pJK391	<i>lpxH<sub>Ec</sub>purE<sub>Aa</sub>purK<sub>Ec</sub></i> ligated into NotI and BamHI sites in pKOV	This study
pJK412	<i>orf1<sub>Aa</sub>purE<sub>Aa</sub></i> ligated into PstI and HindIII sites in pUC118	This study
pJK413	<i>lpxH<sub>Ec</sub>purE<sub>Aa(ATG)</sub>purK<sub>Ec</sub></i> ligated into NotI and BamHI sites in pET23a	This study
pJK414	<i>lpxH<sub>Ec</sub>purE<sub>Aa(ATG)</sub>purK<sub>Ec</sub></i> ligated into NotI and BamHI sites in pKOV	This study
pJK415	<i>orf1<sub>Aa</sub>purE[32]<sub>Aa</sub>purK<sub>Aa</sub></i> ligated into PstI site in pUC118	This study
pJK418	<i>lpxH<sub>Ec</sub>purE<sub>Ec</sub>purK<sub>Aa</sub>ycbF<sub>Ec</sub></i> ligated into NotI and BamHI sites in pET23a	This study
pJK419	<i>purE<sub>Ec</sub>purK[62]<sub>Ec</sub></i> in pET23a	This study
pJK420	<i>lpxH<sub>Ec</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>ycbF<sub>Ec</sub></i> ligated into NotI and BamHI sites in pET23a	This study
pJK421	<i>lpxH<sub>Ec</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>ycbF<sub>Ec</sub></i> ligated into NotI and BamHI sites in pKOV	This study
pJK426	<i>purEH45N<sub>Ec</sub>purK<sub>Ec</sub></i> ligated into NdeI and HindIII sites in pET23a	This study
pJK429	<i>purE<sub>Aa</sub>purK<sub>Aa</sub></i> ligated into HindIII and SacI sites in pET23a	This study
pJK430	<i>ycbF<sub>Ec</sub></i> ligated into SacI and BamHI sites in pET23a	This study
pJK431	<i>purK<sub>Aa</sub></i> ligated into Sall and SacI sites in pET23a	This study
pJK432	<i>lpxH<sub>Aa</sub>purE<sub>Ec</sub></i> ligated into NotI and Sall sites in pET23a	This study
pJK435	<i>purEH45N<sub>Ec</sub>purK[62]<sub>Ec</sub></i> in pET23a	This study
pJK440	<i>purE<sub>Aa</sub>purK<sub>Aa</sub>ycbF<sub>Ec</sub></i> <sup>b</sup> ligated into HindIII and BamHI sites in pET23a	This study
pJK442	<i>lpxH<sub>Ec</sub>purE<sub>Aa</sub>purK<sub>Aa</sub></i> ligated into NotI and SacI sites in pET23a	This study
pJK443	<i>lpxH<sub>Ec</sub>purE<sub>Ec</sub>purK<sub>Aa</sub></i> <sup>b</sup> ligated into NotI and Sall sites in pET23a	This study
pJK444	<i>purE<sub>Aa</sub>purK<sub>Aa</sub>ycbF<sub>Ec</sub></i> ligated into HindIII and BamHI sites in pET23a	This study
pJK446	<i>lpxH<sub>Ec</sub>purE<sub>Ec</sub>purK<sub>Aa</sub></i> ligated into NotI and Sall sites in pET23a	This study
pJK448	<i>lpxH<sub>Ec</sub>purE<sub>Ec</sub>purK<sub>Aa</sub>ycbF<sub>Ec</sub></i> ligated into NotI and BamHI sites in pKOV	This study
pJK455	<i>orf1<sub>Aa</sub>purE[64]<sub>Aa</sub></i> in pUC118	This study

<sup>a</sup>Two letter subscript code indicates origin of gene, (<sub>Ec</sub>) *E. coli* or (<sub>Aa</sub>) *A. aceti*. The *purE<sub>Aa</sub>* gene begins with a GTG start codon, subscript (<sub>ATG</sub>) denotes strains in which start codon for *purE<sub>Aa</sub>* has been changed to ATG. In some plasmids the majority of a gene was deleted. A number in [ ] represents number of amino acids from protein still encoded. <sup>b</sup>Inserts contain internal introduced restriction sites (SacI or Sall) that were introduced in cloning and later reverted to wild type sequence by mutagenesis.

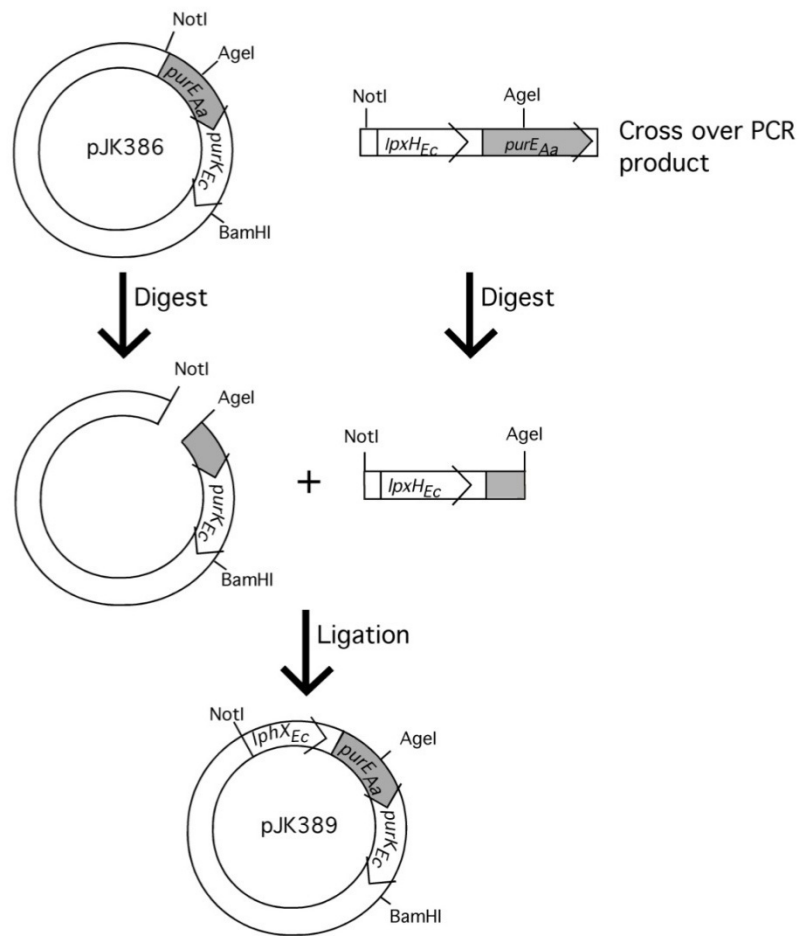
Listed below are the conditions used to construct plasmids necessary for obtaining the precise gene replacement plasmids pJK391, pJK413, pJK421, and pJK448. Also described are the conditions to create plasmid pJK390, a plasmid initially created for the precise deletion of *purE<sub>EC</sub>*. For clarity, plasmids will be reported under subsections according to the product for which they were made: a) pJK391 and pJK414; b) pJK421, c) pJK448 and d) pJK390

**a) Construction of plasmid pJK391 and pJK414 for the precise insertion of *purE<sub>Aa</sub>* into the *E.coli* genome**

The cloning strategy in this section is illustrated by Figure 4.1 shown on the following page.

*pJK386*. The 1.1 kb *purK<sub>EC</sub>* gene was amplified by PCR from plasmid pNC2, using ODNs 1273 and 1274, with an annealing temperature of 60 °C and melting, annealing, and extension times of 30, 60 and 90 s for 25 cycles. The PCR (50 µL) contained 16.5 pmol each ODN, 10 nmol of each dNTP, and 1 unit of Vent DNA polymerase. The crude product was used in subsequent cross-over PCR. The 0.6 kb *purE<sub>Aa</sub>* gene was amplified by PCR from plasmid pJK173, using ODNs 1275 and 1276, with an annealing temperature of 55 °C and melting, annealing, and extension times of 30, 60, and 30 s for 25 cycles. The PCR (100 µL) contained 37 pmol each ODN, 20 nmol of each dNTP and 2 units of Vent DNA polymerase. A portion of crude product was saved, and the majority was purified with a Qiagen PCR purification kit. The purified product was





**Figure 4.1:** Illustration of method used to construct plasmid pJK389 for the precise insertion of *purE<sub>Aa</sub>*. Plasmid pJK386 was obtained by half-blunt ligation of the *purE<sub>Aa</sub>* + *purK<sub>EC</sub>* cross-over product, into pET23a using the BamHI restriction site. The *lpxH<sub>EC</sub>* + *purE<sub>Aa</sub>* cross-over product was then cloned into pJK386 using NotI restriction site of pet23a and an internal Agel restriction site present in *purE<sub>Aa</sub>*. Plasmid pJK389 contains the insert which was cloned into the pKOV vector to yield plasmid pJK391, and was mutated to create plasmid pJK413 which contains the insert which was cloned into pKOV vector to yield plasmid pJK414.

used in the subsequent cross-over PCR. A 1.7 kb cross-over product was obtained by PCR using 1 ng of each of the *purK<sub>EC</sub>* and *purE<sub>Aa</sub>* products as template, and ODNs 1273 and 1305 (ODN 1275 was replaced with 1305 to avoid primer dimers), with an annealing temperature of 60°C and melting, annealing, and extension times of 30, 60 and 120 s for

25 cycles. The PCR (50  $\mu$ L) contained 10 pmol each ODN, 20 nmol of each dNTP and 1 unit of Vent DNA polymerase. The cross-over product was purified by Qiagen PCR purification kit. The cross-over PCR product was digested with BamHI and purified following digestion. Vector pET23a was prepared for ligation by digestion with Sall and purification using a Qiagen PCR purification kit. The 5' overhangs left by digestion were filled in using Vent DNA polymerase in a reaction at a total volume of 107  $\mu$ L that contained 2 units of Vent, and 13 nmol of each dNTP, at 72  $^{\circ}$ C for 15 min. The filled-in vector was purified using a Qiagen PCR purification kit, and then digested with BamHI and treated with 1.2 units of calf intestinal alkaline phosphatase (CIP). The vector was then purified with a PCR purification kit. The cross-over product was half blunt-ligated into the BamHI restriction site of the prepared vector in a 10  $\mu$ L reaction containing 200 units of T4 ligase at 15  $^{\circ}$ C to yield plasmid pJK386. Sequencing of plasmid pJK386 indicated the plasmid contained a deletion, but this was within a region not needed for subsequent cloning steps.

*pJK389*. The 0.9 kb region upstream of *purE<sub>Ec</sub>*, including the entire *lpxH<sub>Ec</sub>* gene, was amplified by PCR using gDNA and ODNs 1269 and 1271, with an annealing temperature of 65  $^{\circ}$ C and melting, annealing and extension times of 30, 60, and 90 s for 25 cycles in a total volume of 50  $\mu$ L. The PCR contained 16.5 pmol of each ODN, 10 nmol of each dNTP, and 1 unit of Vent DNA polymerase. A 1.4 kb cross-over product was obtained by PCR using 0.8 ng of the *lpxH<sub>Ec</sub>* crude PCR product and 0.4 ng of the *purE<sub>Aa</sub>* crude PCR product as template, and ODNs 1269 and 1276, with an annealing temperature of 65  $^{\circ}$ C and melting, annealing and extension times of 30, 60 and 90 s for

30 cycles. The 200  $\mu$ L PCR (four 50 $\mu$ L reactions) contained 40 pmol of each ODN, 80 nmol of each dNTP, and 4 units of Vent DNA polymerase. The cross-over product was purified using a PCR purification kit. The product was sequentially digested with NotI and AgeI, with enzyme removal using a PCR purification kit between digests. The digested product was run on a 1% agarose gel and the 1.0 kb fragment was purified by a Qiagen gel extraction kit. Plasmid pJK386 was sequentially digested with NotI and AgeI as above. The digested plasmid was run on a 1% agarose gel, and the 3.9 kb fragment of vector was purified by gel extraction kit. The purified 1.0 kb fragment of the cross-over product was ligated into the prepared vector using the AgeI and NotI sites, yielding plasmid pJK389. Sequencing of plasmid pJK389 indicated a point mutation that alters the *lpxH* Ala7 codon (GCA) to a threonine codon (ACA).

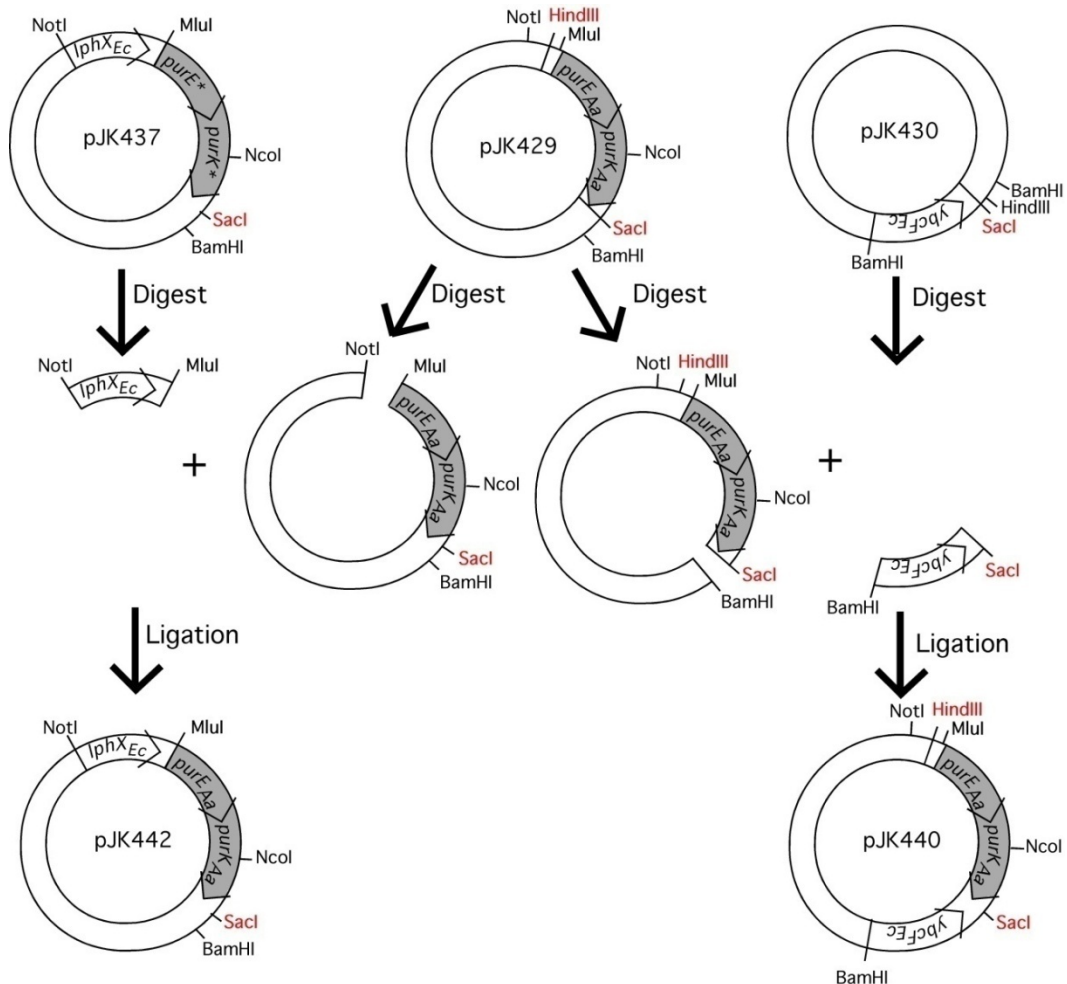
*pJK391.* Plasmid pJK389 was digested with BamHI and NotI, and the insert was purified by gel extraction. The insert was ligated into vector pKOV using the BamHI and NotI restriction sites, yielding plasmid pJK391.

*pJK413.* The start codon of *purE<sub>Aa</sub>* in plasmid pJK389 was changed from GTG to ATG by Quikchange mutagenesis using plasmid pJK389 and ODNs 1372 and 1373. The mutagenesis yielded plasmid pJK413. Sequencing indicated that the desired sequence was obtained.

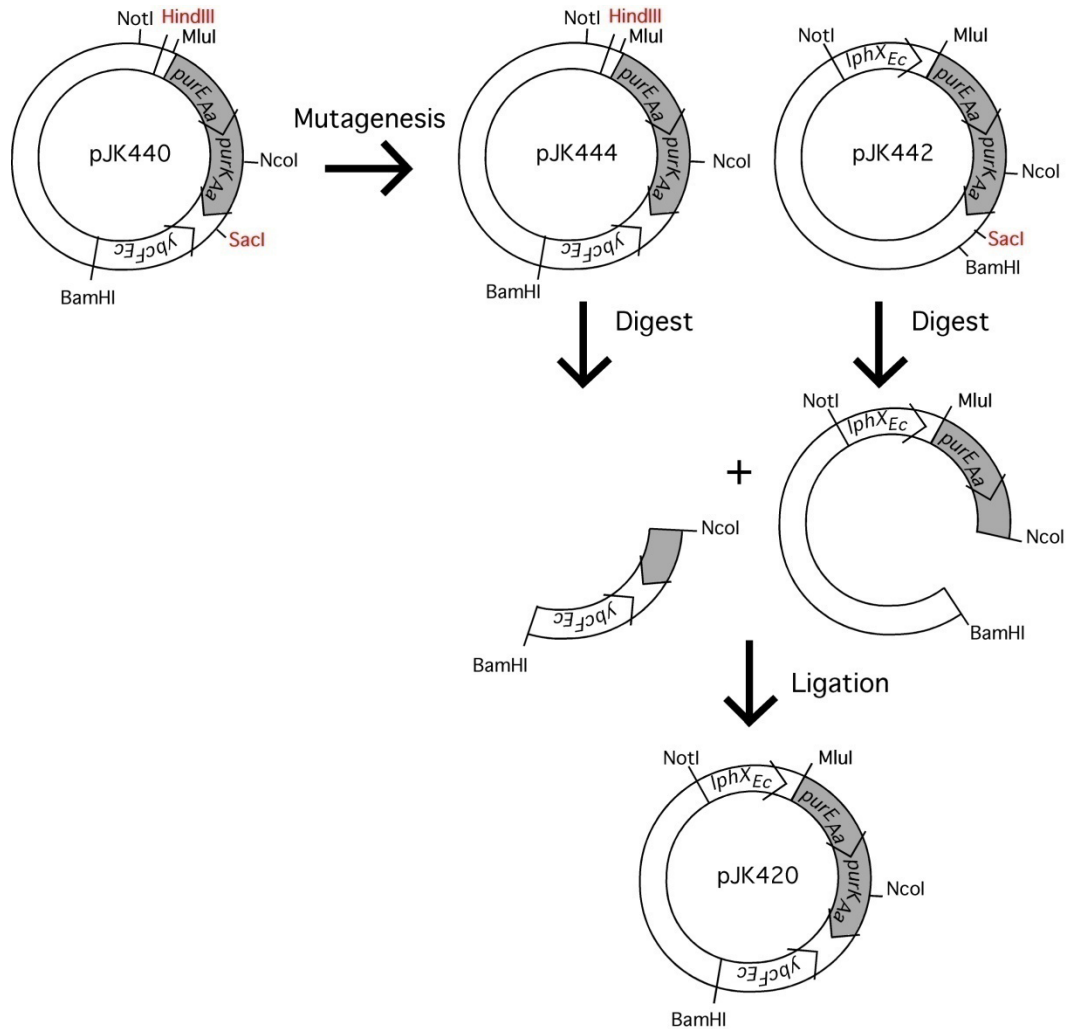
*pJK414.* Plasmid pJK413 was digested with BamHI and NotI, and the insert was purified by gel extraction. The insert was cloned into vector pKOV using the BamHI and NotI restriction sites, yielding plasmid pJK414.

**b) Construction of plasmid pJK421 for the precise insertion of *purE<sub>Aa</sub>purK<sub>Aa</sub>* into the *E. coli* genome**

The cloning strategy used in this section is illustrated by Figures 4.2 and 4.3.



**Figure 4.2:** Assembly of the insert of plasmid pJK421 (Part 1). “Introduced” restriction sites are shown in Red. The vector pET23a does not contain an MluI restriction, however, an MluI site is present at the junction of *purE<sub>Aa</sub>* and the region upstream of *purE<sub>Ec</sub>*. This site was used for the assembly of the region upstream of *purE<sub>Ec</sub>* and *purE<sub>Aa</sub>-purK<sub>Aa</sub>* segments. The upstream region was initially cloned into a subclone of pJK429 that was found to have mutations. The NotI-MluI insert of pJK437 was cloned into plasmid pJK429 using the NotI and MluI restriction sites to yield plasmid pJK442. The “introduced” SacI site was used to assemble the *purE<sub>Aa</sub>-purK<sub>Aa</sub>* segment and the region downstream of *purK<sub>Ec</sub>*. The BamHI-SacI insert of plasmid pJK430 was cloned into pJK429 using the BamHI and SacI restriction sites to yield plasmid pJK440. Construction of insert for pJK421 continued in Figure 4.3.



**Figure 4.3:** Assembly of insert of pJK421 (Part 2). “Introduced” restriction sites shown in red. Quikchange mutagenesis was used to revert the “introduced” *SacI* restriction site to wild type sequence. The inserts of plasmids pJK444 and 442 were combined using a unique *NcoI* restriction site present in *purK<sub>Aa</sub>*. The BamHI-*NcoI* insert of pJK444 was cloned into pJK442 using the BamHI and *NcoI* restriction sites to yield plasmid pJK420. Plasmid pJK420 contains the insert which was cloned into the pKOV vector.

*pJK429*. A 1.8 kb product containing the *purE<sub>Aa</sub>purK<sub>Aa</sub>* genes was amplified by PCR from plasmid pJK173 and ODNs 1394 and 1395, using an annealing temperature of 58 °C with melting, annealing and extension times of 30, 60, and 110 s for 30 cycles. The 100 μL PCR contained 60 pmol of each ODN, 20 μmol of each dNTP, and 2 units of Vent

DNA polymerase. The PCR product was purified by gel extraction, and digested with HindIII and SacI. The digested product was purified by PCR purification kit, and cloned into the HindIII and SacI restriction sites of pET23a, yielding pJK429. ODN 1394 introduces a HindIII cut site immediately upstream of a unique MluI restriction site that is created at the junction of *purE<sub>Aa</sub>* and the region upstream of *purE<sub>Ec</sub>*. The MluI site will be used below. Sequencing of plasmid pJK429 indicated a silent point mutation that alters the *purK* Arg215 codon CGA to CGG.

*pJK437*. A 0.9 kb product containing the region upstream of *purE<sub>Ec</sub>* including the entire *lpxH<sub>Ec</sub>* gene, was amplified by PCR from plasmid gDNA and ODNs 1269 and 1271, and an annealing temperature of 65 °C with melting, annealing and extension times of 30, 60, and 60 s for 25 cycles. The PCR (100 µL) contained 66 pmol of each ODN, 40 µmol of each dNTP, and 2 units of Vent DNA polymerase. The PCR product was purified using a gel extraction kit. The product was digested with NotI and MluI and cloned into the restriction sites of a mutant subclone of pJK429, yielding pJK437. Sequencing indicated that plasmid pJK437 contained the desired sequence.

*pJK430*. A 1.1 kb product containing the region downstream of *purK<sub>Ec</sub>* including the entire divergently-oriented *ybcF<sub>Ec</sub>* gene, was amplified by PCR from plasmid gDNA and ODNs 1386 and 1396, and an annealing temperature of 50 °C with melting, annealing and extension times of 30, 60, and 75 s for 30 cycles. The PCR (50 µL) contained 30 pmol of each ODN, 20 µmol of each dNTP, and 1 units of Vent DNA polymerase. The PCR product was purified using a gel extraction kit, and digested with SacI and BamHI. The digest was purified by PCR purification kit and cloned into the SacI

and BamHI restriction sites of pET23a, yielding pJK430. Sequencing indicated that pJK430 contained the expected sequence.

*pJK440.* The SacI-BamHI insert of pJK430 was cloned into the SacI and BamHI restriction sites of pJK429 to yield pJK440

*pJK442.* The NotI-MluI insert of pJK437 was cloned into the NotI and MluI restriction sites of pJK429 to yield plasmid pJK442.

*pJK444.* Quikchange mutagenesis was used to revert the SacI of pJK440 back wild type sequence using ODNs 1401 and 1402. Sequencing of pJK444 indicated the desired sequence was obtained.

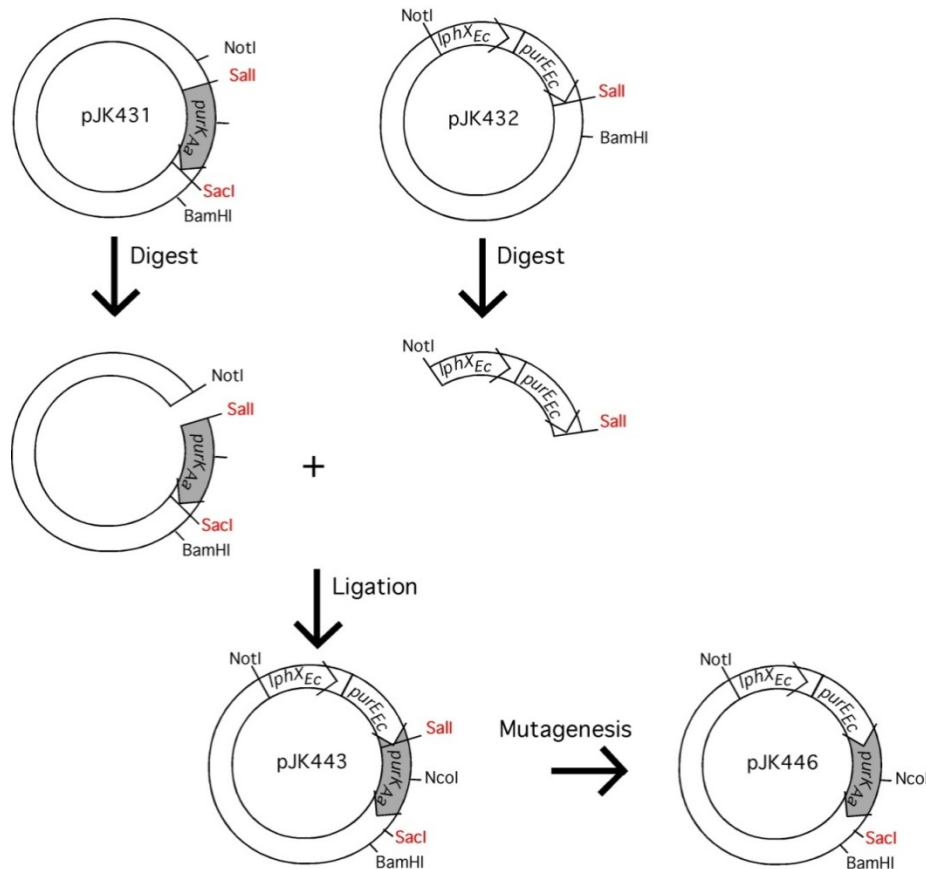
*pJK420.* The inserts of plasmids pJK444 and pJK442 were then combined using a unique NcoI restriction site found in *purK<sub>Aa</sub>*. The BamHI-NcoI insert of pJK444 was cloned into pJK442 using the BamHI and NcoI restriction sites to yield plasmid pJK420.

*pJK421.* The BamHI-NotI insert of pJK420 was cloned into vector pKOV using the BamHI and NotI restriction sites to yield plasmid pJK421.

### **c) Construction of plasmid pJK448 for the precise insertion of *purK<sub>Aa</sub>* into the *E. coli* genome**

The cloning strategy used in this section is illustrated by Figures 4.4 and 4.5.

*pJK432.* A 1.4 kb product containing the region upstream of *purK<sub>Ec</sub>*, including the entire *lpxH<sub>Ec</sub>* and *purE<sub>Ec</sub>* genes, was amplified by PCR using *E. coli* gDNA and ODNs 1269 and 1398, with an annealing temperature of 62 °C and melting, annealing and extension times of 30, 60, and 85 s for 30 cycles. The PCR (50 µL) contained 30 pmol of each ODN, 20 µmol of each dNTP, and 1 units of Vent DNA polymerase. The PCR product

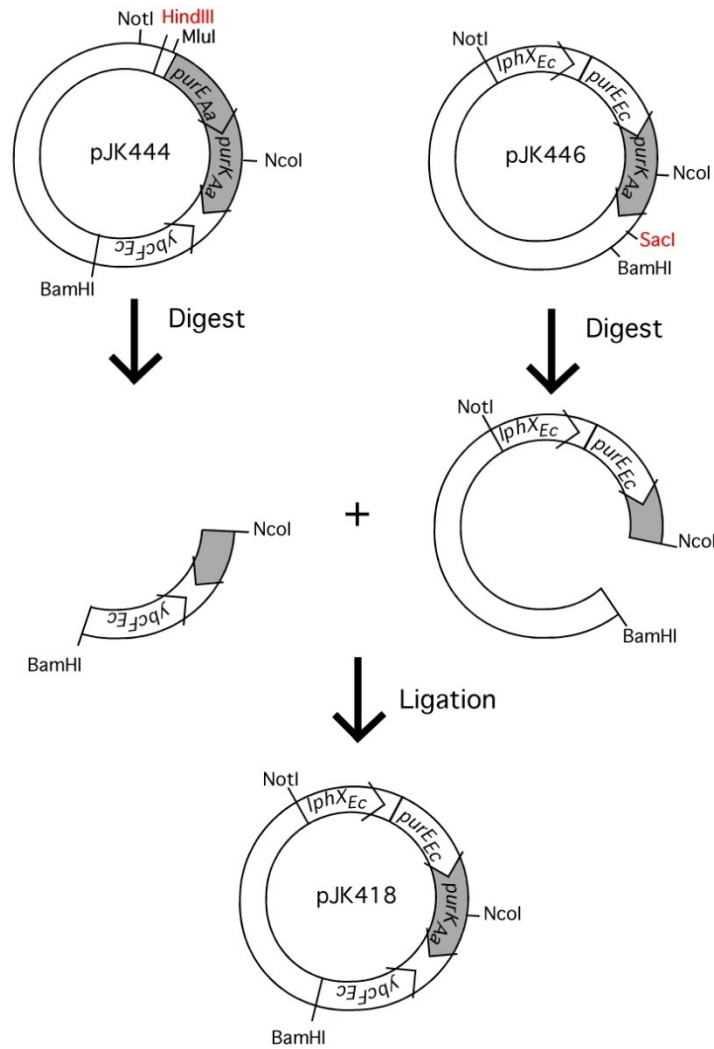


**Figure 4.4:** Assembly of insert of pJK448 (Part 1). “Introduced” restriction sites are shown in red. The “introduced” Sall restriction site was used to assemble the *purK<sub>Aa</sub>* (pJK431) and the region upstream of *purK<sub>Ec</sub>* (pJK432) segments. The NotI-Sall insert of pJK432 was cloned into pJK431 using the NotI and Sall restriction sites to yield plasmid pJK443. The “introduced” Sall site was reverted back to wild type sequence by Quikchange mutagenesis to yield plasmid pJK446. Construction of insert for pJK448 is continued in Figure 4.5.

was purified by a gel extraction kit, and was digested with NotI and Sall. The digest was purified with a PCR purification kit and was cloned into pET23a using the NotI and Sall restriction sites, yielding pJK432. Sequencing indicated that plasmid pJK432 contained the expected sequence.

*pJK431*. A 1.2 kb product containing the *purK<sub>Aa</sub>* gene was amplified by PCR using plasmid pJK173 and ODNs 1395 and 1397, with an annealing temperature of 58 °C and





**Figure 4.5:** Assembly of insert of pJK448 (Part 2). The inserts of plasmids pJK444 and 446 were combined using a unique NcoI restriction site present in *purK<sub>Aa</sub>*. The BamHI-NcoI insert of pJK444 was cloned into pJK446 using the BamHI and NcoI restriction sites to yield plasmid pJK418. Plasmid pJK418 contains the insert which was cloned into the pKOV vector.

melting, annealing and extension times of 30, 60, and 75 s for 30 cycles. The PCR (100  $\mu$ L) contained 60 pmol of each ODN, 40  $\mu$ mol of each dNTP, and 2 units of Vent DNA polymerase. The PCR product was purified by a gel extraction kit and was digested with Sall and SacI. The digested product was purified with a PCR purification kit and cloned into pET23a using the Sall and SacI restriction sites, yielding pJK431. Sequencing

indicated that plasmid pJK431 contained the same silent point mutation in pJK429 that alters the *purK* Arg215 codon CGA to CGG.

*pJK443*. The NotI-Sall insert of plasmid pJK432 was cloned into pJK432 using the NotI and Sall restriction sites to yield plasmid pJK443.

*pJK446*. Quikchange mutagenesis was used to revert the Sall of pJK443 back wild type sequence using ODNs 1399 and 1400. Sequencing of pJK446 indicated the desired mutation had been obtained.

*pJK418*. The inserts of plasmids pJK444 and pJK446 were combined using the BamHI restriction site, and a unique NcoI restriction site found in *purK<sub>Aa</sub>*. The BamHI-NcoI insert of pJK444 was cloned into pJK442 using the BamHI and NcoI restriction sites to yield plasmid pJK418.

*pJK448*. The BamHI-NotI insert of pJK418 was cloned into vector pKOV using the BamHI and NotI restriction sites to yield plasmid pJK448.

#### **d) Construction of plasmid pJK390 for the precise deletion of *purE<sub>Ec</sub>* in the *E.coli* genome**

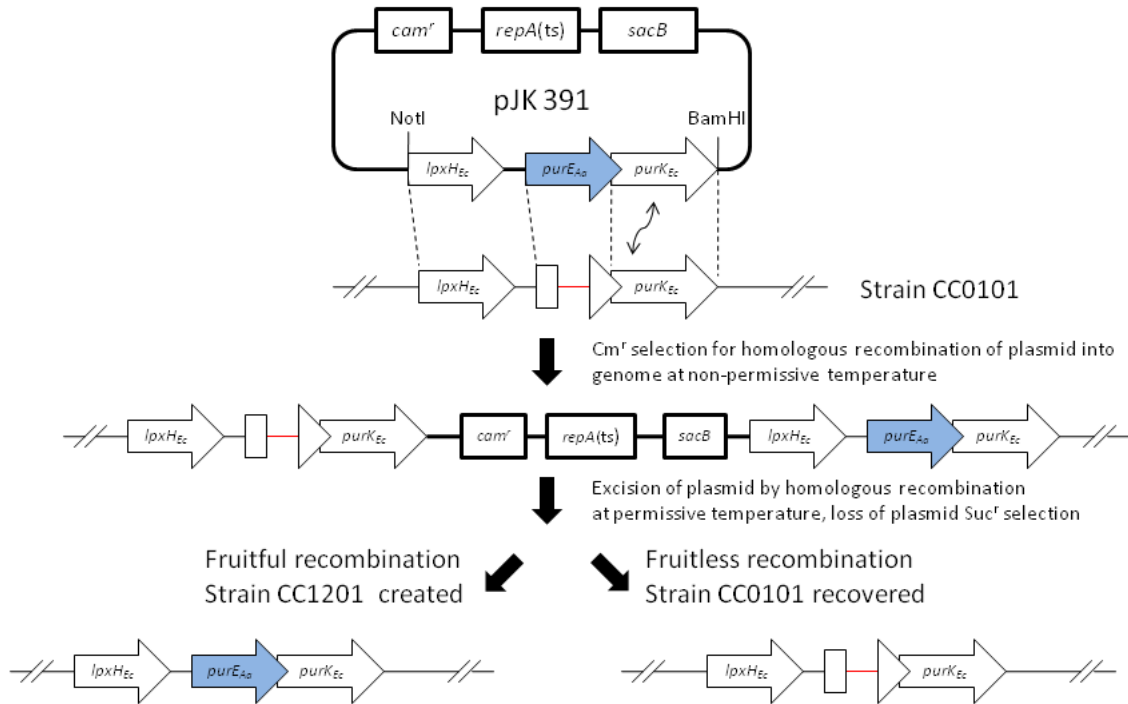
*pJK388*. A 0.9 kb product containing the region upstream of *purE<sub>Ec</sub>* and the first 24 bp of the *purE<sub>Ec</sub>* gene was amplified by PCR from gDNA using ODNs 1269 and 1270, with an annealing temperature of 65 °C and melting, annealing and extension times of 30, 60, and 90 s for 25 cycles. The PCR (50 µL) contained 25 pmol of each ODN, 20 nmol of each dNTP, and 1 unit of Vent DNA polymerase. A 1.2 kb product containing the last 66 bp of *purE<sub>Ec</sub>* and the *purK<sub>Ec</sub>* gene was amplified by PCR from plasmid pNC2 using

ODNs 1301 and 1302, with an annealing temperature of 65 °C and melting, annealing and extension times of 30, 60 and 90 s for 25 cycles. The PCR (50 µL) contained 50 pmol of each ODN, 20 nmol of each dNTP and 1 unit of Vent DNA polymerase. A 2.0 kb cross-over PCR product was obtained using crude products above as template and ODNs 1301 and 1269, with an annealing temperature of 60 °C and melting, annealing and extension times of 30 60 and 150 s for 30 cycles. The PCR (50 µL) contained 25 pmol of each ODN, 40 nmol of each dNTP, 1 unit of Vent DNA polymerase, and the standard Thermopol buffer supplemented with 200 nmol of MgSO<sub>4</sub> to a 6 mM total concentration. The product was purified by PCR purification kit, and the purified product was digested with NotI and BamHI. The digested product was purified by PCR purification kit and was cloned into pET23a using the NotI and BamHI restriction sites to yield plasmid pJK388. Sequencing indicated that the plasmid contained a silent mutation that alters the *purK* Gly31 codon from GGG to GGA.

*pJK390.* The NotI-BamHI insert of pJK388 was cloned into pKOV using the NotI and BamHI restriction sites to yield plasmid pJK390.

*Creation of insertion strains.* Insertion strains were created using the pKOV precise gene replacement vector system developed by the lab of George Church (13, 14). This plasmid allows for the selection of colonies that have inserted the plasmid into the genome by homologous recombination, and allows for later selection of excision and loss of the plasmid (Scheme 4.5). When the plasmid is excised from the genome, it

may either leave the genome unaltered, or replace the genomic DNA with a portion of the plasmid insert.



**Scheme 4.5:** Method for precise gene insertion using pKOV vector system. Illustrated for the creation of *purE<sub>Δa</sub>* insert strain. Unlabeled box and triangle shown for strain 1611 represents remainder of *purE<sub>Ec</sub>* gene left in construction of strain as described in experimental section. Representations are not to scale.

The replacement of genomic DNA with the plasmid insert will be referred to as “fruitful” below. In principle, there is a 50:50 chance of obtaining fruitful and fruitless recombination products, but additional selection factors may bias the ratio of products obtained.

The pKOV vector contains: a chloramphenicol marker, that allows for selection of cells containing the plasmid; a temperature-sensitive origin of replication, that allows for selection of insertion of the plasmid at a non-permissive temperatures and excision

at permissive temperature; and a counter-selectable marker, the *B. subtilis sacB* gene. The *sacB* gene encodes for levansucrase, which is toxic in *E. coli* grown in the presence of sucrose. The counter-selectable marker allows for selection of the loss of the plasmid. The general method for the creation of the precise gene replacement strains is described below, and illustrated in Scheme 4.5 for the creation of the *purE<sub>Aa</sub>* insert strain.

Insertion of plasmid into the genome of *E. coli* was selected for by growth on chloramphenicol plates at a non-permissive temperature for plasmid replication (42 °C). Insertion of plasmid was confirmed by PCR for each strain (individual conditions used for each strain are listed below). Subsequent excision and concomitant loss of plasmid was selected by growth at a permissive temperature for plasmid replication (30 °C) and growth on LB agar plates supplemented with 5% sucrose (w/v). The loss of plasmid was confirmed by sensitivity to chloramphenicol and by PCR.

*Creation of precise purE<sub>Aa</sub> insert strain CC1201.* Several attempts were made to obtain strain CC1201 by transforming W3110 cells with plasmid pJK391. After several failed attempts, the  $\Delta purE_{Ec}$  strain CC0101 was used instead of W3110. This was done owing to the possibility that the fruitless excision of the genome-integrated plasmid may be favored. A fruitless excision of the plasmid would result in strain W3110, and might have a selective advantage because it contains an unaltered copy of *purE*. The  $\Delta purE_{Ec}$  strain CC0101 lacks the *purE<sub>Ec</sub>* gene, and should therefore confer no advantage on fruitless recombination products.

Strain CC1201 was obtained by transforming pJK391 into strain CC0101. After growth at a non-permissive temperature, insertion of the plasmid into the genome was confirmed by whole colony PCR, using the size of products obtained with ODNs pairs 1161 and 1278, and 1277 and 1164, and an annealing temperature of 58 °C and a melting, annealing and extension time of 30, 60 and 120 s for 30 cycles. ODN 1277 and 1278 are sequencing primers for vector pKOV, and ODN 1161 and 1164 are complementary to a region upstream and downstream of *purE<sub>Ec</sub>*. The 1.4 kb size of the product obtained with ODNs 1161 and 1278 indicated that plasmid pJK391 had inserted into the genome. The size of the product indicated that the deleted *purE<sub>Ec</sub>* of strain CC0101 had been amplified and not the *purE<sub>Aa</sub>* of the plasmid. If insertion did not occur or had occurred upstream of ODN 1161, a 1.8 kb product would have been obtained (if the plasmid had inserted upstream of ODN 1161, the reaction with ODNs 1277 and 1164 would have indicated insertion had occurred). The PCR (25 µL) contained 8.25 pmol of each ODN, 10 nmol of each dNTP, and 0.5 units of NEB Taq in Thermopol buffer. Following selection at 30 °C on solid LB/5% sucrose media, colonies were re-passaged on LB/5% sucrose to ensure single colonies were obtained. Whole-colony PCR was used to confirm the excision of plasmid and insertion of *purE<sub>Aa</sub>* into the genome with ODNs 1162 and 1378 and an annealing temperature of 50 °C and melting, annealing and extension times of 30, 60 and 180 s for 30 cycles. The PCR (50 µL) contained 66 pmol of each ODN, 20 nmol of each dNTP, and 1 unit of NEB Taq in Thermopol buffer. Sequencing of the 2.8 kb product indicated the intended strain had been obtained.

*Creation of precise  $purE_{Aa(ATG)}$  insert strain CC1202.* Strain CC1202 was obtained by transforming pJK414 into strain CC0101. After growth at a non permissive temperature, insertion of the plasmid into the genome was confirmed by whole colony PCR, using the size of the product obtained with ODNs pairs 1161 and 1278; and 1164 and 1277 with an annealing temperature of 58 °C and a melting, annealing and extension time of 30, 60 and 120 s for 30 cycles. The 1.3 kb size of the product obtained with ODNs 1277 and 1164 indicated that plasmid pJK414 had inserted into the genome. The size of the product indicated that the deleted  $purE_{Ec}$  of strain CC0101 had been amplified and not the  $purE_{Aa}$  of the plasmid. If insertion had not occurred, or had occurred upstream of ODN 1164, a 1.7 kb product would be obtained (if the plasmid had inserted upstream of ODN 1164, the reaction with ODNs 1278 and 1161 would have indicated insertion had occurred). The PCR (25  $\mu$ L) contained 8.25 pmol of each ODN, 10 nmol of each dNTP, 25 nmol  $MgCl_2$ , and 0.5 units of Taq (Promega) in GoTaq Flexi buffer. Following selection at 30 °C on solid LB/5% sucrose, colonies were re-passaged on LB/5% sucrose to ensure single colonies were obtained. Whole-colony PCR was used to confirm the excision of plasmid and insertion of  $purE_{Aa}$  into the genome with ODNs 1162 and 1378 and an annealing temperature of 50 °C and melting, annealing and extension times of 30, 60 and 180 s for 30 cycles. The PCR (50  $\mu$ L) contained 66 pmol of each ODN, 20 nmol of each dNTP, and 1 unit of Taq in Thermopol buffer. Sequencing of the 2.8 kb product indicated the intended strain had been obtained.

*Creation of precise  $purK_{Aa}$  insert strain CC1203.* Strain CC1203 was obtained by transforming pJK448 into strain CC0103. After growth at a non permissive temperature,

insertion of the plasmid into the genome was confirmed by whole-colony PCR, using the size of the product obtained with ODNs pairs 1161 and 1278; and 1162 and 1277 with an annealing temperature of 58 °C and a melting, annealing and extension time of 30, 60 and 180 s for 30 cycles. The 1.9 kb size of the product obtained with ODNs 1277 and 1162 indicated that plasmid pJK448 had inserted into the genome. The size of the product indicated that the deleted *purK<sub>Ec</sub>* of strain CC0103 had been amplified and not the *purK<sub>Aa</sub>* of the plasmid. If insertion had not occurred, or had occurred upstream of ODN 1162, a 2.9 kb product would be obtained (if the plasmid had inserted upstream of ODN 1162, the reaction with ODNs 1277 and 1161 would have indicated insertion had occurred). The PCR (50 µL) contained 33 pmol of each ODN, 20 nmol of each dNTP, 100 nmol MgCl<sub>2</sub>, and 1 units of Taq (Promega) in GoTaq Flexi buffer. Following selection at 30 °C on solid LB/5% sucrose, colonies were re-passaged on LB/agar/5% sucrose to ensure single colonies were obtained. Whole-colony PCR was used to confirm the excision of plasmid and insertion of *purK<sub>Aa</sub>* into the genome with ODNs 1403 and 1378 and an annealing temperature of 50 °C and melting, annealing and extension times of 30, 60 and 240 s for 30 cycles. The PCR (50 µL) contained 66 pmol of each ODN, 20 nmol of each dNTP, 100 nmol MgCl<sub>2</sub>, and 1 unit of Taq in GoTaq Flexi buffer. Sequencing of the 3.8 kb product indicated the intended strain had been obtained.

*Creation of precise purE<sub>Aa</sub>purK<sub>Aa</sub> insert strain CC1204.* Strain CC1204 was obtained by transforming pJK421 into strain CC0102. After growth at a non permissive temperature, insertion of pJK421 into the genome was confirmed by whole colony PCR, using the size of the product obtained with ODNs pairs 1161 and 1278; and 1162 and



1277 with an annealing temperature of 58 °C and a melting, annealing and extension time of 30, 60 and 180 s for 30 cycles. The 1.4 kb size of the product obtained with ODNs 1278 and 1161 indicated that plasmid pJK421 had inserted into the genome. The size of the product indicated that the deleted *purK<sub>Ec</sub>purK<sub>Ec</sub>* of strain CC0103 had been amplified and not the *purE<sub>Aa</sub>purK<sub>Aa</sub>* of the plasmid. If insertion had not occurred, or had occurred upstream of ODN 1161, a 3.0 kb product would be obtained (if the plasmid had inserted upstream of ODN 1161, the reaction with ODNs 1277 and 1162 would have indicated insertion had occurred). The PCR (50 µL) contained 33 pmol of each ODN, 20 nmol of each dNTP, 100 nmol MgCl<sub>2</sub>, and 1 units of Taq (Promega) in GoTaq Flexi buffer. Following selection at 30°C on solid LB/5% sucrose, colonies were re-passaged on LB/5% sucrose to ensure single colonies were obtained. Whole-colony PCR was used to confirm the excision of plasmid and insertion of *purE<sub>Aa</sub>purK<sub>Aa</sub>* into the genome with ODNs 1403 and 1378 and an annealing temperature of 50 °C and melting, annealing and extension times of 30, 60 and 240 s for 30 cycles. The PCR (50 µL) contained 66 pmol of each ODN, 20 nmol of each dNTP, 100 nmol MgCl<sub>2</sub>, and 1 unit of Taq in GoTaq Flexi buffer. Sequencing of the 3.8 kb product indicated the intended strain had been obtained.

*Creation of precise purE<sub>Aa(ATG)</sub>purK<sub>Aa</sub> insert strain CC1205.* Strain CC1205 was obtained by transforming pJK421 into strain CC1202. Strain CC1202 was used because it already contained *purE<sub>Aa</sub>* with an ATG start codon. After growth at a non-permissive temperature, insertion of pJK421 into the genome was confirmed by whole colony PCR, using ODNs pairs of 1164 and 1277; and 1167 and 1278 with an annealing temperature

of 58 °C and a melting, annealing and extension time of 30, 60 and 120 s for 30 cycles. ODNs 1164 and 1167 are complementary to portions of *purK<sub>EC</sub>*, and would only produce PCR products if the plasmid had integrated into the genome. The 1.6 kb product obtained with ODNs 1278 and 1167 indicated that plasmid pJK421 had inserted into the genome. The PCR (50 µL) contained 33 pmol of each ODN, 20 nmol of each dNTP, 100 nmol MgCl<sub>2</sub>, and 1 units of Taq in GoTaq Flexi buffer. Following selection at 30 °C on solid LB/5% sucrose, colonies were re-passaged on LB/5% sucrose to ensure single colonies were obtained. Whole-colony PCR was used to confirm the excision of plasmid and insertion of *purE<sub>Aa</sub>purK<sub>Aa</sub>* into the genome with ODNs 1403 and 1378 and an annealing temperature of 50 °C and melting, annealing and extension times of 30, 60 and 240 s for 30 cycles. The 50 µL PCRs contained 66 pmol of each ODN, 20 nmol of each dNTP, 100 nmol MgCl<sub>2</sub>, and 1 unit of Taq in GoTaq Flexi buffer. Sequencing of the 3.8 kb product indicated the intended strain had been obtained, with *purE<sub>Aa</sub>* having an ATG start codon.

*Construction of purE<sub>Aa</sub> plasmids for functional complementation studies.* Plasmid pJK173 was made by Christopher Mill of this lab (18). Plasmid pJK173 contains both *purE<sub>Aa</sub>purK<sub>Aa</sub>* and includes a 318 bp 5' upstream region containing *orf1* and a 36 bp 3' untranslated region, cloned into pUC118 via *SphI*. Plasmid pJK173 was used as the template for all plasmids containing *A. aceti* sequence, as listed below.

*Construction of orf-1<sub>Aa</sub>purEH59D<sub>Aa</sub>purK plasmid pJK347.* The codon for His59 in *purE<sub>Aa</sub>* was changed to encode Asp by Quikchange mutagenesis of plasmid pJK173 using ODNs 1017 and 1018. This process yielded plasmid pJK347.

*Construction of orf-1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK plasmid pJK348* – The codon for His59 in *purE<sub>Aa</sub>* was changed to encode Asn by Quikchange mutagenesis of plasmid pJK173 using ODNs 1019 and 1020. This process yielded plasmid pJK348.

*Construction of purE<sub>Aa</sub> single-construct plasmid pJK324.* The *purE<sub>Aa</sub>* region of pJK173 (the *purE* gene, with a 39 bp 5'-UTR and 14 3'-UTR) was amplified by PCR from pJK173 using ODNs 1078 and 1079, with an annealing temperature of 55 °C and melting, annealing and extension times of 30, 60, and 60 s for 25 cycles. The PCR was performed in a 100 µL Hot Start tube with Triton 100-X, and contained 31.3 pmol of each ODN and 2 units of Vent. The product was purified by Qiagen PCR purification kit, and digested with EcoRI and NotI. The digested product was purified by PCR purification kit and cloned into the EcoRI and HindIII restriction sites of pUC118 to yield plasmid pJK324.

*Construction of purEH59D<sub>Aa</sub> plasmid pJK346.* The codon for His59 in *purE<sub>Aa</sub>* was changed to encode Asp by Quikchange mutagenesis of plasmid pJK173, using ODNs 1017 and 1018. This process yielded plasmid pJK346.

*Construction of purEH59N<sub>Aa</sub> plasmid pJK340.* The codon for His59 in *purE<sub>Aa</sub>* was changed to encode Asn by Quikchange mutagenesis of plasmid pJK173 using ODNs 1019 and 1020. This process yielded plasmid pJK340.

*Construction of orf1<sub>Aa</sub>-purE<sub>Aa</sub> plasmid pJK412.* The *purE<sub>Aa</sub>* gene and all of the region upstream of the gene in pJK173 was amplified by PCR from pJK173 using ODNs 617 and 1079, with an annealing temperature of 55 °C and melting, annealing, and extension times of 30, 60 and 60 s for 30 cycles. The 200 µL (for 50 µL reactions) PCR

contained 125 pmol of each ODN, 80 nmol of each dNTP, and 4 units of Vent DNA polymerase. The product was purified using a Qiagen PCR purification kit, and was digested with HindIII and PstI. The digested product was purified by gel extraction and was cloned into the HindIII and PstI restriction sites of pJK173 to yield plasmid pJK412.

*Constructs of  $purE_{Ec}$  for function complementation studies.* Plasmid pNC2 was previously made as reported (19). Plasmid pNC2 contains  $purE_{Ec}purK_{Ec}$  cloned into pET23a via *NdeI* and *HindIII* restriction sites.

*Construction of  $purEH45N_{Ec}purK_{Ec}$  plasmid pJK426.* The codon for His45 in  $purE_{Ec}$  in pNC2 was changed to encode Asn by Quikchange mutagenesis of plasmid pNC2 using ODNs 1388 and 1389. This process yielded plasmid pJK426.

*Construction of  $purE_{Ec}$  single construct plasmid pJK419.* The majority of the  $purK_{Ec}$  gene (bases 193-1066 of the gene) was removed from pNC2 using unique restriction sites *Ascl* and *HincII*. Plasmid pNC2 was digested with restriction enzymes *Ascl* and *HincII* and run on a 1% agarose gel. The 4.3 kb fragment was purified by gel extraction. The 5' overhang left by *Ascl* in the purified fragment was filled in a 150  $\mu$ L reaction, containing 60 nmol of each dNTP and 3 units of Vent DNA polymerase in Thermopol buffer, which was incubated at 72 °C for 15 min. The filled-in fragment was purified by Qiagen PCR purification kit, and ligated to itself in a standard ligation reaction run at room temperature for 3 h, to yield plasmid pJK419. In this plasmid the first 64 amino acids of *EcPurK* are still encoded in an 86 amino acid protein, of which the final 22 amino acid are coded for by the vector.

*Construction of purEH45N<sub>Ec</sub> single construct plasmid pJK435.* The majority of the *purK<sub>Ec</sub>* gene (bases 193-1066 of the gene) was removed from pJK426 using unique restriction sites *Ascl* and *HincII*. Plasmid pJK426 was digested with restriction enzymes *Ascl* and *HincII*, and run on a 1% agarose gel. The 4.3 kb fragment was purified by gel extraction. The 5' overhang left by *Ascl* in the purified fragment was filled in a 100  $\mu$ L reaction, containing 20 nmol of each dNTP and 2 units of Vent DNA polymerase in Thermopol buffer incubated at 72 °C for 15 min. The filled in fragment was purified by Qiagen PCR purification kit, and ligated to itself in a standard ligation reaction run at room temperature for 3 h to yield plasmid pJK435. The first 64 amino acids of *EcPurK* are still encoded in an 86 amino acid protein, of which the final 22 amino acids are coded for by the vector.

*Construction of orf1<sub>Aa</sub>purK<sub>Aa</sub> plasmid pJK415.* A product of the insert of pJK173 with an in-frame deletion removing a large portion of *purE<sub>Aa</sub>* was made by cross-over PCR. ODNs were designed to remove 151 residues of *AaPurE*. A 450 bp product including the multi-cloning site and the first 36 bp of *purE<sub>Aa</sub>* was amplified from pJK173 by PCR using ODNs 617 and 1371, with a 58 °C annealing temperature and melting, annealing and extension times of 30 s each for 30 cycles. The PCR (50  $\mu$ L) contained 16.5 pmol of each ODN, 20 nmol of each dNTP, and 1 unit of Vent DNA polymerase. The product was purified using a PCR purification kit and was used as template (4 ng) in the subsequent cross-over PCR. A 1.3 kb product containing the last 96 bp of the *purE<sub>Aa</sub>* gene, and all of *purK<sub>Aa</sub>* and the *HindIII* site derived from pUC118 was amplified from pJK173 by PCR using ODNs 601 and 1370, with a 62 °C annealing temperature, and

melting, annealing and extension times of 30, 30 and 80 s for 30 cycles. The PCR (50  $\mu$ L) contained 16.5 pmol of each ODN, 20 nmol of each dNTP, and 1 unit of Vent DNA polymerase. The product was purified using a PCR purification kit and was used as template (4 ng) in the subsequent cross-over PCR.

A 1.7 kb cross-over PCR product was obtained using the above products as template and ODNs 601 and 607, with an annealing temperature of 58 °C and melting, annealing and extension times of 30, 30 and 110 s for 30 cycles. The PCR (100  $\mu$ L) contained 33 pmol of each ODN, 40 nmol of each dNTP, and 1 unit of Vent DNA polymerase. The product was purified by gel extraction, and digested with HindIII and PstI. Following digestion, the product was purified by gel extraction and was ligated into pJK173. Sequencing indicated the plasmid contained a silent point mutation in *purK<sub>Aa</sub>* that has been seen previously (Arg215: CGA→CGG).

*Construction of purK<sub>Ec</sub> plasmid pJK427.* A 1.2 kb product containing *purK<sub>Ec</sub>* and a truncated *purE<sub>Ec</sub>* gene was amplified from pJK388, using ODNs 1390 and 1391 with an annealing temperature of 58 °C and melting, annealing and extension times of 30, 60, and 75 s for 30 cycles. The product was gel purified and digested with SphI and NdeI. The digested product was purified using a Qiagen PCR purification kit and was ligated into SphI and NdeI restriction sites of pNC2.

*Construction of orf1<sub>Aa</sub> plasmid.* A large portion of *purE<sub>Aa</sub>*, and all of the *purK<sub>Aa</sub>* gene was removed from pJK173 by digestion with AgeI and HindIII. The 3.7 kb fragment was purified by gel extraction. The overhangs left by digestion were filled in a 55  $\mu$ L

reaction, containing 20 nmol of each dNTP, 1 unit of Vent DNA polymerase in Thermopol buffer, incubated at 72 °C for 10 min. The filled-in product was purified by PCR purification kit and ligated to itself, in a standard ligation reaction run at room temperature for 3 h. The plasmid encodes for *orf1<sub>Aa</sub>* and for the first 64 amino acids of AaPurE fused to the  $\alpha$ -LacZ fragment of pUC118.

*Description of media.* Two types of media lacking purines were used in liquid and solid form (containing 1.5% agar). For all experiments, positive controls were run containing exogenous purines. The addition of purines in the controls allows for the growth of auxotrophic strains on minimal media. Minimal medium (A) was the media described by Schrimsher et al. (20), further supplemented with 10  $\mu$ M thiamine and 100  $\mu$ g/mL carbenicillin. For all experiments involving medium A, the positive controls contained 15  $\mu$ g/mL hypoxanthine added to the media prior to autoclaving. Medium (B) was composed of M9 minimal salts (21) with 0.4% glucose as a carbon source, supplemented with 10  $\mu$ M thiamine and 50  $\mu$ g/mL carbenicillin. For all experiments involving medium B, the positive controls contained 100  $\mu$ g/mL adenine added from an autoclaved 5 mg/mL stock solution of the chloride salt of adenine.

*Initial strain PCO135 functional complementation experiments – Plasmids* pJK173, pJK324, pJK346, pJK347 and pUC118 (negative control) were transformed into strain PCO135 cells and their ability to complement purine auxotrophy on solid media A examined. Cells were grown for 12 h at 37 °C on LB/Amp plates and a single colony was used to inoculate a 5 mL LB/Amp culture which was grown at 37 °C for 10 h. Using  $A_{600}$

measurements, the culture was diluted in LB to a final  $A_{600} = 0.0004$ . A portion of this diluted culture was then spread on medium A. For all plasmids except pJK347, a series of plates were also grown in a candle jar (22). To create the candle jar, plates were placed in a large glass desiccator. A lit candle was placed in the desiccator, and the desiccator was sealed. The candle was allowed to burn until it extinguished, and the desiccator was placed at 37 °C. Plasmids which complemented the purine auxotrophy on solid medium A were also assessed for their ability to complement the purine auxotrophy in liquid medium A.

*Functional complementation of PCO135.* Plasmids pJK173, pJK324, pJK412 and pUC118 (negative control) were transformed into strain PCO135 cells and their ability to grow on solid medium A was assessed. Cells were grown at 37 °C on LB/Amp plates for 12 h, and a single colony was used to inoculate a 5 mL LB/Amp starter culture, which was grown to mid-log phase. The  $A_{600}$  of the cultures were measured and the value used to make dilutions in LB to a final  $A_{600} = 0.00005$ . A portion of the diluted culture was spread on solid medium A and grown at 37 °C. A duplicate set of plates was also grown in a candle jar.

*Solid medium functional complementation studies of strain CC0102 and PCO135 involving  $purE_{Ec}$  and  $purEH45N_{Ec}$  constructs.* Strain CC0102 and PCO135 cells were transformed with plasmids pNC2 ( $purE_{Ec}purK_{Ec}$ ), pJK419 ( $purE_{Ec}$ ), pJK426 ( $purEH45N_{Ec}purK_{Ec}$ ); and pJK435 ( $purEH45N_{Ec}$ ), or pUC118 as a negative control. Plasmids were assessed for their ability to restore prototrophy to those strains. LB/Amp



cultures (5 mL) were inoculated from single colonies and grown for 12 h at 37 °C. Cells were harvested by centrifugation (4000 g for 10 min), and resuspended in 5 mL medium B. Cells were harvested by centrifugation again, and resuspended in a final 5 mL of medium B. The  $A_{600}$  of the cultures were measured and the cultures were diluted in minimal media B to a final  $A_{600} = 0.00005$ . A portion of the diluted culture was spread on solid medium A and grown at 37 °C. For strain CC0102, a portion of the diluted culture was also spread on solid medium B.

*Functional complementation studies of CC0101 in liquid medium B.* Plasmids pJK173, pJK347, pJK348, pJK324, pJK412 and pNC2 were assessed for their ability to restore prototrophy to strain CC0101. The negative control was CC0101/pUC118. Cells were transformed with plasmids listed and placed on LB/Amp plates and incubated at 37 °C for 16 h. Satellite colonies were noted on the plates. A single large colony from each plate was passaged on LB/Amp plates and incubated at 37 °C overnight. LB/Amp cultures (5 mL) were inoculated from single colonies and grown for 12 h at 37 °C. Cells were harvested by centrifugation (4000 g for 15 min), and were resuspended in 5 mL medium B. Cells were harvested by centrifugation again, and resuspended in a final 5 mL of medium B. The  $A_{600}$  of the cultures was measured, and 50 mL medium B cultures were inoculated to an  $A_{600} = 0.01$ . The adenine that was added to positive controls in these experiments was sterile filtered. Cultures were grown at 37 °C with shaking at 200 rpm in a New Brunswick Scientific Innova 4430 shaking incubator. The growth of cultures was monitored by measuring  $A_{600}$  over time. Cultures were observed for 186 h.

*Study  $purE_{EC}$  and  $purE_{EC}purK_{EC}$  constructs on strains CC0101 ( $\Delta purE_{EC}$ ), CC1201 and CC1202, on solid medium B.* Strain CC1201, CC1201 and CC0101 cells were transformed with plasmids pJK419 or pNC2. Plasmids were assessed for their ability to restore growth on solid minimal medium A and solid minimal medium B. Cells were placed on LB/Amp plates which were supplemented with 100  $\mu\text{g}/\text{mL}$  adenine, 0.1 % glucose and 10  $\mu\text{M}$  thiamine and grown at 37 °C for 12 h. LB/Amp starter cultures (5 mL) supplemented with 100  $\mu\text{g}/\text{mL}$  adenine, 0.1 % glucose and 10  $\mu\text{M}$  thiamine, were inoculated from a single colony, and grown for 19 h at 37 °C. The  $A_{600}$  of the cultures was measured, and dilutions made in LB to a final  $A_{600} = 0.002$ . A portion of the diluted culture was spread on solid medium A or solid medium B and grown at 37 °C. A duplicate set of plates for CC0101/pJK419 on solid medium A was also grown in a candle jar.

*Functional complementation of strain CC0101 with single  $purE_{EC}$  construct in liquid medium B.* Strain CC0101 cells were transformed with plasmids pJK419 ( $purE_{EC}$ ). The negative control was CC0101/pUC118. Plasmid pJK419 was assessed for its ability to restore prototrophy to strain CC0101 in liquid media. Transformed cells were placed on LB/Amp plates and incubated at 37 °C for 11 h. LB/Amp cultures (5 mL) were inoculated from single colonies and grown to mid log phase at 37 °C. Cells were harvested by centrifugation (4000 g for 10 min), and were resuspended in 5 mL medium B. Cells were harvested by centrifugation again, and resuspended in a final 5 mL of media B. The  $A_{600}$  of the cultures was measured, and 50 mL medium B cultures were inoculated to an  $A_{600} = 0.005$ . Cultures were grown at 37 °C with shaking at 200 rpm in a

New Brunswick Scientific Innova 4430 shaking incubator. The growth of cultures was monitored by measuring  $A_{600}$  over time.

*Functional complementation studies of CC0102 intended to assess the in vivo activity of  $purE_{Aa}$ ,  $purEH59D_{Aa}$ ,  $purEH59N_{Aa}$ , and  $purEH45N_{Ec}$ .* Plasmids pJK173, pJK347, pJK426, pNC2 and pJK348 were assessed for their ability to restore prototrophy to strain CC0102. The negative control was CC0102/pUC118. Cells were transformed with plasmids listed and a single colony was used to inoculate a 5 mL LB/Amp starter culture which was grown to mid-log phase. Cells were harvested by centrifugation at 4000 g for 10 min. Cell pellets were resuspended in 5 mL of medium B, harvested as before, and resuspended in 5 mL of medium B. The  $A_{600}$  of resuspended cells was measured, and 50 mL medium B cultures were inoculated to an  $A_{600} = 0.005$ . Flasks were incubated at 37 °C with shaking at 200 rpm in a New Brunswick Scientific Innova 4430 shaking incubator. The growth of cultures was monitored by measuring  $A_{600}$  over time.

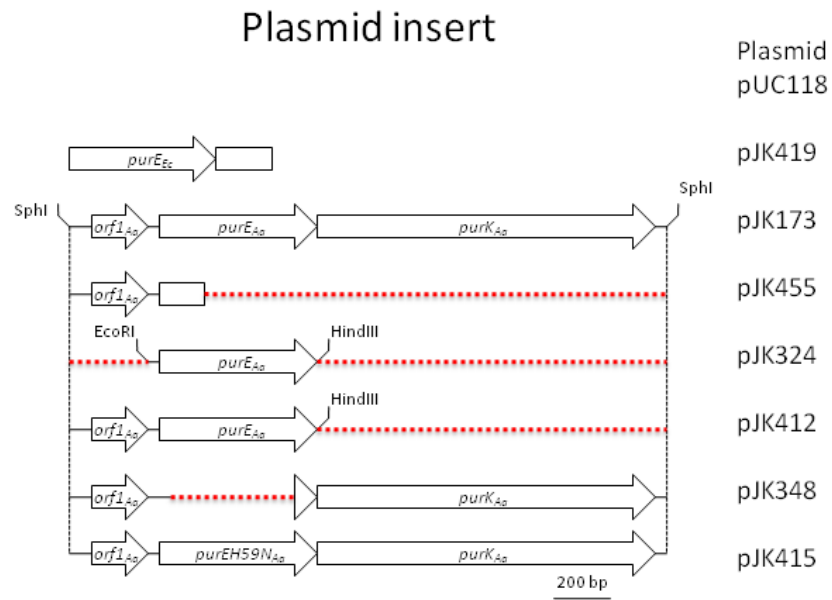
*Growth of strains CC0102 ( $\Delta purE_{Ec} \Delta purK_{Ec}$ ), CC0103( $\Delta purK_{Ec}$ ), CC1203 ( $purK_{Ec} \leftrightarrow purK_{Aa}$ ), CC1204 ( $purE_{Ec} purK_{Ec} \leftrightarrow purE_{Aa} purK_{Aa}$ ) and CC1205 ( $purE_{Ec} purK_{Ec} \leftrightarrow purE_{Aa(ATG)} purK_{Aa}$ ) on solid medium B.* All strains were transformed with pUC118 and their ability to grow on solid medium B was assessed. Strain CC0102 was also transformed with pJK173 as a positive control. Following transformation cells were spread on LB/Amp plates, and grown for 11 h at 37 °C. LB/Amp starter cultures (5 mL) were inoculated from a single colony and grown at 37 °C to mid log phase. Cells were harvested by centrifugation (4000 g, 10 min) and resuspended in 5 mL of medium B.

Cells were harvested again by centrifugation (4000g, 10 min), and resuspended in 5 mL of medium B. The  $A_{600}$  of the cultures were measured, and the cultures were diluted in minimal medium B to a final  $A_{600} = 0.0001$ . A portion of the dilution was spread on solid media B and the plates grown at 37 °C.

*Initial functional complementation studies for CC1203, CC1204, and CC1205 in liquid medium B.* The ability of the *purK<sub>Aα</sub>*, *purE<sub>Aα</sub>purK<sub>Aα</sub>*, and *purE<sub>Aα(ATG)</sub>purK* insert strains to grow in medium B was assessed in reference to strain W3110. All four strains were transformed with pUC118. Additionally the effects of pJK173 or pJK455 on the growth of the *purE<sub>Aα</sub>purK<sub>Aα</sub>*, and *purE<sub>Aα(ATG)</sub>purK* insert strains were assessed. Strain CC0102/pUC118 was used as a negative control. Cells were transformed with the listed plasmids. A single colony was picked from a plate grown at 37 °C for 12 h and used to inoculate a 5 mL LB/Amp starter culture, which was grown to mid-log phase. Cells were harvested by centrifugation at 4000 g for 10 min. Cell pellets were resuspended in 5 mL of medium B, harvested as before, and resuspended in 5 mL of medium B. The  $A_{600}$  of resuspended cells was measured, and 50 mL medium B cultures were inoculated to an  $A_{600} = 0.005$ . Flasks were incubated at 37 °C with shaking at 200 rpm in a New Brunswick Scientific Innova 4430 shaking incubator. The growth of cultures was monitored by measuring  $A_{600}$  over time, and measurements taken only after growth was observed. The negative control was observed for 1000 h.

*Growth experiments involving strains CC1202 and CC1205.* In separate experiments for each strain, plasmid pJK348 and plasmids containing portions of pJK173 (pJK324, pJK412, pJK415, and pJK455) were assessed for their effects on the growth of

strains in medium B (Scheme 4.6). Strains were also transformed with plasmid pJK419 or pUC118. For each growth experiment, a strain and plasmid combination that was not expected to grow in medium B was used. For strain CC1202, CC0101/pUC118 was used as a negative control. For growth experiments with strain CC1205, CC0102/pJK348 was used as a negative control. Single colonies, picked from plates incubated for 12 h at



**Scheme 4.6:** Diagram showing inserts of plasmids. Plasmids were used in growth experiments involving strains CC1202 and CC1205. Regions deleted from plasmids are shown as dashed red line.

37 °C, were used to inoculate 5 mL LB/Amp starter cultures. Starter cultures were grown at 37 °C to mid-log phase. Cells were harvested by centrifugation at 4000 g for 10 min. Cell pellets were resuspended in 5 mL of medium B, harvested as before, and resuspended in 5 mL of medium B. The  $A_{600}$  of resuspended cells was measured, and 50 mL medium B cultures were inoculated to an  $A_{600} = 0.005$ . Flasks were incubated at

37°C with shaking at 200 rpm in a New Brunswick Scientific Innova 4430 shaking incubator. The growth of cultures was monitored by measuring  $A_{600}$  over time, and measurements taken only after growth was observed. Negative controls were observed for 1000 h.

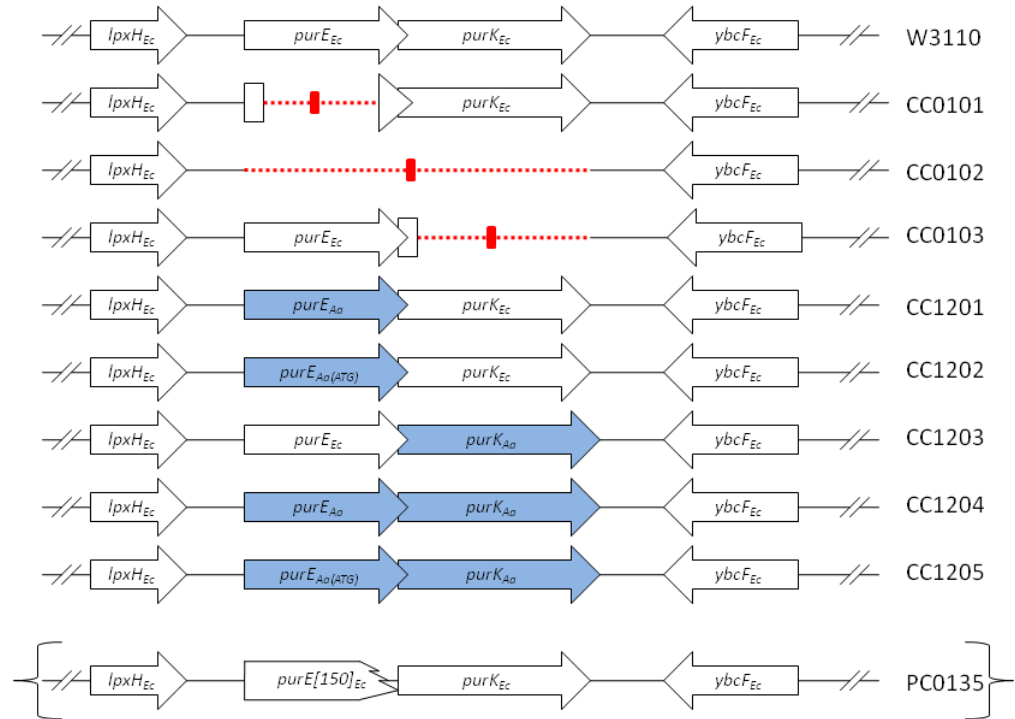
### 4.3 Results

*Construction of plasmids.* Standard molecular biology techniques were used in the construction of almost all of the plasmids used in the functional complementation studies. This process yielded plasmids that either contained the expected sequence, or silent point mutations. To construct the plasmids for manipulation of genomic DNA, two different approaches were taken to combine DNA products from different sources (*E. coli* and *A. aceti*). The first approach used the technique of cross-over PCR. The second approach relied on assembling PCR products amplified from different sources by the introduction of restriction sites and cloning the products into pET23a. After the pieces were assembled, mutagenesis was then used to remove the restriction sites and revert the sequence to “wild type.” Both of these methods provided final plasmids with the expected sequence, or contained a mutation that did not affect the strains. The sequencing of strains indicated that no mutations present in a plasmid had been incorporated into the flanking regions of genomic DNA (i.e., the mutation present in the insert of pJK389 in the *lphX<sub>EC</sub>* Ala7 codon was not incorporated into the genome).

The inserts of plasmids pJK388 and pJK389 were constructed using cross-over PCR. Attempts were made to construct the inserts of plasmids pJK418 and pJK420 by a

similar method but were unsuccessful. While in theory it would seem that cross-over PCR would provide the easiest route, in practice it can be difficult. The increased size (~2.5 kb vs ~1.5 kb) of the cross-over products in the construction of *purE<sub>Aa</sub>purK<sub>Aa</sub>* and *purK<sub>Aa</sub>* insert plasmids likely contributed to the failure of cross-over PCR in the construction of these plasmids. Cross-over PCR was very effective for rapidly creating plasmid pJK388 and plasmid pJK415, each of which was built from comparatively smaller pieces of DNA.

*Construction of new strains.* A schematic representation of the *purEpurK* region for all strains is shown (Scheme 4.7). Initially plasmids were designed to create a precise gene replacement strain *purE<sub>Ec</sub><>purE<sub>Aa</sub>* and an in-frame deletion strain of *purE<sub>Ec</sub>* using the pKOV system (13, 14). These plasmids were designed to install the *purE<sub>Aa</sub>* gene in the precise location of *purE<sub>Ec</sub>* and to create a precise in-frame deletion of *purE<sub>Ec</sub>* without a stop codon, in such a way as to not disturb translational coupling or regulatory control of the targeted genes. After several failed attempts to use these plasmids to introduce changes into the genome of strain W3110, an alternate route was designed to obtain strains for study.



**Scheme 4.7:** Schematic representation of the *purEpurK* region of strains used in this study. The representations of genes are not to scale. Colored genes represent genes from *A. aceti*. The dashed red line in strains CC0101, CC0102, and CC0103 represents deleted region of genome and the small red box represents the exogenous 82 bp scar left by the method used. Strain PC0135 is shown separately and in brackets because it is not isogenic outside the region of the genome shown.

The method of Datsenko and Wanner (15) provided a rapid means to obtain deletion strains, CC0101 ( $\Delta purE_{Ec}$ ), CC0102 ( $\Delta purE_{Ec} \Delta purK_{Ec}$ ), and CC0103 ( $\Delta purK_{Ec}$ ). With this method, the desired strains were usually obtained on the first attempt, with the exception of strain CC0102, which was obtained on the second attempt.

After the  $\Delta purE_{Ec}$  deletion strain CC0101 was created, it was used to obtain the precise gene replacement strain CC1201 ( $purE_{Ec} \leftrightarrow purE_{Ac}$ ). This was based on the assumption that attempts to obtain the precise gene replacement strain CC1201 ( $purE_{Ec} \leftrightarrow purE_{Ac}$ ) using W3110 cells may have failed because a fruitless excision might confer a selective advantage. This is because a fruitless excision would contain an



unaltered copy of the *purE<sub>EC</sub>-purK<sub>EC</sub>* operon. The  $\Delta purE_{EC}$  strain CC0101 lacks the *purE<sub>EC</sub>* gene, and should therefore confer no advantage on fruitless recombination products. Using the deletion strain CC0101, we were able to obtain the precise gene replacement strain CC1201.

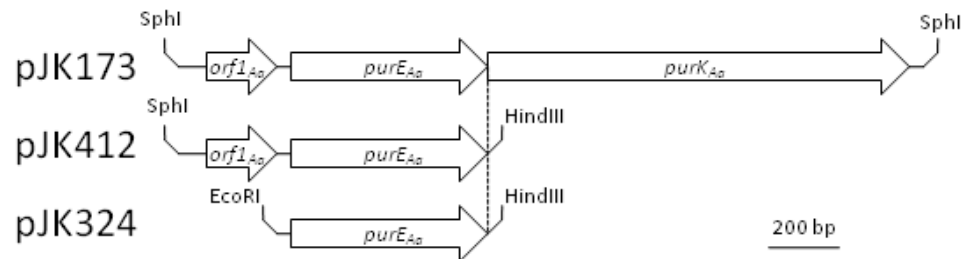
A similar method was used to obtain strains CC1202 (*purE<sub>EC</sub><>purE<sub>Aa(ATG)</sub>*), CC1203 (*purK<sub>EC</sub><>purK<sub>Aa</sub>*), and CC1204 (*purE<sub>EC</sub>purK<sub>EC</sub><>purE<sub>Aa</sub>purK<sub>Aa</sub>*), using the deletion strains CC0101 ( $\Delta purE_{EC}$ ), CC0103 ( $\Delta purK_{EC}$ ), and CC0102 ( $\Delta purE_{EC}\Delta purK_{EC}$ ) respectively. To obtain strain CC1205 (*purE<sub>EC</sub>purK<sub>EC</sub><>purE<sub>Aa(ATG)</sub>purK<sub>Aa</sub>*), the precise gene replacement strain CC1202 was used because it already contained the *purE<sub>Aa</sub>* gene with an ATG start codon. The desired strain was obtained using this method, suggesting that the absence of *purE<sub>EC</sub>* in the altered *purE<sub>Aa(ATG)</sub>purK<sub>EC</sub>* operon did not convey the same benefit to fruitless recombination as the unaltered *purE<sub>EC</sub>* in W3110 apparently did.

*Functional complementation of strain PC0135 by A. aceti genes.* *AaPurE* and *AaPurE His59* mutants were initially tested for their ability to restore prototrophy to strain PC0135 cells on solid and liquid minimal media. This was done to correlate in vivo activities with in vitro activities (or lack thereof) previously observed (Chapter 3). Based on the detected in vitro activities, it was expected that *AaPurE* would be active, that the *AaPurE-H59D* mutant would be less active, and that the *AaPurE-H59N* mutant would be inactive. More definitive results were obtained with strain CC0102 ( $\Delta purE_{EC}\Delta purK_{EC}$ ) and are reported in a following section.

In initial experiments, plasmids were based on pJK173 or pJK324, both of which are derived from the high copy number pUC118 vector (Scheme 4.8). Plasmid pJK173

was initially identified by functional complementation of PC0135 (18). It contains all of the *purE<sub>Aa</sub>purK<sub>Aa</sub>* coding region and surrounding areas including *orf1<sub>Aa</sub>*. Plasmid pJK324 contains *purE<sub>Aa</sub>* and a 39 nt portion of the 5'-untranslated region, which should contain upstream regulatory elements needed for the transcription and translation of *purE<sub>Aa</sub>*. Cells transformed with pUC118 were used as a negative control. The results of the initial findings are summarized below in Table 4.5.

For later experiments, plasmid pJK412 was created. Plasmid pJK412 contains *purE<sub>Aa</sub>* and the 318 bp 5' upstream region present in pJK173 that contains *orf1<sub>Aa</sub>*. The ability of pJK412 to complement the auxotrophy was assessed in comparison to pJK173, pJK324 and pUC118.



**Scheme 4.8:** Illustration of the inserts of plasmids used in the functional complementation studies of the *purE* strain PC0135. Plasmids pJK173 and pJK324 were changed by mutagenesis to obtain the *AaPurE*-H59D and *AaPurE*-H59N mutants.

As expected, wild type *AaPurE* and the *AaPurEH59D* mutant were able to complement the purine auxotrophy in strain PC0135; however, cells transformed with a plasmid containing the apparently inactive mutant *AaPurEH59N* also complemented the purine auxotrophy. A similar result has also been seen in *EcPurE* with the analogous H45N mutant (23). The ability of these mutants, which have no detectable in vitro activity, to complement the PC0135 auxotrophy is proposed to be due to the formation

of heterooctamers that mix the truncated *EcPurE* encoded by PC0135 with the full-length mutant (23).

**Table 4.5:** Functional complementation studies of *purE* deficient strain *E. coli* PC0135

Plasmid ( <i>A. acetii</i> genes)	Medium A (solid)+ hypoxanthine <sup>a</sup>	Medium A (solid)	Medium A (solid) + high CO <sub>2</sub> <sup>b</sup>	Medium A (liq)
pJK173 ( <i>orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub></i> ) <sup>ef</sup>	++++ <sup>c</sup>	++++	+++	++++
pJK347 ( <i>orf1<sub>Aa</sub>purEH59D<sub>Aa</sub>purK<sub>Aa</sub></i> ) <sup>e</sup>	+++	+++ <sup>d</sup>	Nd	++ <sup>d</sup>
pJK348 ( <i>orf1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK<sub>Aa</sub></i> ) <sup>e</sup>	+++	++ <sup>d</sup>	++	-
pJK324 ( <i>purE<sub>Aa</sub></i> ) <sup>ef</sup>	+++	-	++	-
pJK346 ( <i>purEH59D<sub>Aa</sub></i> ) <sup>e</sup>	+++	-	-	-
pJK340 ( <i>purE<sub>Aa</sub></i> ) <sup>e</sup>	+++	-	-	-
pJK412 ( <i>orf1<sub>Aa</sub>purE<sub>Aa</sub></i> ) <sup>f</sup>	+++	+	++	Nd
pUC118 <sup>ef</sup>	+++	-	-	Nd

<sup>a</sup>Minimal medium A as defined in Experimental Section. Hypoxanthine was added at 15 µg/mL. <sup>b</sup>Plates were placed in a candle jar as described in the Experimental Section. <sup>c</sup>Key to symbols: +++++, large colonies/saturated culture at 24 h; +++ small colonies at 24 h; ++, large colonies/light turbidity at 72 h; +, small colonies at 72 h; -, no growth observed at 144 h; Nd, not determined. <sup>d</sup>The indicated plasmids was recovered from this culture, and its DNA sequence was obtained to confirm that the *purE* region had not been modified. <sup>e</sup>Initial experiments. <sup>f</sup>Later experiments.

Although the *AaPurE* and the *AaPurEH59D* mutant were able to restore prototrophy, the ability to do so was only observed when they were co-expressed with either *orf1<sub>Aa</sub>* or *orf1<sub>Aa</sub>* and *purK<sub>Aa</sub>*. This suggests a necessary role for the product of *orf1<sub>Aa</sub>* in the complementation of the purine auxotrophy.

All of the plasmids except pJK347 were also assessed for their ability to functionally complement strain PC0135 in the CO<sub>2</sub> rich environment of a candle jar. Growth in the high pCO<sub>2</sub> environment of the candle jar allowed for functional complementation of strain PC0135 by plasmid pJK324, which was not observed in a normal atmosphere. The phenotype exhibited by PC0135/pJK324 is reminiscent or very

similar to that demonstrated by *purK* auxotrophs (12, 22), which is curious because PC0135 contains an intact, endogenous *purK<sub>Ec</sub>* gene and an exogenous *purE<sub>Aa</sub>* gene present on a high copy number plasmid.

The growth of PC0135/pJK324 under high levels of CO<sub>2</sub> indicated that *purE<sub>Aa</sub>* from pJK324 was expressed in an active form. Given that PC0135/pJK324 gave the same phenotype as a *purK* auxotroph, this suggested that *AaPurE* required *AaPurK* to function properly. Such a requirement suggests that channeling of the substrate *N*<sup>5</sup>-CAIR may occur. Channeling is the direct transfer of a metabolite between successive enzymes in a metabolic pathway. While metabolite channeling is consistent with this result, it is not the only possible explanation. It is inadvisable to interpret these results to imply channeling (23), without performing further experiments to eliminate other possibilities.

*What channeling would require.* If channeling were to occur between PurE and PurK, those enzymes must form a complex. This complex would not necessarily be static, and could form a transient ternary complex with the shared metabolite *N*<sup>5</sup>-CAIR. In either case, ‘matched’ enzymes from the same organism would have compatible surface features. The simplest model is that a strain would grow on minimal medium if it contains a PurE and a PurK from the same source (i.e., *purE<sub>Ec</sub>purK<sub>Ec</sub>* or *purE<sub>Aa</sub>purK<sub>Aa</sub>*). For strains containing a PurE and a PurK from different sources (i.e., *purE<sub>Aa</sub>purK<sub>Ec</sub>* or *purE<sub>Ec</sub>purK<sub>Aa</sub>*), a *pur* phenotype is expected on minimal medium. The formation of a *AaPurK-AaPurE* channeling complex might also require *AaOrf1*, as will be discussed later. Because *E. coli* lacks an *orf1<sub>Aa</sub>* homologue, the interaction between *EcPurK* and

*EcPurE* might be different than that between *AaPurK* and *AaPurE*, leading to an inability to recognize a mismatched partner.

*Characterization of E. coli strain PC0135.* Although strain PC0135 has been used in elucidating the purine biosynthetic pathway and mapping the genes involved (24), no published work has reported the molecular defect. Primers were thus designed to amplify the *purE<sub>Ec</sub>purK<sub>Ec</sub>* genes and surrounding areas. The resulting product was sequenced in both directions and only a single nucleotide difference from wild type was identified: a point mutation in *purE<sub>Ec</sub>* that changes wt-*EcPurE* Trp151 to a stop codon (TGG → TGA). This mutation would cause the final helix of *E. coli* PurE to be deleted, which might interfere with octamer assembly, and could explain the auxotrophic phenotype.

Given that most of *purE<sub>Ec</sub>* is still encoded by strain PC0135, including all active site residues, the strain may have a leaky phenotype due to expression of the truncated *EcPurE*. If the *AaPurEH59N* or *EcPurEH45N* mutants stabilize the truncated *EcPurE*, this leaky phenotype explains functional complementation of the purine auxotrophy with otherwise 'dead' mutants (Chapter 3)(23). The unusual observation that the apparently inactive *AaPurEH59N* mutant and analogous *EcPurEH45N* mutant restore prototrophy to strain PC0135 (23) has led to speculation that the truncated PurE encoded by PC0135 may be stabilized by the inactive PurEs thereby allowing for complementation. The possible leaky phenotype necessitated the construction of better auxotrophic strains.

*Deletion strains CC0101, CC0102, and CC0103.* Of the three deletion strains, strain CC0102 ( $\Delta purE_{Ec}\Delta purK_{Ec}$ ) was selected for most studies because it contained a complete deletion of the genomic genes, and interpretation of results would not require a consideration of the effects of the genomic *purK<sub>Ec</sub>*.

*Functional complementation of strains PC0135 and CC0102 by *purE<sub>Ec</sub>* genes.* The leaky PC0135 phenotype was examined by comparing the effects of *purE<sub>Ec</sub>purK<sub>Ec</sub>* and *purE<sub>Ec</sub>* constructs, along with those plasmids with the *EcPurEH45N* mutation, on the growth of strain PC0135 and deletion strain CC0102 ( $\Delta purE_{Ec}\Delta purK_{Ec}$ ) on solid minimal medium A (Table 4.6). For strain CC0102, experiments were also performed on solid minimal medium B (Table 4.6).

<b>Table 4.6:</b> Functional complementation studies of <i>purE<sub>Ec</sub></i> deficient strain PC0135 and <i>purE<sub>Ec</sub>purK<sub>Ec</sub></i> deficient strain CC0102 by <i>purE<sub>Ec</sub></i> constructs on solid media. <sup>a</sup>			
Plasmid ( <i>E. coli</i> genes)	Medium A (solid) <sup>b</sup> PC0135	Medium A (solid) <sup>b</sup> CC0102	Medium B (solid) <sup>b</sup> CC0102
pUC118	- <sup>c</sup>	-	-
pNC2 ( <i>purE<sub>Ec</sub>purK<sub>Ec</sub></i> )	+++	+++	+++
pJK426 ( <i>purEH45N<sub>Ec</sub>purK<sub>Ec</sub></i> )	++	-	-
pJK419 ( <i>purE<sub>Ec</sub></i> )	++	-	-
pJK435 ( <i>purEH45N<sub>Ec</sub></i> )	+	-	-
<sup>a</sup> A set of plates supplemented with exogenous purines in the form of 15 µg/mL hypoxanthine for medium A or 100 µg/mL adenine for medium B. These controls gave (+++) growth except for strain CC0102/pUC 118 which gave (++) growth. <sup>b</sup> minimal media A and B as described in Experimental section. <sup>c</sup> key to symbols: +++, large colonies at 36 h; ++, small colonies at 36 h; +, small colonies at 48 h; -, tiny colonies which do not increase in size with time up to 144 h.			

As expected, plasmids pNC2 and pJK419 were able to functionally complement the auxotrophy in strain PC0135. Plasmid pNC2 was able to complement the auxotrophy of strain CC0102. Because strain CC0102 does not contain a genomic copy


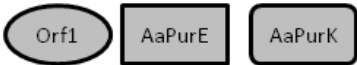

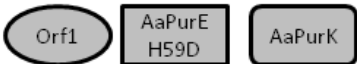

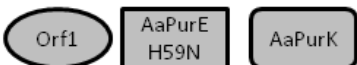





of *purK<sub>EC</sub>*, plasmid pJK419 was not expected to, and did not, restore prototrophy to CC0102.

Plasmids that contained the *EcPurEH45N* mutant (pJK426 and pJK435) restored growth to strain PC0135. This was expected from previously reported studies (23), and from the results we obtained for *AaPurEH59N* above. If the *EcPurEH45N* mutant is active in vivo, plasmid pJK426 should restore prototrophy to strain CC0102. The lack of growth of strain CC0102/pJK426 (*purEH45N<sub>Ec</sub>purK<sub>Ec</sub>*) on solid minimal medium A and on solid minimal medium B indicates that *EcPurEH45N* lacks the activity required to complement the auxotrophy when all of *purE<sub>EC</sub>* is deleted. This result indicates that the inactive mutant *EcPurEH45N* only restores prototrophy in strain PC0135 because it is able to combine with the truncated PurE expressed by this strain.

In strain PC0135, growth is faster when either wild type or mutant *EcPurE* is co-expressed with *EcPurK*. This result suggests that the molecular lesion, which causes the *purE* auxotrophy in strain PC0135, either has a polar effect on the expression of *purK<sub>EC</sub>* or the selective overexpression of *purE<sub>EC</sub>* inhibits growth.

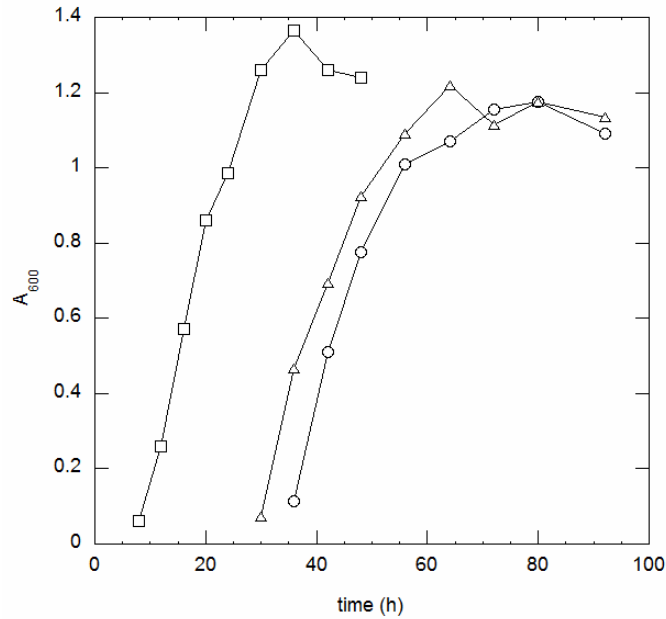
*Functional complementation studies of strain CC0102 ( $\Delta purE_{EC}\Delta purK_{EC}$ ) in liquid medium.* Plasmids bearing *orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>* and *purE<sub>Ec</sub>purK<sub>Ec</sub>*, or their respective PurE H59N (*E. coli*: H45N) mutants, were assessed for their ability to restore prototrophy to strain CC0102 in liquid medium B. Plasmid pUC118 was used as a negative control. Positive controls contained exogenous purines in the form of 100  $\mu$ g/mL adenine, which had been used previously in the solid medium B experiments in Table 4.6 (15  $\mu$ g/mL hypoxanthine was used in medium A experiments). The results of these functional

complementation studies are shown in Figure 4.6, which is a summary of the growth curves shown in Figure 4.7. These results indicate that only PurE constructs with an *in vitro* activity are able to restore growth.

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
1)	CC0102		pJK173		30 h
2)	CC0102		pJK347		36 h
3)	CC0102		pJK348		-
4)	CC0102		pNC2		16 h
5)	CC0102		pJK426		-
6)	CC0102		pUC118		-

**Figure 4.6.** Schematic representation of strain CC0101 growth curves. Summary of results of minimal medium B growth curves shown in Figure 4.7. The first column gives an entry number, and the strain used. The second column indicates proteins provided by genomic DNA, PurE (rectangle) and PurK (rounded rectangle). Proteins are shaded according to origin *E. coli* (white) and *A. aceti* (grey). For deletion strains, genes which have been deleted are denoted by a large black X. The third and fourth columns report the plasmids used, and the proteins provided by those plasmids. Proteins are represented by different shapes, Orf1 (oval), PurE (rectangle), PurK (rounded rectangle). Proteins are shaded according to origin *E. coli* (white) and *A. aceti* (grey). In the 'Growth' column is the time taken for each culture to reach  $A_{600} = 0.3$ .







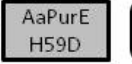



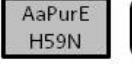


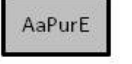


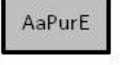

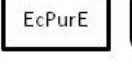
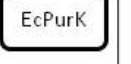





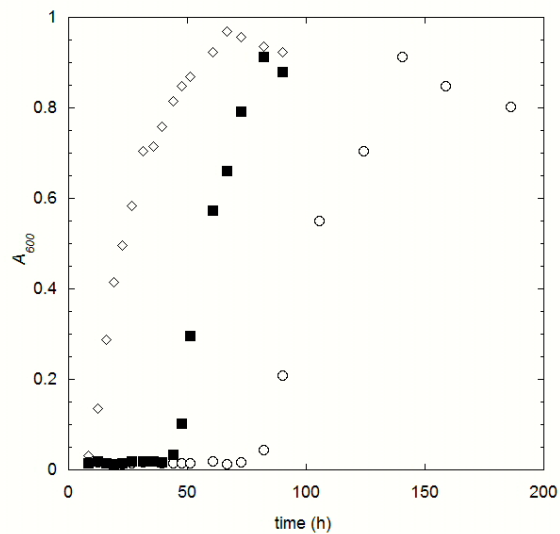
**Figure 4.7:** Growth curves of strain CC0102 in minimal medium B. Strain CC0102 was transformed with plasmids pNC2 (*purE<sub>Ec</sub>purK<sub>Ec</sub>* open squares), pJK173 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>* open triangles), or pJK347 (*orf1<sub>Aa</sub>purEH59D<sub>Aa</sub>purK<sub>Aa</sub>* open circles). Growth was not observed in cultures of CC0102/pUC118, CC0102/pJK348 (*orf1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK<sub>Aa</sub>*) or CC0102/pJK426 (*purEH45N<sub>Ec</sub>purK<sub>Ec</sub>*) through 1000 h. Growth was noted in all positive controls which were supplemented at 100 µg/mL adenine (not shown).

*Functional complementation of strain CC0101 ( $\Delta$ *purE<sub>Ec</sub>*) in liquid minimal medium*

B. Plasmids pJK173, pJK347, pJK348, pJK324, pJK412 and pNC2 were assessed for their ability to restore prototrophy to strain CC0101 in minimal medium B. Plasmid pUC118 was used as a negative control. Positive controls contained exogenous purines in the form of 100 µg/mL adenine. The results of these functional complementation studies are shown in Figure 4.8, which is a summary of the growth curves shown in Figure 4.9.

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
7)	CC0101	 EcPurK	pJK173	 Orf1  AaPurE  AaPurK	60 h
8)	CC0101	 EcPurK	pJK347	 Orf1  AaPurE H59D  AaPurK	105 h
9)	CC0101	 EcPurK	pJK348	 Orf1  AaPurE H59N  AaPurK	-
10)	CC0101	 EcPurK	pJK324	 AaPurE	-
11)	CC0101	 EcPurK	pJK412	 Orf1  AaPurE	-
12)	CC0101	 EcPurK	pNC2	 EcPurE  EcPurK	19 h
13)	CC0101	 EcPurK	pUC118		-

**Figure 4.8:** Schematic representation of strain CC0101 growth curves. Results of minimal medium B growth curves shown in Figure 4.9. In the ‘Growth’ column is the time taken for each culture to reach  $A_{600} = 0.3$ .



**Figure 4.9:** Growth curves of strain CC0101 in minimal medium B. Strain CC0101 was transformed with pNC2 (open diamonds), pJK173 (filled squares), and pJK347 (open circles). Strain CC0101 was also transformed with pUC118, pJK324, pJK348 and pJK412 but no growth was measured.

As expected, plasmids which contained active forms of PurE (pNC2, pJK173 and pJK347) were able to restore prototrophy to strain CC0101. However, constructs that

contained an active PurE but lacked PurK (pJK324 and pJK412) failed to restore prototrophy to strain CC0101. Although this might indicate that *purE<sub>Aα</sub>* requires *purK<sub>Aα</sub>* to function, later experiments (see following sections) indicate that a plasmid containing just *purE<sub>Ec</sub>* restores prototrophy to strain CC0101 poorly on solid and in liquid media. This suggests that expression of *purK<sub>Ec</sub>* is affected in strain CC0101. Because cultures were only observed for 186 h, it is uncertain if growth would have been observed at a later time. Additionally these cultures were inoculated from stationary phase cultures. For all other liquid medium experiments, cultures were inoculated from log phase cultures. Growth was noted at later time for CC0101 than with CC0102 when the strains were transformed with the same plasmids (pJK173, entries 1 and 7; pJK347, entries 2 and 8; pNC2, entries 5 and 12). Although these results are from different strains, this may indicate that cells taken during the log-phase are healthier.

*Effects of purE<sub>Ec</sub> and purE<sub>Ec</sub>purK<sub>Ec</sub> constructs on the growth of strains CC0101, CC1201 and CC1202 on solid media.* *EcPurE* constructs were tested for their ability to restore prototrophy to strains CC0101, CC1201 and CC1202 on solid medium A or solid medium B. Experiments were performed on both types of media so the results of the candle jar experiment with CC0101/pJK419 could be directly compared to the identical conditions used with PC0135. As expected, both plasmids pJK419 (*purE<sub>Ec</sub>*) and pNC2 (*purE<sub>Ec</sub>purK<sub>Ec</sub>*) restored growth to the strains (Table 4.7). However, the degree to which plasmid pJK419 restored growth was dependent on the strain.

Plasmid pJK419 supported the growth of strain CC1202 at levels similar to pNC2. In strain CC1201, plasmid pJK419 allowed for growth at a slightly lower level, with small

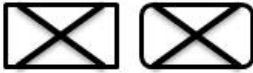


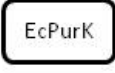



colonies being noted at 24 h. In strain CC0101, plasmid pJK419 gave much slower growth with small colonies being noted at 72 h. At 72 h, the growth of the duplicate plate of CC0101/pJK419 on solid medium A from the candle jar was observed, and significantly larger colonies were found. This indicates that the deletion of *purE<sub>EC</sub>* in strain CC0101 may have a polar effect on *purK<sub>EC</sub>* expression.

**Table 4.7:** Functional complementation studies of strains CC0101, CC1201, and CC1202.<sup>a</sup>

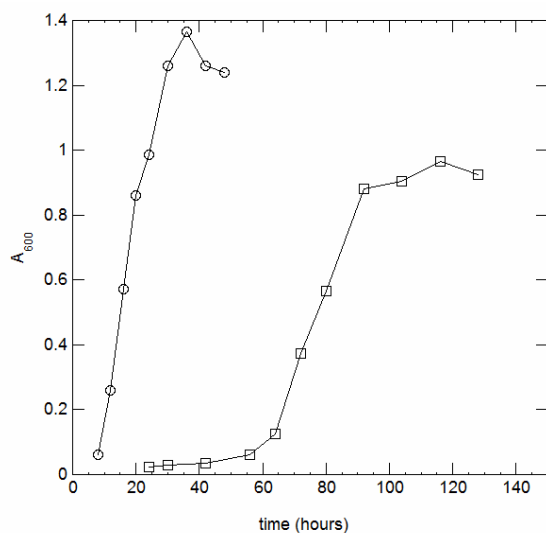
Strain	pNC2		pJK419		
	Medium A <sup>b</sup> (solid)	Medium B <sup>b</sup> (solid)	Medium A (solid)	Medium A (solid) + high CO <sub>2</sub>	Medium B (solid)
CC0101 ( $\Delta purE_{EC}$ )	++++ <sup>c</sup>	++++	+	++	+
CC1201 ( <i>purE<sub>EC</sub></i> <> <i>purE<sub>Aa</sub></i> )	++++	++++	+++	Nd	+++
CC1202 ( <i>purE<sub>EC</sub></i> <> <i>purE<sub>Aa(ATG)</sub></i> )	++++	++++	++++	Nd	++++

<sup>a</sup>A set of plates supplemented with exogenous purines in the form of 15 µg/mL hypoxanthine for medium A or 100 µg/mL adenine for medium B. These controls gave (++++) growth. <sup>b</sup> minimal media A and B as described in Experimental section. <sup>c</sup> key to symbols: +++++, large colonies at 24 h; +++, small colonies at 24 h; ++, large colonies at 72 h; +, small colonies at 72 h; Nd, not determined.

*Functional complementation studies of strain CC0101 ( $\Delta purE_{EC}$ ) in liquid minimal medium.* Plasmid pJK419 was assessed for its ability to restore prototrophy to strain CC0101 in minimal medium B. The results are summarized in Figure 4.10, which is based on the growth curve in Figure 4.11. The results are shown in reference to strain CC0102/pNC2. Plasmid pJK419 was found to complement the auxotrophy less well than pNC2 complemented CC0102. This indicates that the deletion of *purE<sub>EC</sub>* in strain CC0101 either has a polar effect on *purK<sub>EC</sub>* expression or the selective over expression of *purE<sub>EC</sub>* inhibits growth.

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
4)	CC0102		pNC2		16 h
14)	CC0101	 	pJK419		72 h
15)	CC0101	 	pUC118		-

**Figure 4.10:** Schematic representation of strain CC0101 growth curves. Summary of minimal medium B growth curves shown in Figure 4.11. In the ‘Growth’ column is the time taken for each culture to reach  $A_{600} = 0.3$ , dash denotes that no growth was observed up to 1000 h.



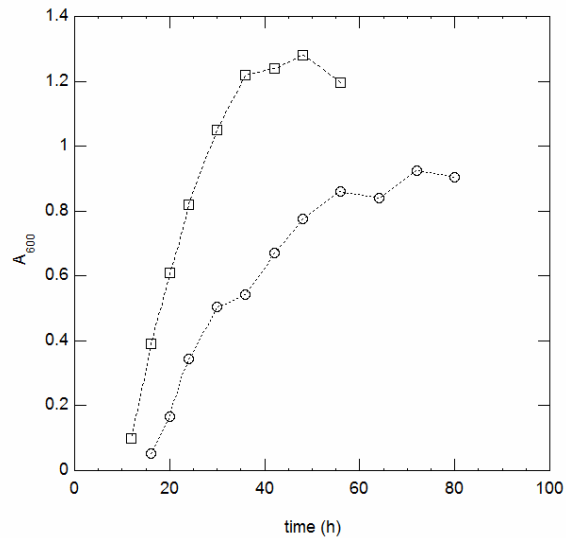
**Figure 4.11:** Growth curves for strain CC0101 in minimal medium B. The strain was transformed with plasmid pJK419 (*purE<sub>Ec</sub>*). Also shown for comparison is the growth curve for CC0102/pNC2 from Figure 4.7. Strain CC0101/pJK419 (open squares); Strain CC0102/pNC2 (open circles).

*Study of wild type strain W3110.* For comparison to other strains, the wild type strain W3110 was transformed with pUC118 and its growth was observed in minimal medium B. Plasmid pJK455 (*orf1<sub>Aa</sub>*) was also assessed for its effect on the growth of strain W3110. Results are shown in Figure 4.12, which is a summary of growth curves shown in Figure 4.13.

The level of expression of each protein is expected to be based on the copy number of the gene. Strain W3110 contains a genomic copy of *purE<sub>Ec</sub>* and *purK<sub>Ec</sub>*, and will express the enzymes encoded at a level that is sufficient to support growth. Plasmid pJK455 is a high-copy number plasmid. This means that when transformed with pJK455, W3110 will contain many copies of the *orf1<sub>Aa</sub>* gene, and should express a comparatively high level of the encoded protein.

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
16)	W3110	<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 2px; text-align: center;">EcPurE</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">EcPurK</div> </div>	pUC118		16 h
17)	W3110	<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 2px; text-align: center;">EcPurE</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">EcPurK</div> </div>	pJK455	<div style="border: 1px solid black; border-radius: 50%; padding: 5px; display: inline-block;">Orf1</div>	24 h

**Figure 4.12.** Schematic representation of strain W3110 growth curves. Results of minimal medium B growth curves shown in Figure 4.13. In the 'Growth' column is the time taken for each culture to reach  $A_{600} = 0.3$ .



**Figure 4.13:** Growth of strain W3110 in minimal medium B. W3110 was transformed with pUC118 (open squares) or pJK455 (open circles).

Strain W3110/pUC118 grew rapidly after a short lag phase in minimal medium B. However, when strain W3110 was transformed with pJK455 (*orf1<sub>Aa</sub>*), the culture grew more slowly (Figure 4.13). The effect of pJK455 on strain W3110 is different than in other strains reported below.

*Study of strains CC0102 ( $\Delta purE_{Ec}\Delta purK_{Ec}$ ), CC0103 ( $\Delta purK_{Ec}$ ), CC1203 ( $purK_{Ec}\langle\rangle purK_{Aa}$ ), CC1204 ( $purE_{Ec}purK_{Ec}\langle\rangle purE_{Aa}purK_{Aa}$ ) and CC1205 ( $purE_{Ec}purK_{Ec}\langle\rangle purE_{Aa(ATG)}purK_{Aa}$ ) on solid medium B.* The ability of strains CC0102, CC1203, CC1204 or CC1205 to grow on solid medium B was assessed (Table 4.8). As expected, strain CC0102/pUC118 and CC0103/pUC118 failed to grow without the addition of exogenous purines. Surprisingly, strains CC1204/pUC118 and CC1205/pUC118 failed to grow in the time that the plates were observed, while strain CC1203/pUC118 and CC0102/pJK173 grew rapidly.

The difference seen between the *purK<sub>Aa</sub>* replacement strain (CC1203/pUC118), and the *purE<sub>Aa</sub>purK<sub>Aa</sub>* double replacement strains (CC1204/pUC118 and CC1205/pUC118) could be due to several possibilities. The replacement of *purE<sub>Ec</sub>* with *purE<sub>Aa</sub>* could change regulatory elements required for *purK* expression. However, the expression level of *purK<sub>Ec</sub>* in strains CC1201 and CC1202 (Table 4.7) is high enough to support growth on solid medium B when a source of *purE<sub>Ec</sub>* is provided. This suggests that the level of expression of *purK<sub>Aa</sub>* in the double replacement strains should be adequate to support growth. Another possibility is that *AaPurE* is not as active as *EcPurE*. However, both enzymes have been purified, and exhibit similar in vitro activities (Chapter 3) (23).

An alternate possibility requiring association of proteins in that PurE takes part in a multi-enzyme complex with other enzymes in purine biosynthesis. The replacement of *EcPurE* with *AaPurE* leads to a failure of the formation of this complex that can be compensated by the increased expression of *AaPurE* from a plasmid. This would allow for CC0102/pJK173 to grow, although CC1204/pUC118 and CC1205/pUC118 do not.

An additional possibility is that strains CC1204 and CC1205 do not contain all of the genes which allow plasmid pJK173 to restore prototrophy to deletion strain CC0102. Contained within the insert of pJK173 is the gene *orf1*. The product of this gene may be required for the proper functioning of *AaPurE* and *AaPurE* as will be discussed in greater detail later.

**Table 4.8:** Assessment of the ability of strains CC0102, CC0103, CC1203, CC1204 and CC1205 to grow on solid medium B.

Strain/plasmid	Medium B (solid)	Medium B + Adenine (solid)
CC0102 ( $\Delta purE_{Ec}\Delta purK_{Ec}$ )/pUC118	-	+++
CC0102 ( $\Delta purE_{Ec}\Delta purK_{Ec}$ )/pJK173	+++	+++
CC0103 ( $\Delta purK_{Ec}$ )/pUC118	-	+++
CC1203 ( $purK_{Ec}\langle\rangle purK_{Aa}$ )/pUC118	+++	+++
CC1204 ( $purE_{Ec}purK_{Ec}\langle\rangle purE_{Aa}purK_{Aa}$ )/pUC118	-	+++
CC1205 ( $purE_{Ec}purK_{Ec}\langle\rangle purE_{Aa(ATG)}purK_{Aa}$ )/pUC118	-	+++

<sup>a</sup>key to symbols: +++, colonies noted at 30 h which become large (> 2mm) by 60h; -, tiny colonies which do not increase in size with time up to 120 h.

*Study of strains CC1201 and CC1202, strains with the precise gene replacement of  $purE_{Ec}$  with  $purE_{Aa}$ . Strain CC1202 ( $purE_{Ec}\langle\rangle purE_{Aa(ATG)}$ ) was chosen for the focus of this*



study because it began with the same start codon as *purE<sub>Ec</sub>*. This can affect expression levels; generally, ATG gives a higher expression level than GTG or other codons (25).

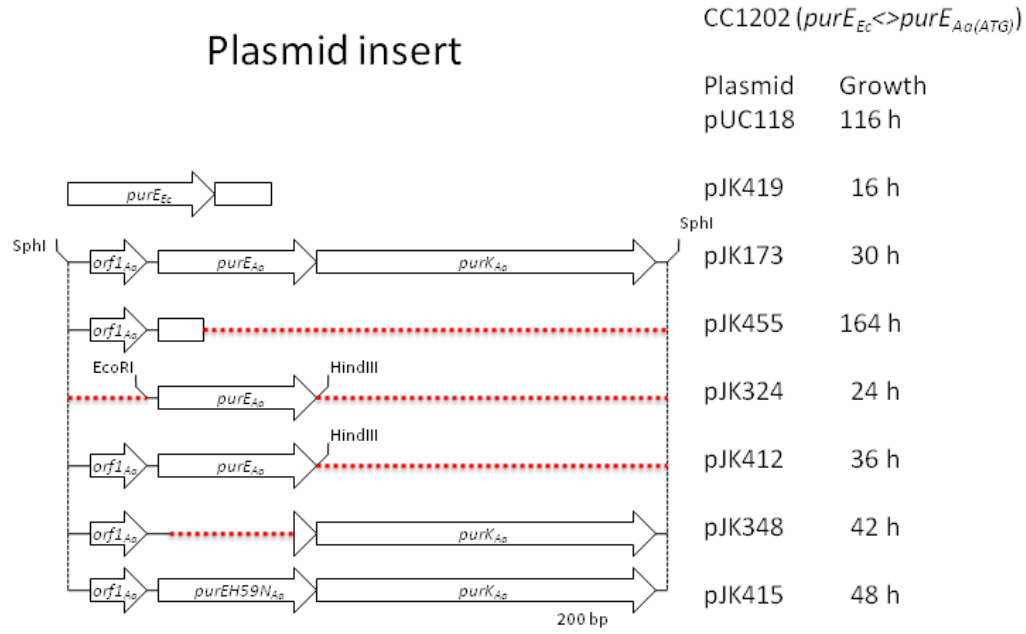
Plasmids pJK348, pJK419, pUC118, pJK173 or plasmids containing portions of the *A. aceti* genome contained in the insert of plasmid pJK173 (pJK324, pJK412, pJK415 and pJK455) were assessed for their effect on insert strains CC1202 in minimal medium B. In this experiment, the *purE<sub>Ec</sub>* deletion strain CC0101 transformed with pUC118 was used as a negative control. The results of these experiments are summarized in Figure 4.14, which includes an illustration of the inserts for each plasmid. Figure 4.14 is a summary of Figure 4.15, which was determined from the growth curves shown in Figures 4.16.

The level of expression of each protein is expected to be based on the copy number of the gene and may be affected by the start codon of the gene. Strain CC1202 contains a genomic copy of *purE<sub>Aa(ATG)</sub>*. This gene starts with an ATG start codon and should express *AaPurE* at similar levels to *EcPurE* in strain W3110. Strain CC1202 contains a genomic copy of *purK<sub>Ec</sub>* and is expected to express *EcPurK* at a level similar to W3110.

Plasmid pJK419 is derived from the moderate-copy number protein expression vector pET23a. All of the plasmids carrying *A. aceti* genes are based on the high copy number vector pUC118. These plasmids should provide many copies of the introduced genes, and lead to an increased level of expression for the encoded proteins.

Strain CC1202/pUC118 grew after a prolonged lag phase indicating inefficient synthesis of purines. Although growth after a prolonged period could indicate

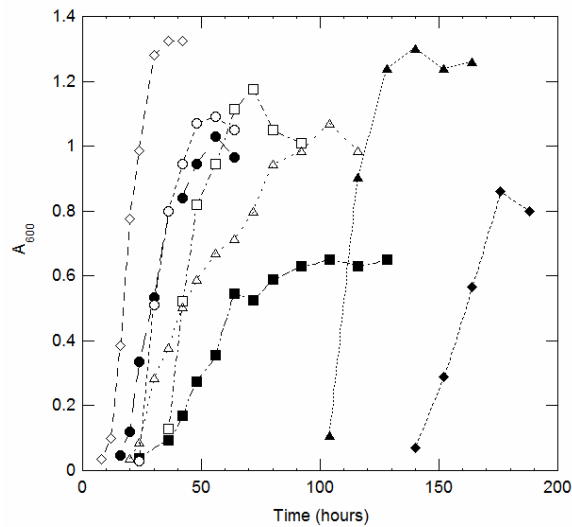
contamination by other bacteria, we note that the negative control CC0101/pUC118 did not show growth at 1000 h.



**Figure 4.14:** Illustration of plasmids used in growth experiments and summary of lag phases from Figure 4.15.

Entry Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
18) CC1202	AaPurE (ATG) EcPurK	pUC118		116 h
19) CC1202	AaPurE (ATG) EcPurK	pJK324	AaPurE	24 h
20) CC1202	AaPurE (ATG) EcPurK	pJK412	Orf1 AaPurE	36 h
21) CC1202	AaPurE (ATG) EcPurK	pJK415	Orf1 AaPurK	48 h
22) CC1202	AaPurE (ATG) EcPurK	pJK348	Orf1 AaPurE H59N AaPurK	42 h
23) CC1202	AaPurE (ATG) EcPurK	pJK173	Orf1 AaPurE AaPurK	30 h
24) CC1202	AaPurE (ATG) EcPurK	pJK455	Orf1	164 h
25) CC1202	AaPurE (ATG) EcPurK	pJK419	EcPurE	16 h

**Figure 4.15:** Schematic representation of strain CC1202 growth curves. Results of medium B growth curves shown in Figure 4.16. In the ‘Growth’ column is the time taken for each culture to reach  $A_{600} = 0.3$ .



**Figure 4.16:** Growth curves for strain CC1202 in minimal medium B. Strain CC1202 was transformed with pUC118 (empty vector control, filled triangles); pJK419 (*purE<sub>Ec</sub>*, open diamonds); pJK173 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>*, open circles); pJK348 (*orf1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK<sub>Aa</sub>*, open squares); pJK412 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>*, open triangles); pJK324 (*purE<sub>Aa</sub>*, filled circles); pJK415 (*orf1<sub>Aa</sub>purK<sub>Aa</sub>*, filled squares); pJK455 (*orf1<sub>Aa</sub>*, filled diamonds).

The CC1202/pJK419 culture was found to grow rapidly. The CC1202/pJK419 culture grew at a rate similar to that seen for the  $\Delta purE_{Ec}\Delta purK_{Ec}$  deletion strain CC0102/pNC2 (entry 4), and W3110/pUC118 (entry 16). This indicates that the substitution of  $purE_{Ec}$  with  $purE_{Aa}$  does not have a polar effect on the expression of genomic  $purK_{Ec}$ . This indicates that the cause of the prolonged lag phase seen in CC1202/pUC118 is due to  $purE_{Aa(ATG)}$  and not to an unforeseen polar effect on  $purK_{Ec}$ .

The addition of plasmid pJK173 ( $orf1_{Aa} purE_{Aa} purK_{Aa}$ ) resulted in CC1202 having a shorter lag phase than CC1202/pUC118. This was expected because plasmid pJK173 restores prototrophy to the  $\Delta purE_{Ec}\Delta purK_{Ec}$  deletion strain CC0102. Clearly  $AaPurK$  and  $AaPurE$  are functional in *E. coli*.

The addition of plasmid pJK324 ( $purE_{Aa}$ ) resulted in CC1202 having a shorter lag phase. The effect of pJK324 was similar to that of pJK173 ( $orf1_{Aa} purE_{Aa} purK_{Aa}$ ). This indicates that increased expression of  $AaPurE$  allows for the cause of the prolonged lag phase to be overcome. While at first it would appear that the cause of the prolonged lag phase seen in CC1202/pUC118 could be due to either poor expression or low activity of the genomic  $purE_{Aa(ATG)}$ , a different result is seen when added  $purE_{Aa}$  is provided by plasmid pJK412 (see next paragraph).

While the addition of plasmid pJK412 ( $orf1_{Aa} purE_{Aa}$ ) resulted in strain CC1202 having a shorter lag phase than CC1202/pUC118, this plasmid complemented less efficiently than pJK324 or pJK173.

The addition of plasmids that contain  $purK_{Aa}$  also decreased the lag phase of strain CC1202. The addition of pJK348 ( $orf1_{Aa} purEH59N_{Aa} purK_{Aa}$ ) shortened the lag

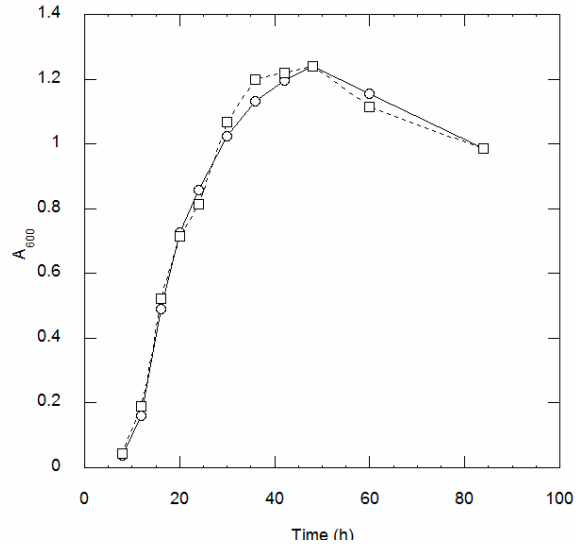
phase of strain CC1202 despite the presence of an entirely inactive *AaPurE*, and complemented almost as well as pJK173. The addition of pJK415 (*orf1<sub>Aa</sub>purK<sub>Aa</sub>*) reduced the lag phase, but this culture grew more slowly and to a lower total cell density than CC1202/pJK348.

Plasmid pJK455 (*orf1<sub>Aa</sub>*) caused an increase in the lag phase of strain CC1202. This indicates that *AaOrf1* exacerbates the cause of the lag phase. In CC1202/pJK455, the amount of *AaOrf1* produced should be significantly higher than the *AaPurE* and *EcPurK* produced from the genome.

*Study of strain CC1203, a strain with the precise gene replacement of purK<sub>Ec</sub> with purK<sub>Aa</sub>.* Strain CC1203 (*purK<sub>Ec</sub><>purK<sub>Aa</sub>*) was transformed with pUC118 and grown in minimal medium B. The growth of CC1203/pUC118 was compared to W3110/pUC118 (Figures 4.17 and 4.18). Strain CC0102/pUC118 was used as a negative control. Strikingly, CC1203/pUC118 grew rapidly, with a profile similar to W3110/pUC118. Thus the substitution of *EcPurK* with *AaPurK* has almost no effect, in marked contrast with the corresponding *PurE* substitution (entry #18).

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
26)	CC1203	<div style="display: inline-block; border: 1px solid black; padding: 2px; margin-right: 10px;">EcPurE</div> <div style="display: inline-block; border: 1px solid black; padding: 2px;">AaPurK</div>	pUC118		16 h
16)	W3110	<div style="display: inline-block; border: 1px solid black; padding: 2px; margin-right: 10px;">EcPurE</div> <div style="display: inline-block; border: 1px solid black; padding: 2px;">EcPurK</div>	pUC118		16 h

**Figure 4.17.** Schematic representation of strain CC1203 and W3110 growth curves. Results of minimal medium B growth curves shown in Figure 4.18. In the ‘Growth’ column is the time taken for each culture to reach  $A_{600} = 0.3$ .



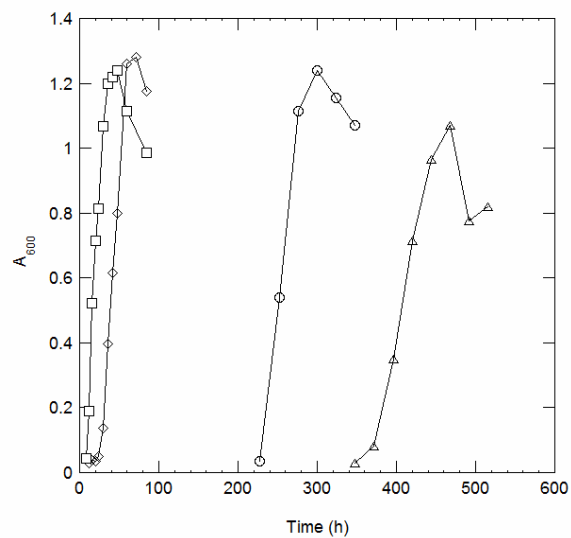
**Figure 4.18:** Growth of strain CC1203 and W3110 in minimal medium B. Strain W3110 (open squares) and strain CC1203 (open circles) were transformed with pUC118.

*Study of strain CC1204, a strain containing a precise gene replacement of  $purE_{Ec}purK_{Ec}$  with  $purE_{Aa}purK_{Aa}$ . Plasmids pJK173, pJK455 and pUC118 were assessed for their effect on strain CC1204 ( $purE_{Ec}purK_{Ec} \leftrightarrow purE_{Aa}purK_{Aa}$ ). These experiments were performed at the same time as entries 26 and 16). The results of this experiment are shown in Figure 4.19, which is a summary of the growth curves shown in Figure 4.20.*

The level of expression of each protein is expected to be based on the copy number of the gene and may be affected by the start codon of the gene. Strain CC1204 contains a genomic copies of  $purE_{Aa}$  and  $purK_{Aa}$ . Because  $purE_{Aa}$  begins with a GTG start codon, the level of expression of  $AaPurE$  in strain CC1204 should be somewhat lower than  $EcPurE$  in strain W3110. The reduced expression level of  $purE_{Aa}$  may cause a reduction in the expression level of the genomic  $purK_{Aa}$ .

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
16)	W3110	EcpurE EcpurK	pUC118		16 h
28)	CC1204	AaPurE AaPurK	pUC118		252 h
29)	CC1204	AaPurE AaPurK	pJK455	Orf1	396 h
30)	CC1204	AaPurE AaPurK	pJK173	Orf1 AaPurE AaPurK	36 h

**Figure 4.19.** Schematic representation of strain CC1204 and W3110 growth curves. Results of minimal medium B growth curves shown in Figure 4.20. In the 'Growth' column is the time taken for each culture to reach  $A_{600} = 0.3$ .



**Figure 4.20:** Growth curves for strain CC1204 and W3110 in medium B. Strain CC1204 was transformed with plasmids pJK173 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>*, open diamonds), pJK455 (*orf1<sub>Aa</sub>*, open triangles) or pUC118 (empty vector control, open squares). Strain W3110 was transformed with pUC118 (open squares). The growth curve for W3110/pUC118 is also shown in Figure 4.18.

Plasmids pJK455, pJK173 and pUC118 are all high-copy number plasmids. When transformed with pJK455 and pJK173, strain CC1204 will contain many copies of the genes encoded by the plasmid, and thus should have increased expression of these proteins.

When transformed with pUC118, strain CC1204 grew but only after a very long lag phase. This lag phase was longer than that seen with CC1202/pUC118, and was longer than the lag phase seen with CC1205/pUC118 below. This is consistent with a lower level of expression of *purE<sub>Aa</sub>*.

Plasmid pJK173 reduced the lag phase in strain CC1204. The growth seen in CC1204/pJK173 was similar to the growth seen previously in other strains transformed with pJK173 and indicates that the increased levels of *AaOrf1*, *AaPurE*, and *AaPurK* compensate for whatever deficiency causes the prolonged lag phase.

Surprisingly plasmid pJK455 caused an increase in the lag phase of strain CC1204. This indicates that *AaOrf1* exacerbates the cause of the lag phase. In strain CC1204/pJK455, the amount of *AaOrf1* produced should be significantly greater than the *AaPurE* and *AaPurK* produced from the genome.

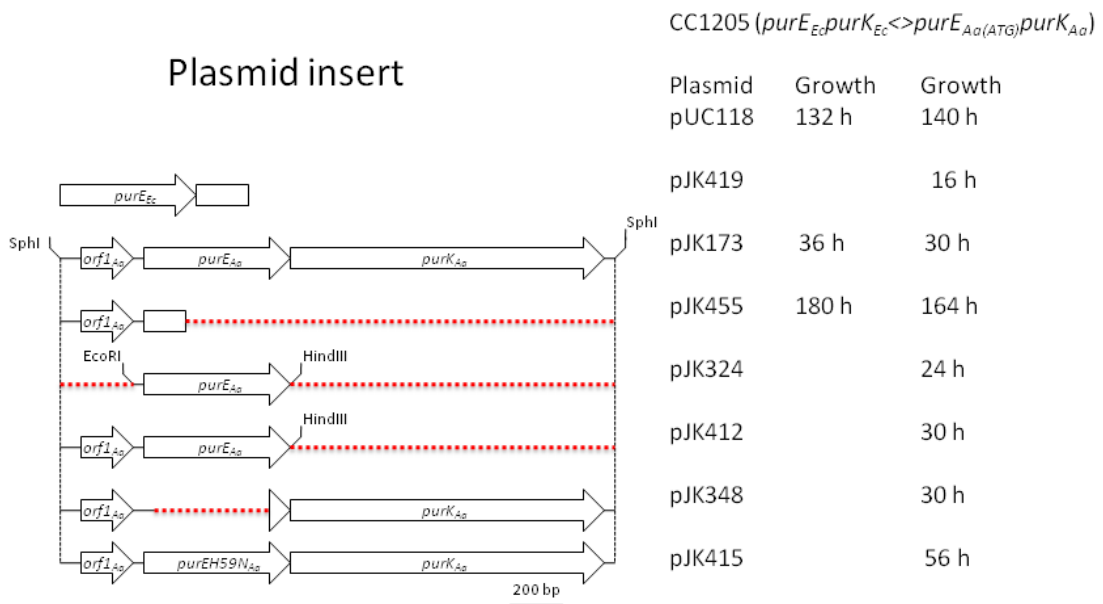
*Study of strain CC1205, a strain containing a precise gene replacement of *purE<sub>Ec</sub>purK<sub>Ec</sub>* with *purE<sub>Aa(ATG)</sub>purK<sub>Aa</sub>*. Plasmids pJK348, pJK419, pUC118, pJK173 or plasmids containing portions of the insert of plasmid pJK173 (pJK324, pJK412, pJK415 and pJK455) were assessed for their effect on replacement strain CC1205 (*purE<sub>Ec</sub>purK<sub>Ec</sub>* <> *purE<sub>Aa(ATG)</sub>purK<sub>Aa</sub>*) in minimal medium B. The results of these experiments are summarized in Figure 4.21, which includes an illustration of the inserts for each plasmid. Figure 4.21 is a summary of Figures 4.22 and 4.24, which were determined from the growth curves shown in Figures 4.21 and 4.23.*

Strain CC1205 contains a genomic copy of *purE<sub>Aa</sub>* and *purK<sub>Aa</sub>*. In strain CC1205 the genomic *purE<sub>Aa</sub>* has an ATG start codon, and should express *purE<sub>Aa</sub>* at a similar level



to *EcPurE* in W3110. Because the *purE<sub>Aa</sub>* gene starts with an ATG start codon, this strain would be expected to express *purK<sub>Aa</sub>* at similar levels to *purK<sub>Ec</sub>* in W3110 and CC1202. Qualitatively similar results were obtained with CC1204 and CC1205 indicating the start codon for *purE<sub>Aa</sub>* has a relatively subtle effect.

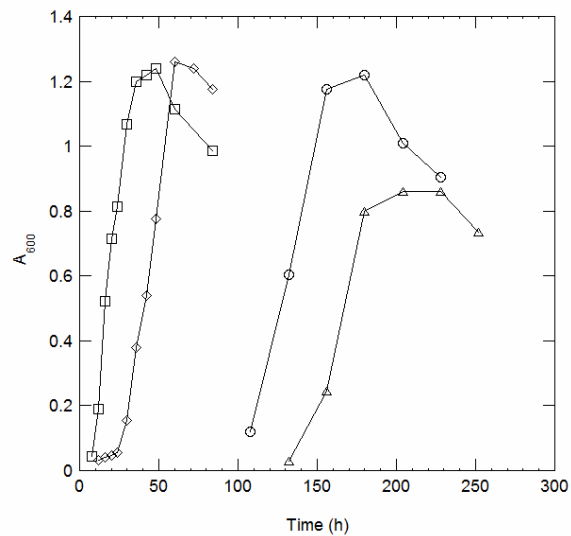
Slightly different times were noted for strain and plasmid combinations in separate experiments. This difference is likely due to the slight variation in disturbing the shaking incubator. In experiments reported in Figures 4.24 and 4.25, more flasks were being grown, requiring the shaking incubator to be disturbed more often.



**Figure 4.21:** Illustration of plasmids used in growth experiments and summary of lag phases from Figure 4.22 (first growth column) and Figure 4.24 (second growth column).

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
16)	W3110	EcpurE EcpurK	pUC118		16 h
31)	CC1205	AaPurE (ATG) AaPurK	pUC118		132 h
32)	CC1205	AaPurE (ATG) AaPurK	pJK455	Orf1	180 h
33)	CC1205	AaPurE (ATG) AaPurK	pJK173	Orf1 AaPurE AaPurK	36 h

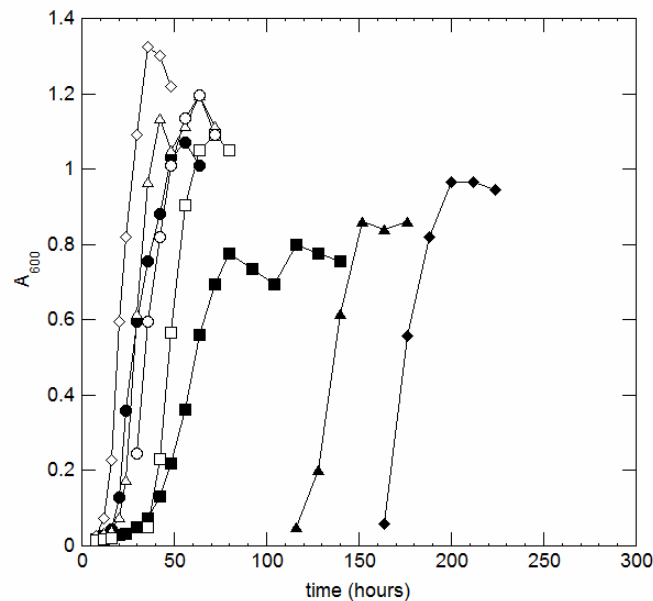
**Figure 4.22.** Schematic representation of strain CC1205 and W3110 growth curves. Results of minimal medium B growth curves shown in Figure 4.23. In the 'Growth' column is the time taken for each culture to reach  $A_{600} = 0.3$ .



**Figure 4.23:** Growth curves for strain CC1205 and W3110 in medium B. Strain CC1205 was transformed with plasmids pJK173 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>*, open diamonds), pJK455 (*orf1<sub>Aa</sub>*, open triangles) or pUC118 (empty vector control, open squares). Strain W3110 was transformed with pUC118 (open squares). The growth curve for W3110/pUC118 is also shown in (Figure 4.18).

Entry	Strain	Proteins from Genome	Plasmid	Proteins from Plasmid	Growth
34)	CC1205	AaPurE (ATG) AaPurK	pUC118		140 h
35)	CC1205	AaPurE (ATG) AaPurK	pJK324	AaPurE	24 h
36)	CC1205	AaPurE (ATG) AaPurK	pJK412	Orf1 AaPurE	30 h
37)	CC1205	AaPurE (ATG) AaPurK	pJK415	Orf1 AaPurK	56 h
38)	CC1205	AaPurE (ATG) AaPurK	pJK348	Orf1 AaPurE H59N AaPurK	48 h
39)	CC1205	AaPurE (ATG) AaPurK	pJK173	Orf1 AaPurE AaPurK	30 h
40)	CC1205	AaPurE (ATG) AaPurK	pJK455	Orf1	164 h
41)	CC1205	AaPurE (ATG) AaPurK	pJK419	EcPurE	16 h

**Figure 4.24:** Schematic representation of strain CC1205 growth curves. Results of medium B growth curves shown in Figure 4.25. In the ‘Growth’ column is the time taken for each culture to reach  $A_{600} = 0.3$ .



**Figure 4.25:** Growth curves for strain CC1205 in medium B. The strain was transformed with plasmids pUC118 (empty vector control, filled triangles); pJK419 (*purE<sub>EC</sub>*, open diamonds); pJK173 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>*, open circles); pJK348 (*orf1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK<sub>Aa</sub>*, open squares); pJK412 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>*, open triangles); pJK324 (*purE<sub>Aa</sub>*, filled circles); pJK415 (*orf1<sub>Aa</sub>purK<sub>Aa</sub>*, filled squares); pJK455 (*orf1<sub>Aa</sub>*, filled diamonds).

Plasmid pJK419 is derived from the moderate-copy number vector pET23a. All of the plasmids carrying *A. acetii* genes are made from the high copy number vector pUC118. All plasmids provide many copies of genes present, and should show an increased level of expression for encoded proteins.

When transformed with pUC118, strain CC1205 grew after a prolonged lag phase. While growth after a prolonged period may indicate contamination by other bacteria, the negative control for each set of experiments did not show any growth at 1000 h (for Figure 4.23 the negative control was CC0102/pUC118, for Figure 4.25 the negative control was CC0102/pJK348).

When strain CC1205 was transformed with plasmid pJK419, the culture grew rapidly. This indicates the deficiency is in apparent PurE activity. The CC1205/pJK419 culture grew at a rate similar to that seen for the  $\Delta purE_{Ec}\Delta purK_{Ec}$  deletion strain CC0102/pNC2 (entry 4), the *purE<sub>Aa(ATG)</sub>* insert strain CC1202/pJK419 (entry 25), the *purK<sub>Aa</sub>* insert strain CC1203/pUC118 (entry 26) and the wild type strain W3110/pUC118 (entries 16 and 35). A secondary result is that the substitution of *purE<sub>Ec</sub>* with *purE<sub>Aa</sub>* does not have a polar effect on the expression *purK<sub>Aa</sub>*, which is expressed in active form.

The addition of plasmid pJK173 (*orf1<sub>Aa</sub> purE<sub>Aa</sub> purK<sub>Aa</sub>*) to CC1205 resulted in a shorter lag phase. This was expected as plasmid pJK173 restores prototrophy to the deletion strain CC0102.

The addition of plasmid pJK324 (*purE<sub>Aa</sub>*) to CC1205 conferred more rapid growth. The effect of pJK324 was similar to that of pJK173. This indicates that increased

expression of *AaPurE* allows for cross-complementation. This result is similar to the result seen with CC1202/pJK324.

The addition of pJK412 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>*) to strain CC1205 had a similar effect to the addition of either pJK324 or pJK173. These results are different from those found in CC1202. The subtle difference in results seen appear to be dependent on the form of *purK* present in the genome of each strain.

The addition of plasmids containing both *orf1<sub>Aa</sub>* and *purK<sub>Aa</sub>* also decreased the lag phase of strain CC1205. The addition of pJK348 (*orf1<sub>Aa</sub> purEH59N<sub>Aa</sub> purK<sub>Aa</sub>*) shortened the lag phase of strain CC1205, and allowed for complementation similar to pJK173. The addition of pJK415 (*orf1<sub>Aa</sub>purK<sub>Aa</sub>*) reduced the lag phase, but the culture grew more slowly and to a lower total cell density than CC1205/pJK348.

Plasmid pJK455 (*orf1<sub>Aa</sub>*) caused an increase in the lag phase of strain CC1205. This indicates that *AaOrf1* exacerbates the cause of the lag phase. In strain CC1205/pJK455, the amount of *AaOrf1* produced should be significantly higher than the *AaPurE* or *AaPurK* produced from the genome. This is very similar to the comparable results with CC1202 (entry #24) and CC1204 (entry # 29).

*Observation of phenotype.* When the insert and deletion strains are transformed with plasmids that lack *purE<sub>Ec</sub>*, satellite colonies begin to form after ~14 h of growth on LB/Amp. Due to this, colonies for starter cultures were picked at or before 12 h incubation. Liquid minimal medium experiments also developed a light yellow color at later time points in growth curves. This is probably due to contamination at

later time points, but contamination does not affect the finding of when initial growth was observed.

#### 4.4 Discussion

*Gene context and expression.* To study the *purE-purK* operon of *A. acetii* in detail, a series of new strains were needed. These strains were designed to create defined systems that would contain a single copy of each gene. For the gene replacement strains, the strains were designed to contain *A. acetii* genes in the precise location of their *E. coli* counterparts. This was done to ensure that these genes would be expressed at the same level and would be under the same regulation as their *E. coli* counterparts in strain W3110.

The expression level of a gene will depend on several factors that were carefully controlled in the current study. One is the number of copies of the gene present. For genomic copies, there will be a single copy per cell. For W3110 and the replacement strains, each contains one copy of a *purE* gene and one copy of a *purK* gene. When a gene is contained on a plasmid, the number of copies of that gene present in a cell will depend on the copy number of the plasmid.

The moderate- and high-copy number vectors pET23a or pUC118 were used for all plasmids in the functional complementation (pET23a was used to construct pNC2, pJK419, pJK426 and pJK435 all of which contain *E. coli* genes; pUC118 was used to construct all plasmids containing *A. acetii* genes). The pET23a vector was used for *E. coli* genes because they are homologous to wild type genes and should be expressed well. Additionally, previous studies of PC0135 involved the pET23a derived pNC2. Cells

containing the pET23a and pUC118 derived plasmids will contain hundreds of copies of each gene present in the insert. The increased number of genes should give an expression level that is significantly higher than that from single-copy genomic inserts.

The start codon of a gene will also affect the expression level of a protein. While most bacterial genes start with an ATG start codon, some start with an alternative codon. Among the alternative start codons, GTG is the most common (25). For the ATG start codon, all three nucleotides are complementary to the Met initiator tRNA anticodon (25). For the GTG start codon, only two of the nucleotides are complementary to the initiator anticodon (25). Thus a gene with an ATG start codon should be expressed at a higher level than the same gene with a GTG codon. Due to the difference in start codons, replacement strains CC1201 and CC1204 would be expected to express less *purE<sub>Aa</sub>* than replacement strains CC1202 or CC1205.

Polar effects can significantly impact levels of expression. In the translationally coupled *trp* operon, the early termination of one gene can lead to an up to 98% reduction in the detectable activity of enzymes encoded by the genes that follow (1). This was more difficult to control in the current experiments, given that the effects of the different *purK* control sequences within *purE<sub>Aa</sub>* are unknown.

*Characterization of auxotrophic strains PC0135.* The cause of the *purE* auxotrophy in strain PC0135 has been found to be a single point mutation which changes residue Trp151 in *EcPurE* to a premature stop codon (TGG → TGA). This results in the deletion of the protein's final helix, which has extensive interaction with the other subunits in the octamer. This led to speculation that this truncated protein allows low-

level complementation by the apparently completely inactive mutants *EcPurEH45N* and *AaPurEH59N* (18, 23).

To explore this possibility, strain CC0102 ( $\Delta purE_{Ec}\Delta purK_{Ec}$ ) was designed to remove the entire *purE<sub>Ec</sub>purK<sub>Ec</sub>* coding region in the genome. This created a strain that lacked any form of *EcPurE*. As expected, plasmids containing the completely inactive mutants *purEH59N<sub>Aa</sub>* (pJK348: *orf1<sub>Aa</sub>purEH59N<sub>Aa</sub>purK<sub>Aa</sub>*) and *purEH45N<sub>Ec</sub>* (pJK426: *purEH45N<sub>Ec</sub>purK<sub>Ec</sub>*) failed to restore prototrophy in strain CC0102. These results indicate that these mutants lack sufficient in vivo activity to support growth, which was expected from the lack of any detectable in vitro activity for these mutants (Chapter 3)(23).

The results from PC0135/pJK426 and CC0102/pJK426 are consistent with the proposal that the truncated *EcPurE* encoded by PC0135 is stabilized by the expressed inactive mutants *EcPurEH45N* leading to complementation. Although most of the following work was done with minimal medium B, solid media experiments with CC0102/pJK426 were also done on same medium as the other PC0135 experiments (minimal medium A). Using this medium controls for the possibility that the difference in media had an effect on functional complementation.

The outcome of the investigation of strain PC0135 also suggests that that the point mutation has a polar effect on the expression of *purK<sub>Ec</sub>*. The addition of single construct *EcPurE* (pJK419) did not restore growth as well as the addition of the dual construct *purE<sub>Ec</sub>purK<sub>Ec</sub>* plasmid pNC2. Although the slow growth of PC0135/pJK419 could be due to the effect of increased expression of *EcPurE*, a similar effect is not seen



with pJK419 in replacement strains CC1202 and CC1205 which contain an uninterrupted *purE*.

Because strain PC0135 has been shown to have a leaky *purE* phenotype that appears to have a polar effect on the expression of *purK*, no further studies were performed with this strain.

*Factors to consider from purine biosynthesis.* The de novo biosynthesis of purines is not an isolated pathway.

The biosynthesis of the cofactor thiamine uses an intermediate from purine biosynthesis. Isotopic labeling experiments in *E. coli* have shown that the amino-methyl-pyrimidine ring of thiamine derives from AIR (26, 27). Although the consumption of AIR in the synthesis of thiamine should be small, care was taken to supplement the minimal media with thiamine. This was done to ensure that thiamine deficiencies did not affect our results in any way.

Although IMP can be synthesized starting with 5-phosphoribosyl 1-pyrophosphate, it is not the only inlet into purine biosynthesis. AICAR, the substrate of the final enzyme in the synthesis of IMP, is also produced during the *de novo* synthesis of histidine. If *E. coli* is grown in minimal media, and the final two steps of the purine biosynthesis are blocked by chemical inactivation, AICAR accumulates. Supplementation of the media with histidine causes a 45% reduction in accumulated AICAR (28). This reduction is likely due to feedback inhibition of histidine biosynthesis,

and indicates that histidine biosynthesis can have a significant contribution to purine biosynthesis.

*Assessment of growth experiments.* Most of the strain and plasmid combinations have only been grown once, so caution is advised. Experiments that have been repeated more than once yielded similar lag times (W3110/pUC118, entries #16 and #27; CC1205/pUC118, entries #31 and #34; CC1205/pJK455, entries #32 and #40; CC1205/pJK173, entries #33 and #39). Moreover, experiments in strain CC1202 (*purE<sub>Ec</sub>*<>*purE<sub>Aa</sub>*) and CC1205(*purE<sub>Ec</sub>purK<sub>Ec</sub>*<>*purE<sub>Aa</sub>purK<sub>Aa</sub>*) generally yield similar results, supporting conclusions obtained with these two closely related strains. In the case of the difference in growth between CC1202/pJK412 (entry #20) and CC1205/pJK412 (entry #36), this observation has been observed once and should be confirmed.

*Prolonged lag phase.* Typically upon inoculation into minimal media, *E. coli* cells quickly acclimate during a brief lag phase and begin to divide. During this lag phase, bacteria produce the enzymes required for the production of all of the compounds necessary for growth formally provided in the richer media in which they were first grown. The lack of growth for some strain-plasmid combinations (CC0102/pJK348, entry #3; CC0102/pUC118, entry #6; CC0102/pJK426, entry #5) and the different length of lag phase for various strains and plasmid combinations in minimal media made it impractical to inoculate observed cultures from starter cultures grown in minimal media.

Although it could be possible that contamination by another bacterium could lead to growth seen at later time points, this seems unlikely for several reasons. 1) Following inoculation, for all but one experiment, flasks were not opened until growth was noted. 2) Several combinations of strains and plasmids in minimal medium B reproducibly did not grow in 1000 h. 3) Strain CC1205 was transformed with pUC118 and pJK455 in two separate experiments, and in both experiments growth was noted at a similar time. 4) Most cultures grew to a similar density. If contamination were to cause the growth seen at later time points in the case of the replacement strains that contain *purE<sub>Aa</sub>* transformed with either pUC118 or pJK455, the results would still indicate that there is a problem in purine biosynthesis, and that the addition of *orf1<sub>Aa</sub>* from pJK455 does not remedy this problem.

Extended lag phases in *E. coli* are often noted as a response to environmental stresses or toxic substances (29-31). The extended lag phase seen in *purE<sub>Aa</sub>* insert strains is unusual in that it is observed in minimal media under normal conditions. A search of the literature found evidence of one other strain of *E. coli*, MH812, which exhibits a prolonged lag phase when grown in minimal media (32). This strain contains a mutation in the promoter region of the gene for dihydrofolate reductase (DHFR), *foIA*, which leads to an overproduction of DHFR. The extended lag phase in this strain is ~20 h longer than the lag phase seen with wild type *E. coli* strain MG1655, and is observed when the strain is grown at 42 °C.

*A possible cause of the lag phase.* The replacement of *purE<sub>Ec</sub>* with *purE<sub>Aa</sub>* may result in a decreased production of purines through purine biosynthesis. During the acclimation to minimal media, these *purE<sub>Aa</sub>* replacement strains may not adequately supply enough purines to support rapid replication of DNA and synthesis of RNA for newly transcribed genes, resulting in a prolonged lag phase. If histidine biosynthesis increases once cells begin to divide, this may provide enough AICAR to supplement the AICAR produced through purine biosynthesis and support rapid growth. For these reasons we suggest that the lag phase is the most reliable indicator of a problem in purine biosynthesis. However, there may be additional information that could be gleaned from strains that grow slowly after the end of the lag phase, or that grow to lower final densities.

*Metabolite channeling.* For channeling of a metabolite between two enzymes to occur, a ternary complex of the two enzymes and the metabolite must form. This means that the two enzymes must interact. This interaction may be transient and of a duration that allows for metabolite transfer.

In vivo evidence of metabolite channeling is difficult to obtain. If proteins are shown to associate in vivo, this only shows that channeling may occur. The best evidence of channeling is found in vivo experiments using the technique of isotopic dilution with either cell components (33, 34) or intact cells (35). This technique utilizes an isotopically labeled precursor, and an unlabeled (or differently labeled) intermediate. Low incorporation of the unlabeled intermediate into the final product suggests that the

intermediate formed from the precursor does not freely dissociate into solution.

Evidence for metabolite channeling can also be found if a reaction results in a symmetric intermediate (36). If an isotopically labeled precursor is converted to a symmetric intermediate, non-random scrambling of the label in the final product suggests that the intermediate is not free in solution.

The possibility that metabolite channeling may occur between PurE and PurK is an appealing proposal. This is because PurK hydrolyzes ATP to produce the acid labile  $N^5$ -CAIR. If  $N^5$ -CAIR were not used efficiently, decarboxylation of  $N^5$ -CAIR to reform the starting substrate AIR would waste energy. This is a particularly acute concern in the acidophile *A. aceti*.

When the initial functional complementation studies of PC0135 with *A. aceti* genes were published (18), great care was taken to say that the results of the experiments suggest that *AaPurK* is needed for the proper functioning of *AaPurE*. Another group interpreted those same results to imply that *AaPurE* could not accept  $N^5$ -CAIR from *EcPurK* (23). This was specifically not stated for many reasons. There are several ways an enzyme could require another enzyme to function without channeling. One possible requirement of a protein for another is that association of the two is required *in vivo* for activity without channeling. However, the main reason we did not claim to have evidence for channeling, is that strain PC0135 has a leaky phenotype. Subsequently we realized that *AaOrf1* might have a role in functional complementation. At the time results were reported it was stated that experiments were planned to

address the possible requirement of *AaPurE* for *AaPurK*. Those experiments are discussed below.

The *purE<sub>Aa</sub>* and *purE<sub>Aa</sub>purK<sub>Aa</sub>* replacement strains were created to explore the unusual finding that PC0135/pJK324 has an unexpected *purK*- phenotype. This suggested that the substrate of *AaPurE* ( $N^5$ -CAIR) might not be produced in a usable manner by *EcPurK*, and could be a sign of substrate channeling from *PurK* to *PurE*. One way this could occur is if  $N^5$ -CAIR does not freely dissociate from the enzyme in the absence of *EcPurE* after it is produced.

Singly expressed *AaPurE* has been successfully isolated as active enzyme by overexpression in *E. coli* (Chapter 3). This indicates that this enzyme should be expressed as an active protein from the genome. The fact that the CC1202 (*purE<sub>Ec</sub>*<>*purE<sub>Aa</sub>*), CC1204 (*purE<sub>Ec</sub>purK<sub>Ec</sub>*<>*purE<sub>Aa</sub>purK<sub>Aa</sub>*), and CC1205 (*purE<sub>Ec</sub>purK<sub>Ec</sub>*<>*purE<sub>Aa</sub>(ATG)purK<sub>Aa</sub>*) replacement strains grow, albeit after a prolonged lag phase, indicates the *AaPurE* is expressed as an active protein from the genomic level under normal atmosphere.

Because the *purK<sub>Aa</sub>* replacement strain CC1203 grows rapidly in minimal media, *AaPurK* is expressed as an active enzyme from the genomic level. This also indicates that *AaPurK* can serve as a source of  $N^5$ -CAIR for *EcPurE*. This could occur either by *AaPurK* releasing  $N^5$ -CAIR into solution, or by *EcPurE* being able to accept  $N^5$ -CAIR if channeling is occurring. The identity of *PurK* seems to make no difference.

Although the *purK*<sub>Aa</sub> replacement strain CC1203 does grow well, the double replacement strains that contained *purE*<sub>Aa</sub>*purK*<sub>Aa</sub> failed to grow better than the *purE*<sub>Aa</sub> single replacement strain CC1202. This means that the protein mismatch between *AaPurE* and *EcPurK* is not adequate to explain CC1202 growth deficiency as would be expected in a channeling model. While *AaPurK* provided adequate *N*<sup>5</sup>-CAIR for *EcPurE*, it failed to provide adequate *N*<sup>5</sup>-CAIR for *AaPurE*. While these results do not preclude the possibility of substrate channeling in *E. coli* PurE and PurK, they indicate that *AaPurK* can partner with either *EcPurE* or *AaPurE*. These results suggest that the biosynthesis of CAIR may be different in *A. acetii*, and they lead us to consider the possible involvement of *orf1*<sub>Aa</sub>.

*Reasons for the possible involvement of AaOrf1.* The possibility of an additional protein involved in the biosynthesis of CAIR raises several interesting possibilities. All of these possibilities reduce the potential loss of *N*<sup>5</sup>-CAIR by decarboxylation.

There are several possible roles that *AaOrf1* could serve:

(1) *AaOrf1* could serve as an *N*<sup>5</sup>-CAIR binding protein. *AaOrf1* would either bind *N*<sup>5</sup>-CAIR in solution or accept it directly from *AaPurK* and protect it from decarboxylation. *AaOrf1* would then associate with *AaPurE* to deliver *N*<sup>5</sup>-CAIR.

(2) *AaOrf1* could be an additional subunit of *AaPurE*. *AaPurE* would bind *N*<sup>5</sup>-CAIR poorly in the absence of *AaOrf1*.

(3) *AaOrf1* could bind *AaPurE* and prevent the production of CAIR in the absence of *AaPurK*. The production of CAIR may thus be dependent on a three protein complex.

(4) *AaOrf1* could bind *AaPurK* and prevent the production of  $N^5$ -CAIR in the absence of *AaPurE*. The production of  $N^5$ -CAIR may thus be dependent on a three protein complex.

(5) *AaOrf1* could serve as a scaffold for the assembly of a *AaOrf1:AaPurE:AaPurK* complex. This complex could be transient and result in the channeling of  $N^5$ -CAIR, from *AaPurK* to *AaPurE*, without a direct interaction between *AaOrf1* and  $N^5$ -CAIR. This complex could also be stable and result in the co-localization of PurE and PurK activities.

(6) *AaOrf1* is a transcriptional regulator of *AaPurE*.

Models 3 and 4 will be similar to model 5 if a three-protein complex of *AaOrf1:AaPurE:AaPurK* is required. The main difference in these models is if *AaOrf1* associates with *AaPurK* or *AaPurE* prior to complex formation

If *AaOrf1* is required for the proper function of either *AaPurE* or *AaPurK* individually (models 2-4), systems without *orf1<sub>Aa</sub>* paired with its partner would give a *pur* phenotype.

If *AaOrf1* were part of a three-protein complex with *AaPurE* and *AaPurK* (models 3-5), systems lacking any of the three would give a *pur* phenotype. In this case, the relative amounts of *AaOrf1*, *AaPurE*, and *AaPurK* may be important.

If *AaOrf1* serves as an  $N^5$ -CAIR binding protein (model 1), the absence of *AaPurE* or *AaPurK* could result in a *pur* phenotype.

If *AaOrf1* allows for a complex of PurE and PurK to form (models 3, 4 and 5), CAIR may be produced more efficiently. In this complex  $N^5$ -CAIR could be transferred



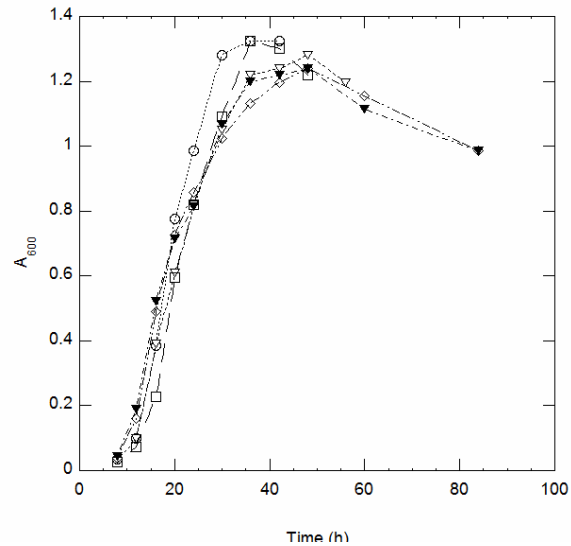
without dissociation into the cytoplasm. If *AaOrf1* binds  $N^5$ -CAIR, it could prevent decarboxylation (model 1). If *AaOrf1* prevents the activity of one component in the absence of the other, energy would not be wasted by the production of  $N^5$ -CAIR when it is not needed (models 3 and 4).

Although much emphasis has been focused on the fate of  $N^5$ -CAIR, the fate of CAIR must also be considered. PurE catalyzes the readily reversible transfer of the carboxylate of  $N^5$ -CAIR from the exocyclic amino group, to the C4 position. PurE also catalyzes the conversion of CAIR back to the more acid-labile  $N^5$ -CAIR. This effect could be more detrimental to *A. acetii* which maintains an acidic cytoplasm. If the activities of these enzymes are coupled through channeling or association (models 3 and 5), an additional benefit could occur. If *AaPurE* is only active in vivo in a complex, this would reduce the possible loss of CAIR by the back-reaction to  $N^5$ -CAIR  $\rightarrow$  AIR + CO<sub>2</sub>.

*Narrowing the possible role of AaOrf1.* The possibility that *orf1* affects purine biosynthesis led to a series of experiments with the replacement strains CC1202 (*purE<sub>Ec</sub>* $\leftrightarrow$ *purE<sub>Aa(ATG)</sub>*) and CC1205 (*purE<sub>Ec</sub>purK<sub>Ec</sub>* $\leftrightarrow$ *purE<sub>Aa(ATG)</sub>purK<sub>Aa</sub>*). For both strains, the effect that different portions of the insert of pJK173 had on their growth was assessed, along with plasmids pJK419, pJK348 and pUC118.

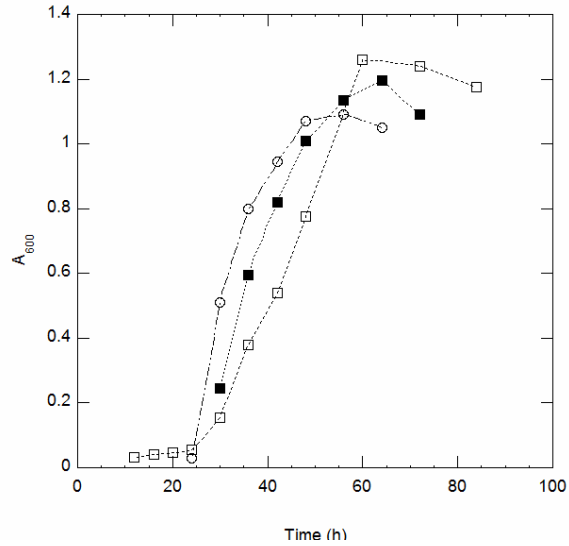
In both strains, the addition of pJK419 (*purE<sub>Ec</sub>*) restored growth similar to that of wild type strain W3110 (Figure 4.26), indicating there is a supply of  $N^5$ -CAIR adequate to support growth. The *purK<sub>Ec</sub>* $\leftrightarrow$ *purK<sub>Aa</sub>* replacement strain also grew rapidly and at a rate similar to strain W3110. Although this does not rule out any of the models, it does

indicate that the reason the replacement strains that contain *purE<sub>Aa</sub>* grow poorly is due to impaired functioning of *AaPurE* in *E. coli*.



**Figure 4.26:** Comparison of growth curves of strains with a source of *EcPurE*. CC1202 (*purE<sub>Ec</sub>* $\leftrightarrow$ *purE<sub>Aa(ATG)</sub>*)/pJK419 (entry #25, open circles), CC1205 (*purE<sub>Ec</sub>purK<sub>Ec</sub>* $\leftrightarrow$ *purE<sub>Aa(ATG)</sub>purK<sub>Aa</sub>*)/pJK419 (entry #31, open squares), CC1203(*purK<sub>Ec</sub>* $\leftrightarrow$ *purK<sub>Aa</sub>*)/pUC118 (open diamonds), W3110/pUC118 (separate experiments, open and filled downward triangles).

The effects of the addition of pJK173 will be used as a reference for the comparisons that follow. The addition of plasmid pJK173 to strains CC1202 and CC1205 gave similar growth profiles (Figure 4.27). Although this finding does not rule out any of the models, it does indicate that an increase in the products of *orf1<sub>Aa</sub>*, *purE<sub>Aa(ATG)</sub>*, and *purK<sub>Aa</sub>* compensates for the problem in replacement strains that contain *purE<sub>Aa</sub>*.



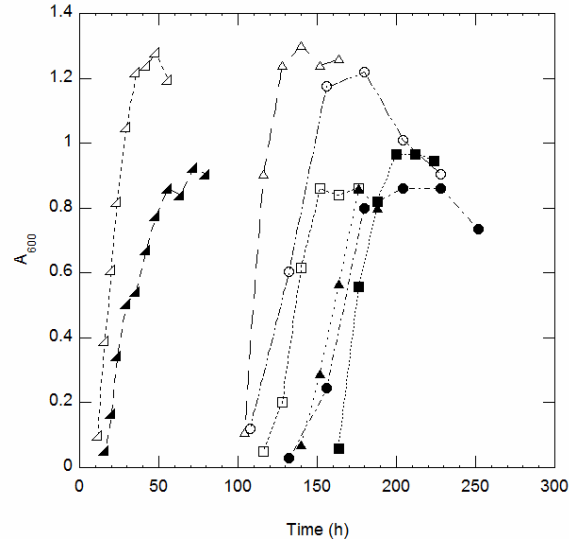
**Figure 4.27:** Comparison the effect of pJK173. The growth curves of replacement strains CC1202 (entry #23, open circles) and CC1205 (entries #33 and #39, open and filled squares respectively) transformed with pJK173 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>*).

For both strains, the addition of pJK455 (*orf1<sub>Aa</sub>*) produced an increased lag phase compared to pUC118 (Figure 4.28). Plasmid pJK455 contains the beginning portion of the insert of pJK173 including *orf1<sub>Aa</sub>* and a portion of *AaPurE* (the first 64 amino acids of *AaPurE* fused to the  $\alpha$ -LacZ fragment of pUC118). This portion of *AaPurE* is too small to contribute an activity. An increased lag phase was also seen in the double replacement strain CC1204 with pJK455 in comparison to pUC118. These results are different from the effect of pJK455 in strain W3110, which showed only a subtle growth inhibition (Figures 4.12 and 4.13).

While pJK455 slowed the growth of W3110, it did not produce an increased lag phase. This suggests that the effect of *AaOrf1* is dependent on the identity of *purE*.

While plasmid pJK455 has the same effect on strains CC1202 and CC1205, one cannot eliminate the possibility that *AaOrf1* interacts with both *AaPurE* and *AaPurK*. It is

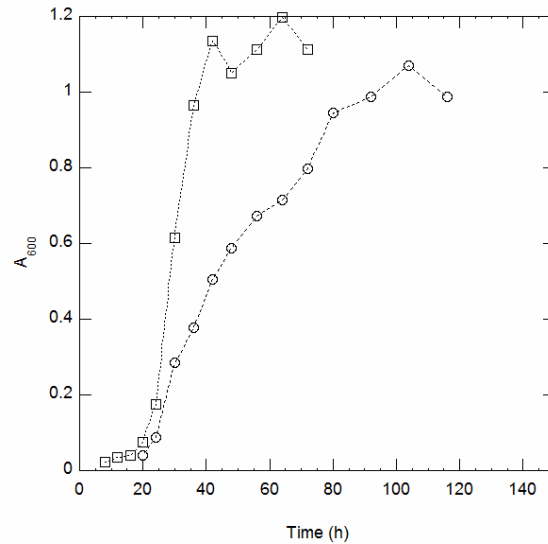
unknown what the effect of increased concentration of one product would have if association of the proteins occurs.



**Figure 4.28:** Comparison of effects of pUC118 and pJK455. Strains were transformed with pUC118 (empty vector) or pJK455 (*orf1<sub>Aa</sub>*). W3110/pJK455 (entry # 17, filled right triangles), W3110/pUC118 (entry #16, open right triangles), Strain CC1205/pJK455 (entries #32 and #40, filled circles and squares), Strain1205/pUC118 (entries #31 and #34, open circles and squares), Strain CC1202/pJK455 (entry #24, filled isosceles triangle), and Strain CC1202/pUC118 (entry #18, open isosceles triangles).

For both strains, the addition of pJK412 (*orf1<sub>Aa</sub>purE<sub>Aa</sub>*) caused a reduction in the lag phase (Figure 4.29). Although plasmid pJK412 reduced the lag phase in both strains, it had slightly different effects. In strain CC1205, which contains *purK<sub>Aa</sub>*, the growth obtained is comparable to fully complemented strain CC1205/pJK173. In strain CC1202, which contains *purK<sub>Ec</sub>*, slower growth was noted than in the fully complemented CC1205/pJK173. This indicates the effect of *AaOrf1* depends on the identity of *purK* but to a lesser degree than on the identity of *purE*. These results coupled with the results of pJK455 indicate that the effect of *AaOrf1* depends upon the identity of both *purE* and

*purK*. This is inconsistent with model 2, in which *AaOrf1* is an additional subunit of *AaPurE* (model 2). This model is also inconsistent with the observation that pure *AaPurE* isolated from *E. coli* is as active as *EcPurE*.

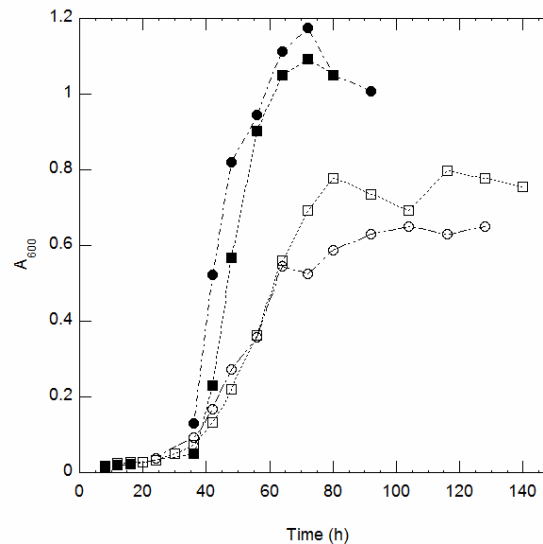


**Figure 4.29:** Comparison of the effect of pJK412. Shown are the growth curves of replacement strains CC1202 (entry #20, open circles) and CC1205 (entry #36, open squares) transformed with pJK412.

The addition of pJK415 (*orf1<sub>AaPurK<sub>Aa</sub></sub>*) had a similar effect in both single- and double-replacement strains (Figure 4.30). Although the lag phase was reduced, both strains grew slowly. A consideration of these results is included with those obtained using pJK348 (below).

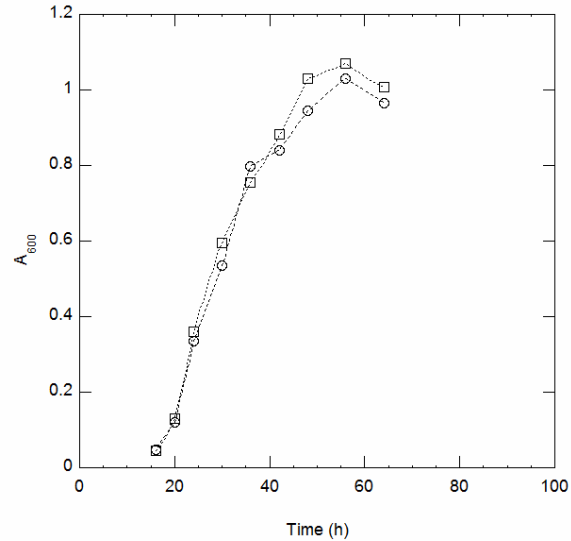
In both strain CC1202 and CC1205, the addition of plasmid pJK348 (*orf1<sub>AaPurEH59N<sub>AaPurK<sub>Aa</sub></sub></sub>*) had a similar effect (Figure 4.30). The lag phase was reduced, and both strains grew rapidly. The growth was similar to growth curves obtained using the fully-complemented strains CC1202/pJK173 and CC1205/pJK173. If *AaPurK* activity were inhibited by *AaOrf1* in the absence of *AaPurE*, the production of the inactive but

properly folded *AaPurEH59N* mutant would not be expected to improve growth. The effect of pJK348 on both insert strains relative to pJK415 argues against model 4 in which *AaOrf1* serves to prevent  $N^5$ -CAIR production by *AaPurK* in the absence of *AaPurE*. Instead the results obtained with these plasmids suggest that similar expression of all three components is beneficial.



**Figure 4.30:** Comparison of the effects of pJK415 and pJK348. Shown are the growth curves of replacement strains in minimal medium B. CC1202/pJK415 (entry #21, open circles), CC1202/pJK348 (entry #22, filled circles), CC1205/pJK415 (entry #37, open squares), CC1205/pJK348 (entry #39, filled squares).

In both replacement strains, the addition of plasmid pJK324 (*purE<sub>Aa</sub>*) improved growth at levels similar to pJK173 (Figure 4.31). This indicates that increased levels of *AaPurE* relative to *AaPurK* may compensate for the absence of *AaOrf1*. These results do not distinguish between models 3 and 5.



**Figure 4.31:** Comparison of the effect of pJK324. Shown are the growth curves of replacement strains CC1202 (entry#19, open circles) and CC1205 (entry #35, open squares) in minimal medium B.

*Orf1.* A search of the *AaOrf1* protein using blastp (37) retrieved 21 hits with an E value of  $\leq 1 \times 10^{-6}$ . All of the hits were annotated as small hypothetical proteins or proteins of unknown function (DUF 465). These hits were found throughout much of the class  $\alpha$ -proteobacteria and with two hits from the  $\gamma$ -proteobacteria genus *Thioalkalivibrio*.

In most cases, the small hypothetical protein is immediately upstream or within one gene of the *purEpurK* genes in these organisms. Among the 19  $\alpha$ -proteobacteria, only four were found with a different synteny; in *Erythrobacter sp SD-21*, *orf1* precedes a probable AraC family transcriptional regulator; in two strains of *Magnetospirillum magnetotacticum*, *orf1* precedes a gene involved in the biosynthesis of tetrahydrofolate (THF), 7,8-dihydroxymethylpterin pyrophosphokinase; and in *Methylocella silvestris BL2*, *orf1* is found divergently oriented from a gene for phosphatidylethanolamine N-

methyltransferase. In the  $\gamma$ -proteobacteria genus *Thioalkalivibrio*, *orf1* precedes a gene that encodes a FAD-dependent pyridine nucleotide-disulfide oxidoreductase.

*AaOrf1* was aligned using ClustalW with BLAST for sequences that have *orf1* in the *purE-purK* cluster (Figure 4.32). Only 12 of the 14 proteins are shown because two entries were subspecies. This alignment identifies three highly conserved regions: residues 16-23 (*A. aceti* numbering), RxEHRDLD; residues 43-48, RLKKxKL; residues 64-68, PDxIA. There are also three other conserved residues: Leu12, Asp53 and Glu58. All conserved residues are shown in bold in Figure 4.32.



**Figure 4.32:** Sequence alignment of Orf1 from *A. aceti* and 12 close homologs.

```

Acetobacter_aceti_orf1          -----MLR-DRDLLLAQLHELRLSEHRDLDT 24
Granulibacter_bethesdensis_orf1 -----MSFQAMLK-DRDALLRQLHELRLSEHRDLDT 29
Gluconobacter_oxydans_orf1     -----MMT-DDNAMLKRLHELRLSEHRDLDT 24
Rhodospirillum_rubrum_orf1    -----MDDHELDGLRAKLEELRTEHRDLDA 25
Acidiphilium_cryptum_orf1     -----MLT-DKDNLLRQLHGLRLSEHRDLDS 24
Stappia_aggregata_orf1        -----MQLAQLRQEHRDLDA 15
Roseovarius_TM1035_orf1       -----MNAFSDLSMKSEDVLRVELEEFRRREHRDLDE 31
Roseovarius_217_orf1          -----MNAYSDL SMKSEDVLRVELEEFRRREHRDLDE 31
Dinoroseobacter_shibae_orf1   -----MGMNGTGAMSR--EEVLKYELEVLRRREHRDLDE 31
Gluconacetobacter_diazotrophic_orf1 -----MLT-DRDTLLRKLHELRLSEHRDLDT 24
Methylobacterium_radiotolerans_orf1 -MLSGDERSSSGSQVTMADEAGEGAQSDLLGELARLREHRDLDS 43
Beijerinckia_indica_orf1     -----MVDRLSEDERISLQTELEHLRQEHRDLDA 29
Rhodobacter_sphaeroides_orf1  -----MNASPELSF--EDMLRIRLEVLRRREHRDLDE 29
                                     . *   : *   * * * * *

Acetobacter_aceti_orf1          VINRMAHETTFIDQLYLQRLKKRKL L L L K D Q I T K V E S L L I P D D I A 68
Granulibacter_bethesdensis_orf1 VISRLADQGA-LDQLQLQRLKKRKL L I K D E V S R L E S N L I P D N I A 72
Gluconobacter_oxydans_orf1     VIERLVHH-P-LNQLQLQRLKKRKL Q L K D E I S W I E T R L I P D N I A 66
Rhodospirillum_rubrum_orf1    VIARITENMP-FDMIQMQRLLKKRKL A L K D Q I S R L E N R M I P D I I A 68
Acidiphilium_cryptum_orf1     VIARLGDQPT-IDQLQIQRLKKRKL L L L R D Q I L R L E S R L I P D S I A 67
Stappia_aggregata_orf1        AVEALASTSN-QDALQLQRLKKK L M I K D R I T A L E D Q L F P D I I A 58
Roseovarius_TM1035_orf1       AIRALQDKGT-ADQLMIQRLKKK K L W L R D M I A R I E D R L Y P D I I A 74
Roseovarius_217_orf1          AIRALQDKGT-ADQLMVQRLKKK K L W L R D M I A R I E D R L Y P D I I A 74
Dinoroseobacter_shibae_orf1   AIEALGERPA-PDQLTLKRLKKRKL A L K D Q I A R I E D E L F P D I I A 74
Gluconacetobacter_diazotrophic_orf1 VISRLALH-P-MDQLQLQRLKKRKL L L L K D E I A W L E S R L I P D N I A 66
Methylobacterium_radiotolerans_orf1 AIEALERSVA-GDQLQIQRLKKRKL T L R D R I F H I E D A L T P D I I A 86
Beijerinckia_indica_orf1     AIEALLHLST-TDRLQVQRLKKRKL V L R D R I V F I E D L L T P D I I A 72
Rhodobacter_sphaeroides_orf1  AIAAIEAGGR-GDQLMLRRLKKQKL A L K D Q I V K I E D R L I P D I I A 72
                                     . :   :           : : : * * * * * : : *   : *   : * * * *

```

Because *AaOrf1* is proposed to interact with *AaPurE* and *AaPurK* (models 3 and 5), these proteins from the *orf1*-containing organisms were also aligned and compared to *EcPurE*.

Sequence alignments showed no obvious conserved differences from *EcPurE*, except that *EcPurE* Lys50 is an arginine in *orf1*-containing organisms (*AaPurE* Arg64). In *A. aceti*, PurE contains an additional ~20 N-terminal residues not found in *EcPurE* (38). Although these residues would appear to be an important difference, they are not present in the other PurEs from *orf1*-containing organisms.

Sequence alignments of PurK from *orf1*-containing organisms indicated three conserved differences from *EcPurK*. Two of these differences are amino acid substitutions. Residues Asp268 and Asp306 in *E. coli* PurK are replaced with Gly in all PurKs from *orf1*-containing organisms. The third difference between the PurKs from *orf1*-containing organisms and *EcPurK* is they all contained ~20 amino acids inserted between residues Ser44 and Val45 in *EcPurK* (Figure 4.33).

Valine 45 in the *EcPurK* crystal structure (pdb 1B6S) is the beginning of a third strand of a  $\beta$ -sheet (39). The sequence of *AaPurK* was submitted to Jpred3 (40), and the 20 amino acid insert was predicted to have a small  $\beta$ -strand comprised of residues Thr56 and Cys 57 followed by an  $\alpha$ -helix comprised of residues Pro 63- Arg71.

**Figure 4.33:** Sequence alignment of PurKs

```

Acetobacter_Aceti_PurK          PEANGPAAQVSHAVTCGKYDD-----PTALDAFARAVDVVTFEFENISAD 86
Granulibacter_bethesdensis_Pur DAPDSPAIQVSAAHTIGAYDD-----PQALRAFANAVDVVTFEFENVSAE 83
Gluconobacter_oxydans_PurK      TEAPSPATEVAASVTVGAYDD-----PAALEDFASRCDVVTTFEFENISAE 84
Rhodospirillum_rubrum_ATCC_Pur PEDNGPASQVCPMVTLLAAYDD-----LDALSRFAAASVDVITTFEFENIPAA 91
Acidiphilium_cryptum_PurK      PEPDAPASQVAAATTRADYDD-----NAALLRFADAVDVITTFEFENVSAE 84
Stappia_aggregata_PurK         PDPNSPAFDVSAIFTVAPYED-----IDALDRFASSVAAVTYEFENVPGP 85
Roseovarius_TM1035_PurK        PGGDCPASHVAHQHIQADYTD-----EDALRRFAGIVDVITYEFENIPTA 83
Roseovarius_217_PurK          PGGDCPASHVAHRHIQAEYS-----EDALRRFAEAVDVITYEFENIPTT 83
Dinoroseobacter_shibae_PurK    PGAAPPAGQVAEAVTTAGYDD-----LDALRRFAEVVDVITYEFENIPTA 83
Gluconacetobacter_diazotrophic_P DEAEGPAAQVAHAVTVGAYDD-----PDVLRRAFASSVDVVTTFEFENISAD 87
Methylobacterium_radiotolerans_P PDADSPAFDVAARTTCAAYDD-----AAALADFARSVDVVTYEFENIPHA 84
Beijerinckia_indica_ATCC_PurK  PEADSPAFAVVAAGHMLAAYED-----EEALAQFAAQVDVITYEFENVPAR 83
Rhodobacter_sphaeroides_PurK   PSANPPAADVAHAVTTAPYED-----EAALRAFATSVDVITYEFENIPTS 83
Escherichia_coli_PurK         LDAEPAAVPFQQS-----VITAEIERWPET 56
Bacillus_subtilis_PurK        PVKDSPCGQVADVEITAHYND-----REAIRKLAEISDIITYEFENIDYD 84
Cryptococcus_neoformans_Ade2   SGSYTPAKQTLLPPPHSHPDGPFTSETHIRKLASACDILTVEIEHVNAD 84
                                ..                                :*  *:*

```

While this insertion is not found in *EcPurK* or any annotated PurK in the order Enterobacteriales, other organisms that lack *orf1* contain a ~20 amino acid insertion in this region of PurK. This insert is found in PurK from Firmicutes bacteria. The enzyme Ade2, which is a bifunctional protein that contains the PurE and PurK activities in fungi, also contains a ~20 amino acid insert in this region that is slightly larger. For examples of the inserts contained in other organisms, *B. subtilis* and *C. neoformans* were chosen. This is because *B. subtilis* is a well studied bacterium with a characterized purine pathway, and the enzyme Ade2 from *C. neoformans* has been characterized (41). The alignment indicates only two conserved residue in all of the ~20 amino acid inserts corresponding to *AaPurK* Asp62 and Ala70. Because this ~20 amino acid insert is present in other organisms, it is unknown if this insert might influence the function of *AaOrf1*.

*Summary.* Whether *A. acetii* Orf1 is actually a new protein involved in the function of PurE and PurK is not proven by these results. However, it does appear likely that it is a new protein involved in purine biosynthesis. The conserved synteny of the gene and the conserved residues present in the protein, along with the results of the replacement strains, indicate has an effect that is dependent on the identity of both PurE and PurK. One exciting possibility is that Orf1 is a channeling factor (model 5) The possibility of the involvement of another protein in the activity of PurE and PurK raises many interesting possibilities for the regulation and control of a set of reactions with labile products.

#### 4.5 Future Directions

Under the current proposal, *AaOrf1* appears to be required for the proper functioning of *AaPurE* and *AaPurK* in *E. coli*. In order to better determine the role of *orf1*, two routes present themselves. An in-frame deletion of *orf1<sub>Aa</sub>* in the host organism would be most useful. Unfortunately *A. aceti* has difficult genetics, and no system is available to create such deletions. While insertional inactivation methodology is available for some acetic acid bacteria (42-44), any such insertion will have unknown polar effects on the adjacent *purE-purK* genes. We predict that a  $\Delta$ *orf1* strain would require exogenous purines for rapid growth. We expect that an *orf1<sub>Aa</sub>purE<sub>Aa</sub>purK<sub>Aa</sub>* replacement strain in *E. coli* should grow on minimal media without a prolonged lag phase.

## 4.6 References

1. Yanofsky, C., and Ito, J. (1966) Nonsense codons and polarity in the tryptophan operon, *Journal of Molecular Biology* 21, 313-334.
2. Oppenheim, D. S., and Yanofsky, C. (1980) Translational coupling during expression of the tryptophan operon of *Escherichia coli*, *Genetics* 95, 785-795.
3. Imamoto, F., and Yanofsky, C. (1967) Transcription of the tryptophan operon in polarity mutants of *Escherichia coli*: I. Characterization of the tryptophan messenger RNA of polar mutants, *Journal of Molecular Biology* 28, 1-23.
4. Imamoto, F., and Yanofsky, C. (1967) Transcription of the tryptophan operon in polarity mutants of *Escherichia coli*: II. Evidence for normal production of *tryp*-mRNA molecules and for premature termination of transcription, *Journal of Molecular Biology* 28, 25-35.
5. Ames, B. N., and Hartman, P. E. The Histidine Operon, *Cold Spring Harbor Symposia on Quantitative Biology* 28, 349-356.
6. Nagy, P. L., McCorkle, G. M., and Zalkin, H. (1993) PurU, a source of formate for PurT-dependent phosphoribosyl-N-formylglycinamide synthesis, *Journal of Bacteriology* 175, 7066-7073.
7. Saxild, H. H., and Nygaard, P. (2000) The *yexA* gene product is required for phosphoribosylformylglycinamide synthetase activity in *Bacillus subtilis*, *Microbiology* 146, 807-814.
8. Hoskins, A. A., Anand, R., Ealick, S. E., and Stubbe, J. (2004) The formylglycinamide ribonucleotide amidotransferase complex from *Bacillus subtilis*: Metabolite-mediated complex formation, *Biochemistry* 43, 10314-10327.
9. Morar, M., Hoskins, A. A., Stubbe, J., and Ealick, S. E. (2008) Formylglycinamide ribonucleotide amidotransferase from *Thermotoga maritima*: Structural insights into complex formation, *Biochemistry* 47, 7816-7830.
10. Ebbole, D., and Zalkin, H. (1987) Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* encoding nine enzymes for de novo purine nucleotide synthesis, *Journal of Biological Chemistry* 262, 8274-8287.
11. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994)  $N^5$ -carboxyaminoimidazole ribonucleotide: Evidence for a new intermediate and two new enzymic activities in the de novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry* 33, 2269-2278.
12. Gots, J. S., Bensen, C. E., Jochimsen B., and Koduri, K. R. (1977) CIBA Foundation Symposium, in *Purine and pyrimidine metabolism*, pp 23-41.
13. Church, G. Lab webpage, <http://arep.med.harvard.edu/labgc/pko3.html>. Gene replacement using pKO vectors. Last updated October 26 2004
14. Link, A. J., Phillips, D. and Church, G. M. (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: Application to open reading frame characterization, *Journal of Bacteriology* 179, 6228-6237.
15. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proceedings of the National Academy of Sciences of the United States of America* 97, 6640-6645.
16. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction, *Gene* 77, 51-59.

17. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension, *Gene* 77, 61-68.
18. Constantine, C. Z., Starks, C. M., Mill, C. P., Ransome, A. E., Karpowicz, S. J., Francois, J. A., Goodman, R. A., and Kappock, T. J. (2006) Biochemical and structural studies of  $N^5$ -carboxyaminoimidazole ribonucleotide mutase from the acidophilic bacterium *Acetobacter aceti*, *Biochemistry* 45, 8193-8208.
19. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) Evidence for the direct transfer of the carboxylate of  $N^5$ -carboxyaminoimidazole ribonucleotide ( $N^5$ -CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by *Escherichia coli* PurE, an  $N^5$ -CAIR Mutase, *Biochemistry* 38, 3012-3018.
20. Schrimsher, J. L., Schendel, F. J., Stubbe, J., and Smith, J. M. (1986) Purification and characterization of aminoimidazole ribonucleotide synthetase from *Escherichia coli*, *Biochemistry* 25, 4366-4371.
21. Maniatis, T., and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.
22. Meyer, E., Leonard, N. J., Bhat, B., Stubbe, J., and Smith, J. M. (1992) Purification and characterization of the *purE*, *purK*, and *purC* gene products: identification of a previously unrecognized energy requirement in the purine biosynthetic pathway, *Biochemistry* 31, 5022-5032.
23. Hoskins, A. A., Morar, M., Kappock, T. J., Mathews, I. I., Zaugg, J. B., Barder, T. E., Peng, P., Okamoto, A., Ealick, S. E., and Stubbe, J. (2007)  $N^5$ -CAIR mutase: Role of a CO<sub>2</sub> binding site and substrate movement in catalysis, *Biochemistry* 46, 2842-2855.
24. Stouthamer A H, d. H. P., Nijkamp HJ. (1965) Mapping of purine markers in *Escherichia coli*, *Genetics Research*, 442-453.
25. Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes, *Gene* 234, 187-208.
26. Estramareix, B., and Therisod, M. (1984) Biosynthesis of thiamin: 5-aminoimidazole ribotide as the precursor of all the carbon atoms of the pyrimidine moiety, *Journal of the American Chemical Society* 106, 3857-3860.
27. Lawhorn, B. G., Mehl, R. A., and Begley, T. P. (2004) Biosynthesis of the thiamin pyrimidine: the reconstitution of a remarkable rearrangement reaction, *Organic & Biomolecular Chemistry* 2, 2538-2546.
28. Nagy, P., Marolewski, A., Benkovic, S., and Zalkin, H. (1995) Formyltetrahydrofolate hydrolase, a regulatory enzyme that functions to balance pools of tetrahydrofolate and one-carbon tetrahydrofolate adducts in *Escherichia coli*, *Journal of Bacteriology* 177, 1292-1298.
29. Sugimoto, S., Nakayama, J., Fukuda, D., Sonezaki, S., Wantanabe, M., Tosukhowong, A., and Sonomoto, K. (2003) Effect of heterologous expression of molecular chaperone DnaK from *Tetragenococcus halophilus* on salinity adaptation of *Escherichia coli*, *Journal of Bioscience and Bioengineering* 96, 129-133.
30. Shapira, J., and Dittmer, K. (1961) Unsaturated amino acids V: Microbiological properties of some halogenated olefinic amino acids, *Journal of Bacteriology* 82, 640-647.
31. Takahashi, C., Takahashi, D., Carvalhal, M., and Alterthum, F. (1999) Effects of acetate on the growth and fermentation performance of *Escherichia coli* KO11, *Applied Biochemistry and Biotechnology* 81, 193-203.

32. Herrington, M. B., MacRae, T. J., Panagopoulos, D., and Wong, S.-H. W. (2002) A mutation in the *folA* promoter delays adaptation to minimal medium by *Escherichia coli* K-12, *Journal of Basic Microbiology* 42, 172-180.
33. Graham, J. W. A., Williams, T. C. R., Morgan, M., Fernie, A. R., Ratcliffe, R. G., and Sweetlove, L. J. (2007) Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling, *Plant Cell* 19, 3723-3738.
34. Czichi, U., and Kindl, H. (1977) Phenylalanine ammonia lyase and cinnamic acid hydroxylases as assembled consecutive enzymes on microsomal membranes of cucumber cotyledons: Cooperation and subcellular distribution, *Planta* 134, 133-143.
35. Shearer, G., Lee, J. C., Koo, J., and Kohl, D. H. (2005) Quantitative estimation of channeling from early glycolytic intermediates to CO<sub>2</sub> in intact *Escherichia coli*, *FEBS Journal* 272, 3260-3269.
36. Sumegi, B., Sherry, A. D., Malloy, C. R., and Srere, P. A. (1993) Evidence for orientation-conserved transfer in the TCA cycle in *Saccharomyces cerevisiae*: carbon-13 NMR studies, *Biochemistry* 32, 12725-12729.
37. Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Research* 25, 3389-3402.
38. Settembre, E. C., Chittuluru, J. R., Mill, C. P., Kappock, T. J., and Ealick, S. E. (2004) Acidophilic adaptations in the structure of *Acetobacter aceti* N<sup>5</sup>-carboxyaminoimidazole ribonucleotide mutase (PurE), *Acta Crystallographica Section D: Biological Crystallography* D60, 1753-1760.
39. Thoden, J. B., Kappock, T. J., Stubbe, J., and Holden, H. M. (1999) Three-dimensional structure of N<sup>5</sup>-carboxyaminoimidazole ribonucleotide synthetase: A member of the ATP grasp protein superfamily, *Biochemistry* 38, 15480-15492.
40. Cole, C., Barber, J. D., and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server, *Nucleic Acids Research* 36, W197-201.
41. Firestine, S. M., Misialek, S., Toffaletti, D. L., Klem, T. J., Perfect, J. R., and Davisson, V. J. (1998) Biochemical role of the *Cryptococcus neoformans* ADE2 protein in fungal de novo purine biosynthesis, *Archives of Biochemistry and Biophysics* 351, 123-134.
42. Holscher, T., and Gorisch, H. (2006) Knockout and overexpression of pyrroloquinoline quinone biosynthetic genes in *Gluconobacter oxydans* 621H, *Journal of Bacteriology* 188, 7668-7676.
43. Katzen, F., Becker, A., Ielmini, M. V., Oddo, C. G., and Ielpi, L. (1999) New mobilizable vectors suitable for gene replacement in gram-negative bacteria and their use in mapping of the 3' end of the *Xanthomonas campestris* pv. *campestris* gum operon, *Applied and Environmental Microbiology* 65, 278-282.
44. Fukaya, M., Takemura, H., Okumura, H., Kawamura, Y., Horinouchi, S., and Beppu, T. (1990) Cloning of genes responsible for acetic acid resistance in *Acetobacter aceti*, *Journal of Bacteriology* 172, 2096-2104.



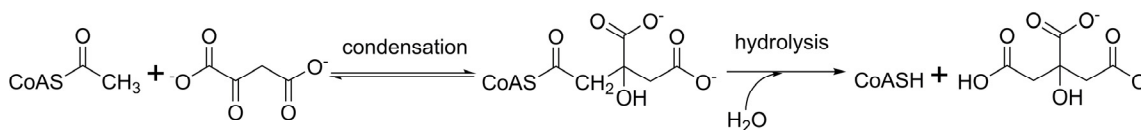


## **Appendix I**

### Citrate Synthase

## I.1 Introduction

*Citrate synthase.* Citrate synthase catalyzes the formation of citrate and coenzyme A (CoA) from oxaloacetic acid (OAA) and acetyl CoA (AcCoA). This is the entry point of AcCoA into the citric acid cycle (CAC). The rest of the CAC produces two molecules of CO<sub>2</sub> from the two carbons introduced by citrate synthase and regenerates OAA. The CAC usually yields energy for the cell in the form of one nucleotide triphosphate (ATP or GTP) and four reduced cofactors which can be used in oxidative phosphorylation to produce ATP. In *A. acetii* the CAC is also involved in acetate removal (1).



**Scheme I.1:** Reaction of citrate synthase

Citrate synthase first catalyzes the formation of citryl-CoA (CitCoA) (Scheme I.1). To accomplish this, CS catalyzes an aldol-Claisen condensation of AcCoA and OAA to yield CitCoA, likely through an AcCoA enolate intermediate (2, 3). CS then catalyzes the subsequent hydrolysis of the thioester of CitCoA to yield citrate, H<sup>+</sup>, and CoA.

When unbound by ligands, citrate synthase is found in an open conformation with the active site exposed to solvent. Upon ligand binding, conformational shifts completely sequester the substrate OAA. Citrate synthase (CS) destabilizes the sequestered OAA via polarization of the carbonyl oxygen, which is stabilized by a hydrogen bond with active site residue His320 (pig CS numbering). This polarization is

evidenced by a change in the NMR chemical shift of the carbonyl carbon and by infrared spectroscopy indicating a diminution of double bond character (4, 5). This polarization prepares the carbonyl carbon for nucleophilic attack. Residue Asp375 of pig CS catalyzes the deprotonation of the AcCoA methyl group, leading to an intermediate of enolate character that interacts through one of the acetyl oxygens with His 274 (2, 3, 6). Following the nucleophilic attack on the polarized carbonyl carbon of OAA by the AcCoA enolate, CitCoA is formed. The same active site catalyzes the subsequent hydrolysis of CitCoA to yield the product citrate, CoA, and a proton.

Crystal structures, mutagenesis studies, and substrate/transition state analogs have yielded a great volume of information about the reaction and mechanism of CS (3-5, 7-13). In particular, mutation of the active site base Asp375 greatly diminishes catalytic activity. However, the identification and detailed study of the kinetics steps involved have proven difficult by virtue of the fact that in most CSs no step is clearly rate limiting (13). This however is not the case for *Thermoplasma acidophilum* citrate synthase (*TpCS*), in which hydrolysis of the CitCoA thioester has been shown to be rate limiting. *TpCS* has also been shown to undergo several changes in fluorescence upon the binding of substrates or substrate analogues (10, 12). The *TpCS* equivalent of pig CS Asp375 is Asp317. In recent unpublished studies, Hong Jiang of this lab has noted that a key mutant (*TpCS*-D317G) undergoes at least a 2 step binding of the substrate analogue CMCoA. Experiments in the presence of glycerol or sucrose indicated that the rates of these steps are independent of viscosity (changes in the viscosity of the buffer by the addition of viscosogens, glycerol and sucrose, can help identify steps which are dependent

on conformational changes of a protein). A triple mutant (TM) form of *TpCS*, which is catalytically active, has only a single Trp residue and much simpler fluorescence responses. Recent unpublished studies by a collaborator have noted that TM also undergoes at least a 2 step binding of CMC<sub>o</sub>A; however, the rate of one of these steps is highly dependent on the viscosity of the buffer. The difference in viscosity-dependent steps suggests that there may be at least a third step involved in the binding of the AcCoA analogue in catalytically active enzymes that is unobservable by fluorescence. The availability of TM, a large number of substrate and transition state analogues (7, 11), and a detailed understanding of the functions of several active site residues (11, 14), suggest that *TpCS* is an appropriate form of CS for detailed characterization of CS reaction kinetics.

Herein is reported the determination of binding constants for *TpCS*-D317G, and the creation and purification of a mutant TM-D317G made to help elucidate the kinetic mechanism of *TpCS*.

## **I.2 Methods**

*Materials* – All materials were from Sigma Aldrich or Fisher Scientific and of the highest purity unless otherwise noted. Mutagenesis was performed using Quikchange Mutagenesis kits from Stratagene. Oligodeoxynucleotides (ODNs) were from Integrated DNA Technologies. PD-10 disposable columns were obtained from GE Health Care (formerly Amersham Biosciences). Centrifugation steps were performed using a Beckman Avanti J-20 centrifuge with a JLA-10.5 or JA-20 rotor. Routine activity assays

were performed on an Agilent 8453 diode array UV/vis spectrophotometer. Circular dichroism spectra were recorded on a Jasco J715 instrument. Cell disruption by sonication was performed using a Virsonic 100 sonicator. Proteins were routinely concentrated using Amicon ultrafiltration devices with a 10 kD molecular weight cutoff (YM-10). AcCoA analogues were prepared by Hong Jiang and Jung Park. Plasmids were sequenced by Sanger sequencing at the Protein and Nucleic Acid Chemistry Laboratory at Washington University in St. Louis. Protein masses were determined by ESI-MS on a Micromass Q-TOF Ultima quadropole-TOF mass spectrometer by Washington University Center for Biomedical and Bioorganic Mass Spectrometry.

*Mutagenesis of TM-D317G*- The plasmid used for expression of TM was a gift from the lab of Linda Kurz (12) and was given the designation pJK393. Plasmid pJK383 (TM-D317G) was created by Quickchange mutagenesis of plasmid pJK393 using and oligodeoxynucleotides 982 (5'-TCCGAACACGGgTTACTTCTCC) and 983 (5'-GGAGAAGTAAcCCGTGTTTCGGA).

*Purification of TM and TM-D317G*- TM was purified by a published method (12), except that streptomycin was used in place of polyamine-p. The protein-containing solute was brought to 1% streptomycin from a 10% stock, following a 15 minute incubation on ice, solids were removed by centrifugation. TM-D317G was purified using the TM procedure except the final column, which was replaced with a Cibracon Blue 3-GA agarose column (20 mL, 2.5 × 4cm, 250 drop fractions) equilibrated in 20 mM Tris pH 8.0, 1 mM EDTA. The column was washed with 10 column volumes (CV) (200 mL) of buffer, and the protein was eluted with 50 mM ammonium sulfate, 20 mM Tris pH 8.0, 1

mM EDTA, fractions containing *TpCS* were pooled and concentrated using an Amicon ultra filtration device to a volume of 2.5 mL. The protein solution was applied to a disposable PD-10 column previously equilibrated in 20 mM Tris 1 mM EDTA pH 8.0, to remove salts and small molecules, and the protein eluted with 3.5 mL of buffer. The protein solution was concentrated by use of a new Amicon ultra filtration device. A portion was saved for immediate use and stored at 4 °C, and the remainder of the purified protein stored as an 85% ammonium sulfate suspension at 4 °C for future use.

*OAA titration of TMD317G.* Titrations were performed in a starting solution of 2.0 mL, 50 mM EPPS 0.1 mM EDTA pH 8.0 and 0.2 μM TM-D317G at 20 °C. The emission spectra were recorded from 300 to 420 nm (5 nm slit width, 0.5 s integration, 1 nm intervals), exciting at 290 nm (1 nm slit width). The intensity of  $\lambda_{\max}$  of the unliganded protein was then measured after small additions from a solution of 20 μM OAA, in 50 mM EPPS, 0.1 mM EDTA pH 8.0. Titrant solution was made from a standardized stock solution of OAA made in 50 mM EPPS, 0.1 mM EDTA at pH 8.5. Fluorescence data ( $F$ ) were fit as a function of  $C$ = [titrant] to four parameters using equation 1 below:  $K$  = dissociation constant,  $E$ = concentration of enzyme,  $F_0 = F$  prior to addition of titrant, and  $\Delta F$  = maximal change in  $F$ .

$$F = F_0 - \frac{\Delta F}{2} \left[ (C + E + K) - \sqrt{(C + E + K)^2 - 4EC} \right] \quad (1)$$

*Titration of TpCS-D317G and TpCS-D317N with AcCoA analogues.* Titrations were performed in a starting solution of 25 mL 50 mM EPPS 0.1 mM EDTA, 5 μM OAA, and 0.5 μM [subunit] at 20 °C in a 10 cm path length cell. Titrations were also performed in 20% glycerol or 30% sucrose as noted. The circular dichroism of the sample was

measured 5 times from 300 to 250 nm, (1 nm step intervals, 10 nm/min acquisition, 500  $\mu$ M slit width, 2 nm bandwidth, 10 mdeg sensitivity). Following all additions, the solution was mixed by gentle rocking and the cell placed in the cell holder. The cell was allowed to equilibrate for 10 min. The ellipticity of the sample was recorded after small additions from a 1.5 mM stock solution of the titrant (CMX or CMCoA). The ellipticity of the sample at 260 nm ( $\theta_{260}$ ) was fit as a function of C = [titrant] concentration, to equation 2 below, with the known constants,  $\theta_{i260}$  = initial ellipticity at zero titrant, E= [subunit], and variables,  $K_d$  = dissociation constant, and  $\Delta\theta_{260}$  = maximal change in  $\theta_{260}$ .

$$\theta_{260} = \theta_{i260} - \frac{\Delta\theta_{260}}{2} \left[ (C + E + K_d) - \sqrt{(C + E + K_d)^2 - 4EC} \right] \quad (2)$$

*Titration of TM-D317G with CMCoA.* Titrations were performed in a starting solution of 2.0 mL 50 mM EPPS 0.1 mM EDTA, 100  $\mu$ M OAA, and 8.98  $\mu$ M [subunit] at 20°C in a 1 cm path length cuvette. The circular dichroism of the sample was recorded from 300 to 250 nm, (1 nm step intervals, 2 nm/min acquisition, 500  $\mu$ M slit width, 2 nm bandwidth, 10 mdeg sensitivity). Following all additions, the solution was mixed gently with a stirrer 20 times, so as not to disturb the cuvette, and the solution allowed to equilibrate for 5 min prior to the next spectrum being recorded. The ellipticity of the sample was recorded after small additions from a 1.0 mM stock solution of the CMCoA. Following a small linear correction to account for the ellipticity of CMCoA, the ellipticity of the sample at 260 nm ( $\theta_{260}$ ) was fit to equation 2.

*Quantum yield determination for TM-D317G* - Values were determined using a previously published method (12).

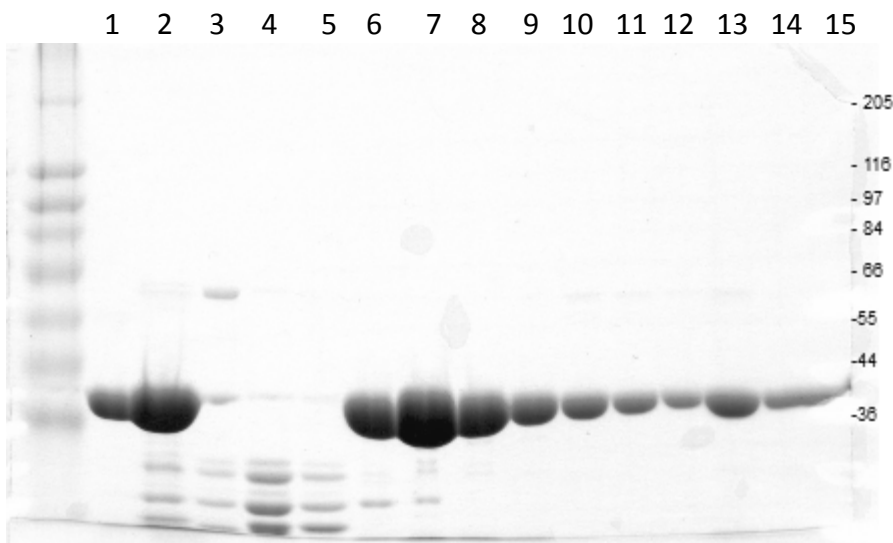


*Oligomerization state of TM-D317G-* Initial titrations of TM-D317G with OAA or CMCoA suggested that TM-D317G may purify with partial occupation of the OAA. To examine this, TM-D317G was analyzed by gel filtration. TM-D317G (2.5 mg) was run on a 120 mL 16/60 Superdex gel filtration column, (1 mL/min flow rate, 20 mM Tris 1 mM EDTA 50 mM KCL buffer). Absorbance of the eluant was monitored as a function of volume. The experiment was repeated including 500  $\mu$ M OAA in the buffer and the elution profiles were compared. BIO-RAD gel filtration standards (thyroglobulin 670,000 Da, bovine gamma globulin 158000 Da, chicken ovalbumin 44,000 Da, equine myoglobin 17,000 Da, Vitamin B<sub>12</sub> 1,350 Da) were run under the initial conditions to generate a calibration curve, which was used to determine the molecular weights of the peaks.

*Stopped flow fluorescence.* Experiments were performed in filtered 50 mM EPPS 0.1 mM EDTA at pHs 7.0, 8.0 or 8.9. Transients were collected using the following parameters: 305 nm cut off filter, 20.0 °C, 1000 points, 285 nm excitation, 2 nm slits, PM voltage = 400 V. Solutions in each syringe contained 200  $\mu$ M OAA. The following concentrations listed for components present in only one syringe, refer to the final concentration in the cell (i.e. 40  $\mu$ M CMCoA means that one syringe contained 80  $\mu$ M CMCoA). Transients were collected under a series of conditions including at pHs 7.0, 8.0 or 8.9 from 0.5  $\rightarrow$  4  $\mu$ M TpcSTMD317G against 40  $\mu$ M CMCoA, or at pH 8.0 with varying CMCoA concentrations (0  $\rightarrow$  210  $\mu$ M CMCoA). Transients were collected for CMX at 40 or 80  $\mu$ M and 2  $\mu$ M enzyme at pH 7.0, 8.0 or 8.9.

### I.3 Results

*Purification*- TM was purified several times according to published methods (12) and yielded protein of high purity. The mutant TM-D317G was found to bind irreversibly to the same column material used in the final step of purification of TM. A series of agarose dye affinity column materials was screened (Cibracon Blue 3GA, Reactive Brown 10, Reactive Green 19, Reactive Blue 4, Reactive Yellow 86, Reactive Red 120). Cibracon Blue-3GA was found to purify the protein and bind TM-D317G reversibly (Figure I.1). The purified protein was analyzed by ESI-MS and found to have a mass of  $42,765.0 \pm 4$  Da. This mass is consistent with the TMD317G – Met1, which has an expected mass of 42,766.0 Da.



**Figure I.1:** 9% SDS-PAGE analysis of Cibracon Blue 3-GA Agarose column purification of *TpCSTMD317G*. Lane 1, Sigma high range marker; lane 2, 5.8  $\mu$ g TM-D317G, lane 3, concentrated Q sepharose eluant; lanes 4-6, Cibracon Blue 3-GA flowthrough; Lanes 7-13, Fractions eluted in buffer with 50 mM ammonium sulfate. Lanes 14 and 15, 1 M ammonium sulfate wash. Fractions containing protein at 50 mM ammonium sulfate were pooled and concentrated.

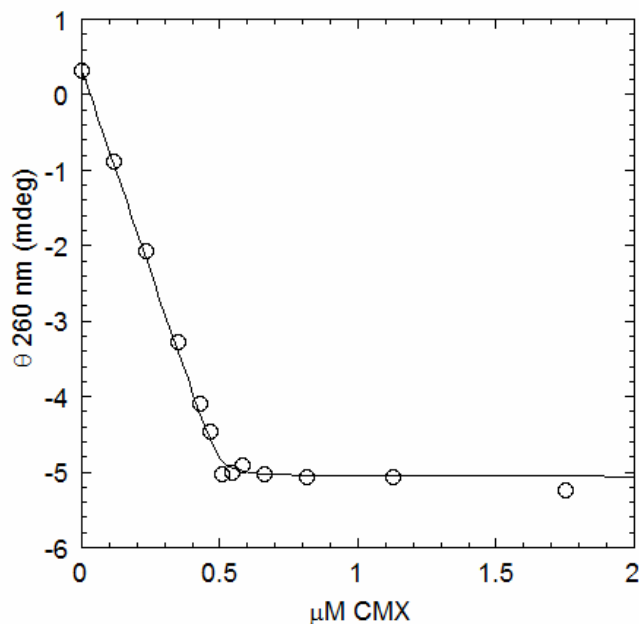
*CMCoA and CMX Titrations.* The binding of AcCoA analogs CMX and CMCoA cause a change in the circular dichroism spectrum of all *TpCS* forms due to the immobilization of the adenine moiety in a chiral protein environment. This change allowed for the determination of binding constants by monitoring the change in ellipticity at 260 nm as a function of titrant (Figures I.2-I.9). Data collected were fit to equation 2 to yield values of  $K_d$  (Table I.1). The binding affinities of *TpCS*-D317G for CMCoA or CMX were determined, and the effects of the viscosogens glycerol and sucrose on the binding affinity assessed. The binding affinity of *TpCS*-D317N and TM-D317G for CMX was also determined. All of the  $K_d$ s determined are much lower than the concentration of enzyme used in the titrations, so some caution is advised in interpreting their values.

**Table I.1.** Acetyl-CoA analogues  $K_d$ s Determined for *TpCS* mutants

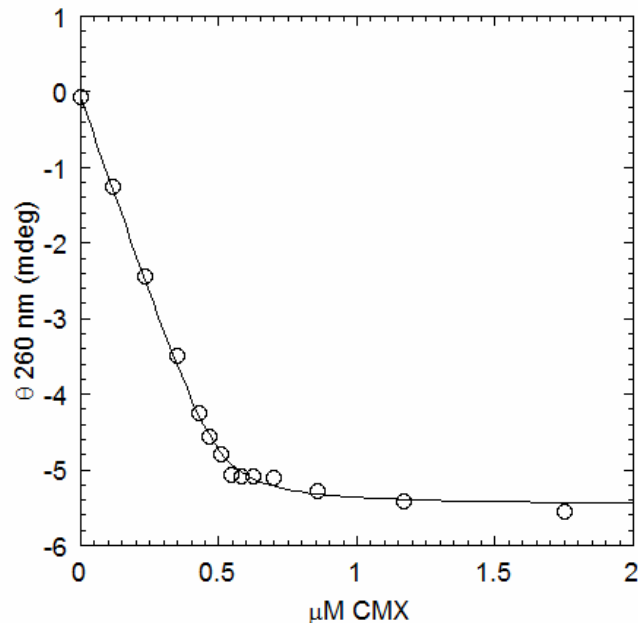
	$K_d$ (CMCoA)			$K_d$ (CMX)		
	Buffer	+20% Glycerol	+30% Sucrose	Buffer	+20% Glycerol	+ 30% Sucrose
<i>TpCS</i> -D317G <sup>a</sup>	16 ± 3 nM	3 ± 1 nM	5 ± 1 nM	11 ± 2 nM	5 ± 1 nM	8 ± 3 nM
<i>TpCS</i> -D317N <sup>a</sup>	-	-	-	1 ± 1 nM	-	-
TM-D317G <sup>b</sup>	60 ± 10 nM	-	-	-	-	-

<sup>1</sup>- Titrations were performed at 20°C in 50 mM EPPS, 0.1 mM EDTA, 5 μM OAA and 0.5 μM enzyme in a 10 cm path length cell. Titrations were performed with the addition of viscosogens glycerol and sucrose as noted. <sup>2</sup>- Titration was performed at 20°C in 50 mM EPPS, 0.1 mM EDTA, 50 μM OAA and 8.98 μM enzyme in a 1 cm path length cell.

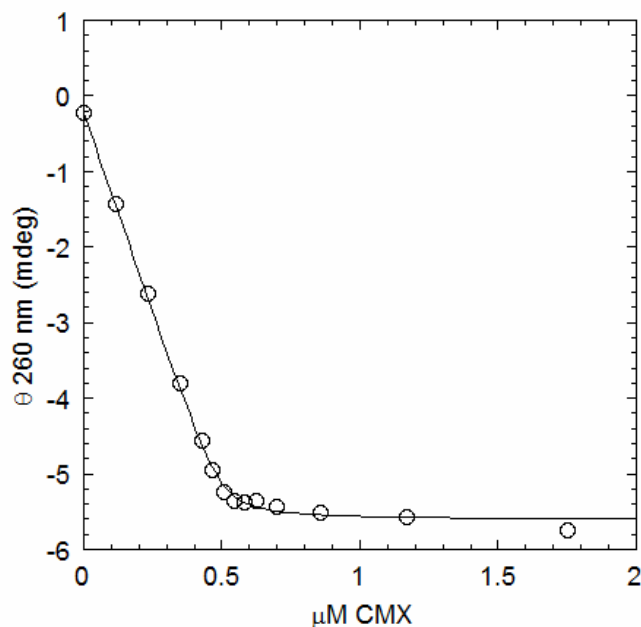
Mutant *TpCS*-D317N bound the AcCoA-enolate analogue CMX more tightly than *TpCS*-D317G. This higher affinity could be due to hydrogen bonding of the carboxylate of CMX by the Asn side chain. Titrations for the *TpCS*-D317G mutant indicate that the addition of the viscosogens glycerol and sucrose caused *TpCS*-D317G to bind CMCoA more tightly, while it had nearly the same affinity for CMX (Table I.1). The TM-D317G mutant had the lowest acetyl-CoA analogue affinity of the enzymes studied (Table I.1).



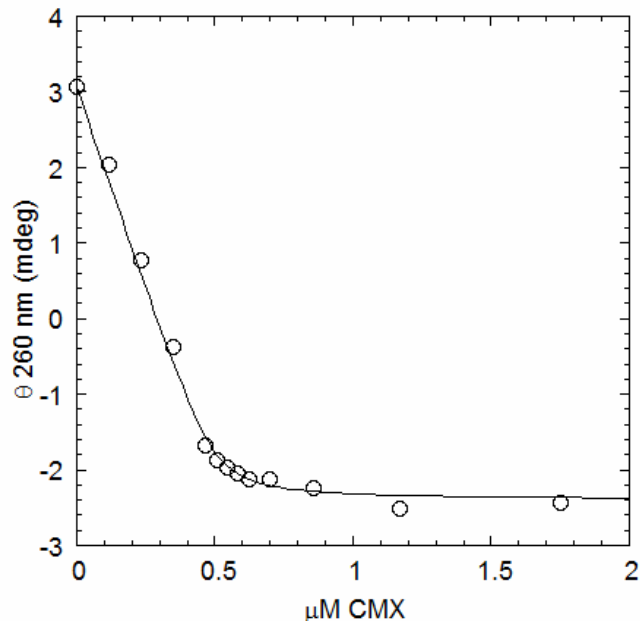
**Figure I.2:** CMX titration of *TpCSD317N*. Ellipticity at 260 nm versus [CMX] for *TpCSD317N* at 0.50 μM [subunit], 50 mM EPPS, and 0.1 mM EDTA pH 8.0 in a 10 cm path length cell. The solid line represents a fit of data to equation 2 with constants  $E = 0.50 \mu\text{M}$ ,  $\theta_{i,260} = 0.3258$ , and variables  $K_{d,\text{CMX}} = 1.1 \pm 1.0 \text{ nM}$ ,  $\Delta\theta_{260} = -5.39 \pm 0.06$ .



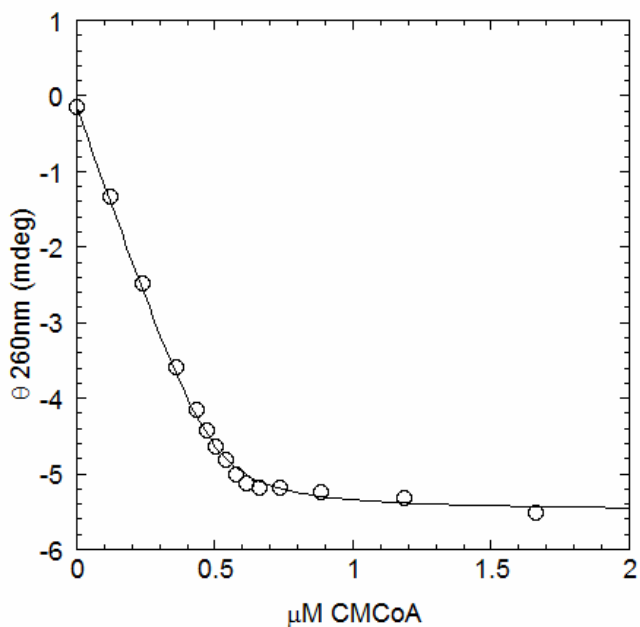
**Figure I.3:** CMX titration of *TpCSD317G*. Ellipticity at 260 nm versus [CMX] for *TpCSD317G* at 0.50  $\mu\text{M}$  [subunit], 50 mM EPPS, 0.1 mM EDTA pH 8.0 and 5  $\mu\text{M}$  OAA in a 10 cm path length cell. The solid line represents a fit of data to equation 2 with constants  $E = 0.50 \mu\text{M}$ ,  $\theta_{i,260} = -0.0727$ , and variables  $K_{d,\text{CMX}} = 11 \pm 2 \text{ nM}$ ,  $\Delta\theta_{260} = -5.41 \pm 0.06$ .



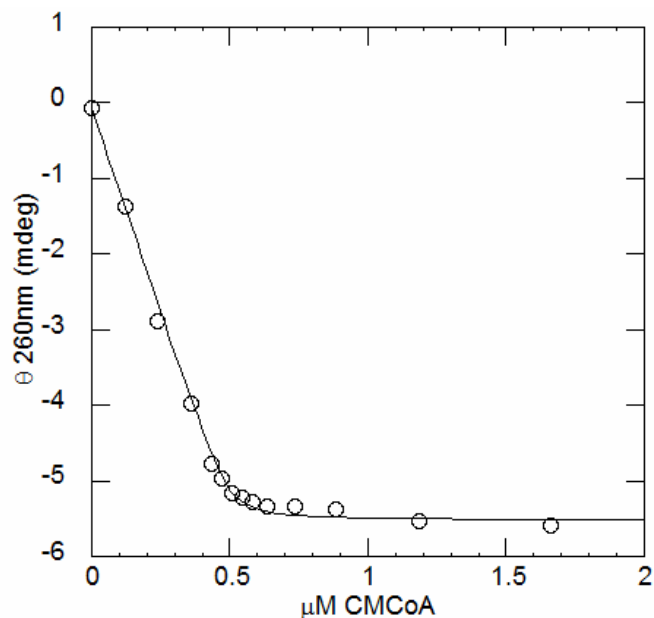
**Figure I.4:** CMX titration of *TpCSD317G* in 20% Glycerol. Ellipticity at 260 nm versus [CMX] for *TpCSD317G* at 0.50  $\mu\text{M}$  [subunit], 50 mM EPPS, 0.1 mM EDTA pH 8.0, 5  $\mu\text{M}$  OAA and 20% glycerol in a 10 cm path length cell. The solid line represents a fit of the data to equation 2 with constants  $E = 0.50 \mu\text{M}$ ,  $\theta_{i,260} = -0.217$  and variables  $K_{d,\text{CMX}} = 5 \pm 1 \text{ nM}$ ,  $\Delta\theta_{260} = -5.39 \pm 0.05$ .



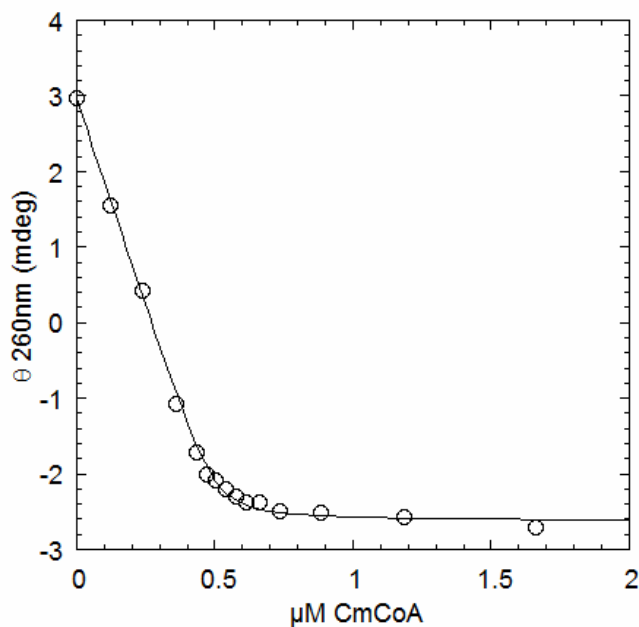
**Figure I.5:** CMX titration of *TpCSD317G* in 30% Sucrose. Ellipticity at 260 nm versus [CMX] for *TpCSD317G* at 0.50 μM [subunit], 50 mM EPPS, 0.1 mM EDTA pH 8.0, 5 μM OAA and 30% sucrose in a 10 cm path length cell. The solid line represents a fit of the data to equation 2 with constants  $E = 0.50 \mu\text{M}$ ,  $\theta_{i,260} = 3.064$ , and variables  $K_{d,\text{CMX}} = 8 \pm 3 \text{ nM}$ ,  $\Delta\theta_{260} = -5.46 \pm 0.09$ .



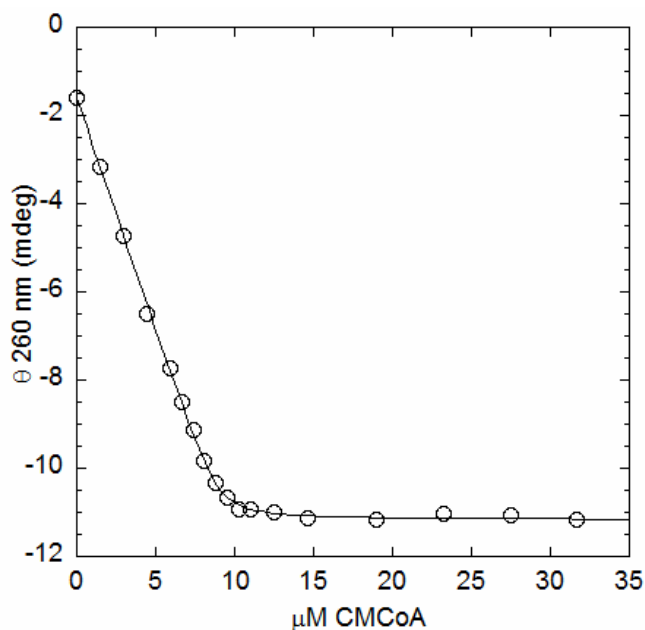
**Figure I.6:** CMCoA titration of *TpCSD317G*. Ellipticity at 260 nm versus [CMCoA] for *TpCSD317G* at 0.50 μM [subunit], 50 mM EPPS, 0.1 mM EDTA pH 8.0 and 5 μM OAA in a 10 cm path length cell. The solid line represents a fit of the data to equation 2 with constants  $E = 0.50 \mu\text{M}$ ,  $\theta_{i,260} = -0.150$  and variables  $K_{d,\text{CMCoA}} = 16 \pm 3 \text{ nM}$ ,  $\Delta\theta_{260} = -5.36 \pm 0.06$ .



**Figure I.7:** CMCoA titration of *TpCSD317G* in 20% Glycerol. Ellipticity at 260 nm versus [CMCoA] for *TpCSD317G* at 0.50  $\mu\text{M}$  [subunit], 50 mM EPPS, 0.1 mM EDTA pH 8.0, 5  $\mu\text{M}$  OAA and 20 % glycerol, in a 10 cm path length cell. The solid line represents fit of data to equation 2 with constants  $E = 0.50 \mu\text{M}$   $\theta_{i,260} = -0.074$  and variables  $K_{d,\text{CMCoA}} = 3 \pm 1 \text{ nM}$ ,  $\Delta\theta_{260} = -5.45 \pm 0.06$ .



**Figure I.8:** CMCoA titration of *TpCSD317G* in 30% Sucrose. Ellipticity at 260 nm versus [CMCoA] for *TpCSD317G* at 0.50  $\mu\text{M}$  [subunit], 50 mM EPPS, 0.1 mM EDTA pH 8.0, 5  $\mu\text{M}$  OAA and 30 % sucrose, in a 10 cm path length cell. The solid line represents a fit of the data to equation 2 with constants  $E = 0.50 \mu\text{M}$ ,  $\theta_{i,260} = 2.968$  and variables  $K_{d,\text{CMCoA}} = 5 \pm 1 \text{ nM}$ ,  $\Delta\theta_{260} = -5.60 \pm 0.05$ .

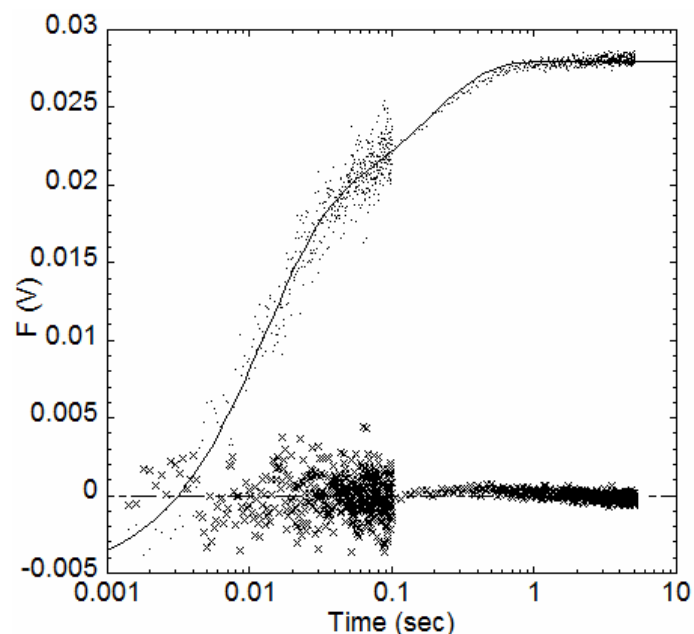


**Figure I.9:** CMCoA titration of *Tp*CSTMD317G. Ellipticity at 260 nm versus [CMCoA] for *Tp*CSTMD317G at 8.98 μM [subunit], 50 mM EPPS, 50 μM OAA, 0.1 mM EDTA pH 8.0, in a 1 cm path length cell. The solid line represents a fit of the data to equation 2 with constants  $E = 8.98 \mu\text{M}$ ,  $\theta_{i,260} = -1.5938$  and variables  $K_{d,\text{CMCoA}} = 60 \pm 10 \text{ nM}$ ,  $\Delta\theta_{260} = -9.58 \pm 0.04$ .

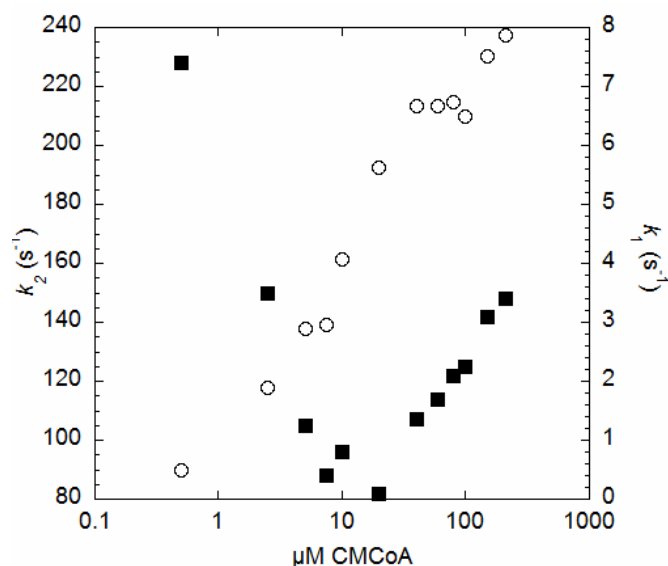
*Stopped flow fluorescence.* The binding of CMCoA to the ternary complex of TM-D317G (0.5 μM [subunit] and 200 μM OAA, 50 mM EPPS 0.1 mM EDTA pH 8.0) was examined at varying concentrations of CMCoA. Attempts to fit the data to single, double, triple and quadruple exponential equations, including attempts with a linear correction, all gave fits with systematic inconsistencies. A double-exponential equation yielded the best fit, as shown for a sample trace at 20 μM CMCoA (Figure I.10). The residuals of the fit are also shown, and as can be seen in both the transient and the residuals, the data from ~0.1-0.5 s and the data beyond 5 s deviate from the fit. Therefore, a plot of the derived parameters  $k_1$  and  $k_2$  versus [CMCoA] cannot be interpreted (Figure I.11). These trends are not seen for *Tp*CS-D317G or *Tp*CS-TM.



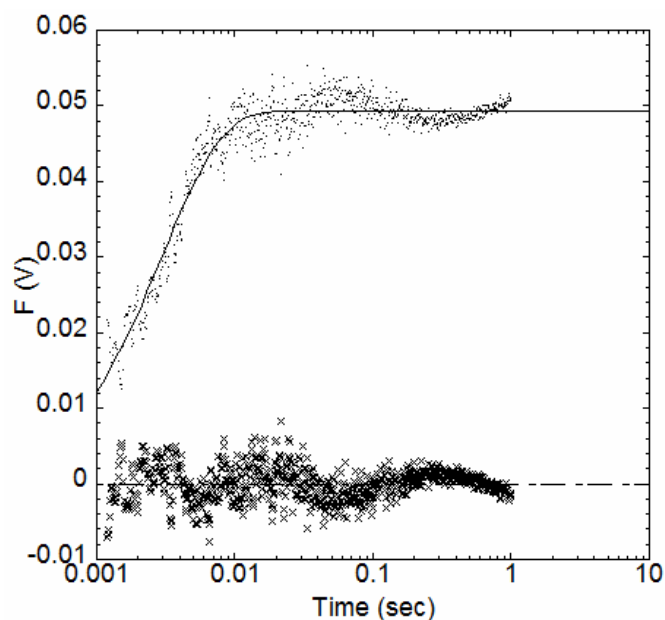
Several conditions were surveyed for TM-D317G and CMCoA to find conditions which would yield data that could be fit. These conditions included: 0.5-4.0  $\mu\text{M}$  TMD317G at 40  $\mu\text{M}$  CMCoA at pH 8.0, at 0.5-3.0  $\mu\text{M}$  subunit; and 2.5  $\mu\text{M}$  subunit with 50 mM KCl at pH 8.9; and 0.5-4.0  $\mu\text{M}$  subunit at 40  $\mu\text{M}$  CMCoA at pH 7.0. Experiments with CMX indicated that the transient fluorescence changes were much smaller in magnitude. Traces were collected at pH 7.0 at 40  $\mu\text{M}$  CMX and pH 8.0 or 8.9 at 80  $\mu\text{M}$  CMX. Attempts to characterize the kinetics of CMX binding proved even more difficult, as the transients were of smaller magnitude and appeared more complex (Figure I.12). None of the equations used above fit the data as well as a single exponential. The residuals from this fit are also shown (Figure 5.13). No suitable conditions were found for the study of CMCoA or CMX binding.



**Figure I.10:** Stopped flow fluorescence of CMCoA binding TM-D317G-OAA at pH 8.0. Traces collected at 0.5  $\mu\text{M}$  TM-D317G, 200  $\mu\text{M}$  OAA and 20  $\mu\text{M}$  CMCoA final mixed concentrations. Voltages for time points shown as dots to highlight double exponential fit to the data (solid line). Residuals of the fit are shown (X-token). Dashed line represents  $y = 0$ . Double exponential fit to parameters  $V_f = 0.02798 \pm 0.00004$ ;  $\Delta V_{k_1} = 0.0101 \pm 0.0003$ ;  $k_1 = 5.6 \pm 0.3 \text{ s}^{-1}$ ;  $\Delta V_{k_2} = 0.0233 \pm 0.0004$ ;  $k_2 = 82 \pm 2 \text{ s}^{-1}$ .



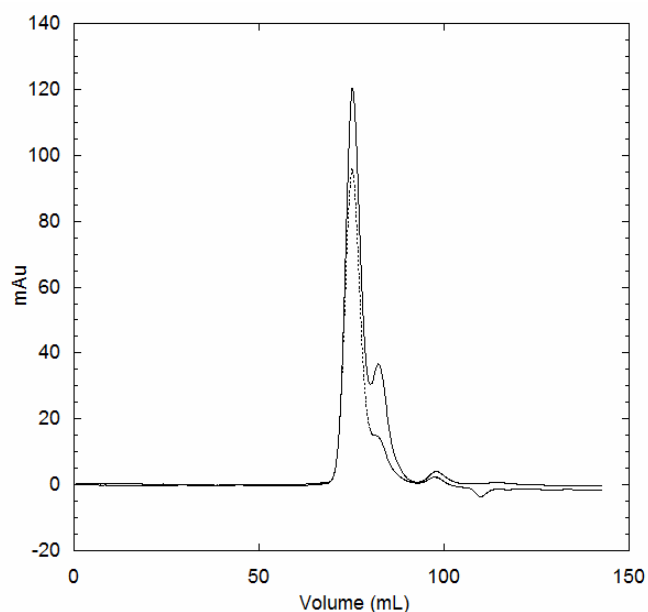
**Figure I.11:** Values of  $k_1$  and  $k_2$  from double exponential fits of TM-D317G as a function of [CMCoA]. Plot shown on log scale to highlight the change in  $k_2$  (filled squares), and  $k_1$  (open circles).



**Figure I.12:** Stopped flow fluorescence of CMX binding TM-D317G-OAA at pH 8.0. Traces collected at 2  $\mu\text{M}$  TM-D317G, 200  $\mu\text{M}$  OAA and 80  $\mu\text{M}$  CMX final mixed concentrations. Voltages for time points shown as dots to highlight double exponential fit of data (solid line). Residuals of the fit are shown (X-tokens). Dashed line represent  $y = 0$ . Double exponential fit to parameters  $V_f = 0.04935 \pm 0.00009$ ;  $\Delta V_{k_1} = 0.0519 \pm 0.0009$ ;  $k_1 = 333 \pm 7 \text{ s}^{-1}$ .

*Gel-filtration profile of TM-D317G.* The inconsistent stopped flow fluorescence data fitting led us to suspect an inhomogeneous aggregation state of the multiply-mutated TMD317G protein. Upon purification, TM-D317G was initially titrated with a solution of OAA. The resulting titration yielded an end point below the concentration of enzyme in the cuvette (~70% [subunit]). A subsequent titration with CMCoA indicated that this was not a result of the concentration of protein being incorrect. This suggested that the enzyme co-purified with partial occupancy of OAA. Ordinarily this would be confirmed by enzyme activity, but the D317G mutant is inactive.

In order to investigate the possibility of partial occupancy of OAA, the protein was analyzed by gel filtration and analyzed again with the addition of OAA to the buffer. The gel-filtration profile indicated two peaks for the purified protein (Figure I.14). When 500  $\mu$ M OAA was added to the buffer, the second peak decreased in magnitude (Figure I.14). This is consistent with TM-D317G co-purifying with tightly bound OAA. Indeed, D317G is known to bind OAA very tightly (12).



**Figure I.13:** Gel filtration profile of TM-D317G. TM-D317G (2.5 mg) was run on a gel filtration column in 20 mM Tris 1 mM EDTA, and 50 mM KCl as buffer with (dashed line) and without (solid line) 500  $\mu$ M OAA in the buffer. TM-D317G yielded two peaks with apparent molecular masses of 59 kDa and 31 kDa. The addition of OAA to the buffer caused reduction in the peak at 80 mL (31 kDa), consistent with TM-D317G purifying in two forms, unliganded and OAA-bound.

*Quantum yield determinations of TM-D317G-* The absolute quantum yield of TM-D317G was determined following a published procedure previously used for other forms of *TpCS* (12) (Table I.2).

	Unliganded protein absolute Quantum Yield in reference to Trp (0.14) (relative to Trp)	OAA complex relative to unliganded protein	Ternary complex
TM-D317G <sup>a</sup>	0.11 (0.79)	0.30	0.46
TM <sup>b</sup>	0.27 (1.99)	0.21	n.d.

<sup>a</sup>-Protein purifies with partial occupancy of OAA. Therefore the actual value of absolute quantum yield of completely unliganded protein should be higher and the ratio of OAA complex relative to unliganded protein and ternary complex ratio to be lower. <sup>b</sup>-values from(12)

#### I.4 Discussion

*Purification of TM forms-* The current method of growth and purification yields very pure protein. However, the total yield of protein is typically low (~3 mg/L culture). The small broth size used during growth of cells (300 mL in 2 L flasks), makes the production of large quantities of protein difficult. If expression of large quantities of protein is required, an alternative system may be needed.

The TM-D317G mutant irreversibly bound to the Reactive Red 120 CL-3000B column used in the previous purifications of TM, necessitating a search for a new method to purify the protein. This was accomplished by the use of a Cibracon Blue 3GA column. This has the additional advantage of being commercially available in large quantities.

The difference in purification for TM-D317G may be due to the enzyme purifying in two forms (Figure I.13); TM does not co-purify with OAA (12). The cause of this difference may be the removal of the catalytically important Asp317. The D317G mutant

cannot convert bound OAA to citrate, and this may lead to accumulation of OAA on the enzyme during expression of the protein.

*Stopped flow fluorescence.* While the data collected for the TM-D317G mutant failed to give the expected simplified kinetics, the data collected can be used to aid qualitatively in the analysis of data collected from the *TpCS*-wt, *TpCS*-D317G, *TpCS*-D317N and TM. The transient fluorescence change for both *TpCSD317G* and TM are biphasic. *TpCSD317G* shows a large rapid increase, followed by a small slow decrease in transient fluorescence upon binding CMC<sub>o</sub>A or CMX. However, TM shows a small rapid increase, followed by a large slow decrease in transient fluorescence upon binding CMC<sub>o</sub>A or CMX. The traces for TM-D317G resemble the traces of the *TpCSD317G* and indicate that the transient fluorescence changes seen in TM are not due to the replacement of three tryptophans with phenylalanine. The lack of a slow decrease in the traces for TM-D317G suggests that the second phase seen in *TpCSD317G* is due to the three tryptophans not present in TM-D317G.

## I.5 References

1. Mullins, E. A., Francois, J. A., and Kappock, T. J. (2008) A specialized citric acid cycle requiring succinyl-Coenzyme A (CoA): Acetate CoA-transferase (AarC) confers acetic acid resistance on the acidophile *Acetobacter acetii*, *Journal of Bacteriology* 190, 4933-4940.
2. Kurz, L. C., Shah, S., Crane, B. R., Donald, L. J., Duckworth, H. W., and Drysdale, G. R. (1992) Proton uptake accompanies formation of the ternary complex of citrate synthase, oxaloacetate, and the transition-state analog inhibitor, carboxymethyl-CoA. Evidence that a neutral enol is the activated form of acetyl-CoA in the citrate synthase reaction, *Biochemistry* 31, 7899-7907.
3. Gu, Z., Drueckhammer, D. G., Kurz, L., Liu, K., Martin, D. P., and McDermott, A. (1999) Solid state NMR studies of hydrogen bonding in a citrate synthase inhibitor complex, *Biochemistry* 38, 8022-8031.
4. Kurz, L. C., Ackerman, J. J. H., and Drysdale, G. R. (1985) Evidence from carbon-13 NMR for polarization of the carbonyl of oxaloacetate in the active site of citrate synthase, *Biochemistry* 24, 452-457.
5. Kurz, L. C., and Drysdale, G. R. (1987) Evidence from Fourier transform infrared spectroscopy for polarization of the carbonyl of oxaloacetate in the active site of citrate synthase, *Biochemistry* 26, 2623-2627.
6. Evans, C. T., Kurz, L. C., Remington, S. J., and Srere, P. A. (1996) Active site mutants of pig citrate synthase: Effects of mutations on the enzyme catalytic and structural properties, *Biochemistry* 35, 10661-10672.
7. Ernst Bayer, B. B., Hermann Eggerer, . (1981) Evidence from inhibitor studies for conformational changes of citrate synthase, *European Journal of Biochemistry* 120, 155-160.
8. Francois, J. A., Starks, C. M., Sivanuntakorn, S., Jiang, H., Ransome, A. E., Nam, J.-W., Constantine, C. Z., and Kappock, T. J. (2006) Structure of a NADH-insensitive hexameric citrate synthase that resists acid inactivation, *Biochemistry* 45, 13487-13499.
9. Kosicki, G. W., and Srere, P. A. (1961) Kinetic studies on the citrate-condensing enzyme, *Journal of Biological Chemistry* 236, 2560-2565.
10. Kurz, L. C., Drysdale, G., Riley, M., Tomar, M. A., Chen, J., Russell, R. J. M., and Danson, M. J. (2000) Kinetics and mechanism of the citrate synthase from the thermophilic archaeon *Thermoplasma acidophilum*, *Biochemistry* 39, 2283-2296.
11. Kurz, L. C., Drysdale, G. R., Riley, M. C., Evans, C. T., and Srere, P. A. (1992) Catalytic strategy of citrate synthase: effects of amino acid changes in the acetyl-CoA binding site on transition-state analog inhibitor complexes, *Biochemistry* 31, 7908-7914.
12. Kurz, L. C., Fite, B., Jean, J., Park, J., Erpelding, T., and Callis, P. (2005) Photophysics of tryptophan fluorescence: Link with the catalytic strategy of the citrate synthase from *Thermoplasma acidophilum*, *Biochemistry* 44, 1394-1413.
13. Remington, S. J. (1992) Structure and mechanism of citrate synthase, *Current Topics in Cellular Regulation* 33, 209-229.
14. Kurz, L. C., Nakra, T., Stein, R., Plungkhen, W., Riley, M., Hsu, F., and Drysdale, G. R. (1998) Effects of changes in three catalytic residues on the relative stabilities of some of the intermediates and transition states in the citrate synthase reaction, *Biochemistry* 37, 9724-9737.

## SPRINGER LICENSE TERMS AND CONDITIONS

Apr 14, 2009

---



---

This is a License Agreement between Charles Z Constantine ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

License Number	2167780657134
License date	Apr 14, 2009
Licensed content publisher	Springer
Licensed content publication	Archives of Microbiology
Licensed content title	The internal pH of <i>Acetobacterium wieringae</i> and <i>Acetobacter aceti</i> during growth and production of acetic acid
Licensed content author	Ulrike Menzel
Licensed content date	Oct 1, 1985
Volume number	143
Issue number	1
Pages	47 - 51
Type of Use	Thesis / Dissertation
Details of use	Print
Requestor Type	Individual
Portion of the article	Figures
Title of your thesis / dissertation	An Investigation of <i>Acetobacter aceti</i> N5-Carboxyaminoimidazole Ribonucleotide Mutase and Its purE-purK Operon
Expected completion date	May 2009
Billing Type	Invoice
Company	Charles Z Constantine
Billing Address	
	St. Louis , MO 63108
	United States
Customer reference info	
Total	0.00 USD
Terms and Conditions	

### Introduction

The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the



following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

#### Limited License

With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry. Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided it is password protected or on the university's intranet, destined to microfilming by UMI and University repository. For any other electronic use, please contact Springer at ([permissions.dordrecht@springer.com](mailto:permissions.dordrecht@springer.com) or [permissions.heidelberg@springer.com](mailto:permissions.heidelberg@springer.com))

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, provided permission is also obtained from the (co) author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well). Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

#### Altering/Modifying Material: Not Permitted

However figures and illustrations may be altered minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at [permissions.dordrecht@springer.com](mailto:permissions.dordrecht@springer.com) or [permissions.heidelberg@springer.com](mailto:permissions.heidelberg@springer.com))

#### Reservation of Rights

Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

#### Copyright Notice:

Please include the following copyright citation referencing the publication in which the material was originally published. Where wording is within brackets, please include verbatim.

"With kind permission from Springer Science+Business Media: <book/journal title, chapter/article title, volume, year of publication, page, name(s) of author(s), figure number(s), and any original (first) copyright notice displayed with material>."

Warranties: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

#### Indemnity

You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

#### No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer Science + Business Media's written permission.

#### No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

#### Objection to Contrary Terms

Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

#### Jurisdiction

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by the country's law in which the work was originally published.

v1.2

**Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.**

**If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 30 days of the license date. Payment should be in the form of a check or money order referencing your account number and this license number 2167780657134.**

**If you would prefer to pay for this license by credit card, please go to <http://www.copyright.com/creditcard> to download our credit card payment authorization form.**

#### **Make Payment To:**

**Copyright Clearance Center  
Dept 001  
P.O. Box 843006  
Boston, MA 02284-3006**

**If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2167780657134 and noting the reason for cancellation.**

**Questions? [customercare@copyright.com](mailto:customercare@copyright.com) or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.**

