METABOLIC EXPANSION IN *ACINETOBACTER BAYLYI* ADP1 FOR ENHANCED AROMATIC COMPOUND CATABOLISM

by

STACY RENE BEDORE

(Under the Direction of Ellen L. Neidle)

ABSTRACT

Acinetobacter baylyi ADP1 degrades many aromatic compounds and has a uniquely tractable genetic system. Recent interest in aromatic compound metabolism stems from diverse biotechnology applications. My research focused on increasing the number and range of aromatic compounds metabolized by *A. baylyi* to enhance this bacterial host's ability to convert a heterogenous lignin-derived mixture to a valuable product. The lignin portion of renewable lignocellulosic biomass is currently vastly underutilized. Foreign DNA segments encoding catabolic pathways were inserted in the ADP1 chromosome and amplified in tandem arrays prior to adaptive laboratory evolution. Additional research goals include understanding how metabolic functions become regulated and integrated in a new host. This approach, therefore, explores fundamental aspects of evolution, horizontal gene transfer, and metabolic regulation. In this dissertation, I describe the chromosomal integration and adaptation in *A. baylyi* of a *meta*-cleavage protocatechuate degradation pathway from *Pseudoalteromonas atlantica*. Similar experiments used a codon-optimized pathway for veratrate and isovanillate degradation from *Pseudomonas putida*. *A. baylyi* strains able to grow on aromatic compounds using these non-native pathways were isolated and characterized. Whole genome sequencing revealed multiple

mutations, and new strains were constructed with different combinations of mutations to determine the significance of genetic changes. A single chromosomal copy of the 7-kbp *meta*-cleavage pathway was sufficient for growth at 25 °C, whereas growth using this pathway at higher temperatures (up to 37 °C) occurred only with multiple chromosomal copies. A remarkable expansion of the chromosome initially added as much as 2 Mbp of repeated DNA to the normally 3.6 Mbp chromosome. After long-term culturing, copy number decreased and mutations accrued that increase the growth temperature at which the foreign pathway is functional. Mutations unveiled the importance of a previously undescribed phosphotransferase system, several peptidyl-prolyl isomerases, and the GroE chaperone. In additional studies, I characterized aspartate metabolism and regulation in ADP1. This research was done in conjunction with teaching an undergraduate microbiology research course. All studies focused on metabolic regulation and involved the development and modification of methods that exploit the exceptional efficiency of natural transformation and allelic replacement in ADP1, an ideal model organism for synthetic biology.

INDEX WORDS: ADP1, *Acinetobacter baylyi*, PCA-4,5, veratrate, aspartate, adaptive evolution, PTS, peptidyl-prolyl isomerase, chaperone, aromatic compound catabolism, thermal tolerance

METABOLIC EXPANSION IN *ACINETOBACTER BAYLYI* ADP1 FOR ENHANCED AROMATIC COMPOUND CATABOLISM

by

STACY RENE BEDORE

BS, Illinois State University, 2016

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

METABOLIC EXPANSION IN *ACINETOBACTER BAYLYI* ADP1 FOR ENHANCED AROMATIC COMPOUND CATABOLISM

by

STACY RENE BEDORE

Major Professor:	Ellen L. Neidle
Committee:	Mark Eiteman
	Zachary Lewis
	Jan Mrázek

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2021

DEDICATION

I dedicate this dissertation to my grandparents, William and Shirley Hamrick, who both passed away before I completed my PhD. They provided me with constant support and love through all the adventures I've embarked on in my life. I hope this makes you two proud.

ACKNOWLEDGEMENTS

So many people have given me support and guidance in completing my PhD and I would like to thank and acknowledge them:

- Dr. Ellen Lee Neidle, my thesis advisor, for years of incredible scientific, professional, and personal support and advice. I have grown so much during my PhD under her direction and guidance.
- Drs. Mark Eiteman, Zachary Lewis, and Jan Mrázek, my committee members, for helpful discussion, feedback, and advice on my research and career.
- Kyle Healy, my husband, for everything. Without his unceasing support and love, I would not be the person I am today. I owe this achievement to him.
- My parents and siblings for a lifetime of love and support through all my journeys
- Past and present members of the Neidle lab, Dr. Melissa Tumen-Velasquez, Dr. Alaa Ahmed, Emily McIntrye, Chantel Duscent-Maitland, Alyssa Baugh, Lauren Slarks, William Christopher Moxley, and Alicia Schmidt.
- The J. Craig Venter Institute for two fantastic summer internships, guidance and mentorship from Dr. Yo Suzuki, and research support from Emanuel Vasquez.
- Neidle lab collaborators, Drs. Gregg Beckham and Chris Johnson at the National Renewable Energy Laboratory, and Drs. Keith Tyo and Bradley Biggs from Northwestern University for great partnerships and publications
- The Microbiology Department office staff for answering all my questions and not letting me miss any administrative deadlines.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Purpose of study

This dissertation focuses on expanding the catabolism of lignin-derived aromatic compounds by the non-pathogenic soil bacterium, *Acinetobacter baylyi* ADP1. Aromatic compounds are highly abundant in the environment, including as components of lignin, a complex polymer that is prevalent in plant biomass. Due to the defining aromatic, or benzene, ring, aromatic compounds are resistant to degradation. The majority of the recycling of carbon from aromatic compounds is done by specialist microorganisms because of the chemical stability of these compounds. Some microbes can utilize the oxidative power of molecular oxygen to break open the benzene ring, despite the risk that oxygen radicals will be produced and can cause damage to the cells. This potential danger has been proposed as the reason that higher organisms evolved to exclude enzymes that use molecular oxygen from most metabolic reactions (1).

Interest in the bacterial degradation of aromatic compounds originally focused on the critical role played by microorganisms in catabolizing lignin and other plant-produced compounds as part of the natural carbon cycle (2). Recently, interest has developed in using lignin as a renewable resource to produce economically valuable chemicals, a process called lignin valorization (3). One promising scheme takes advantage of the ability of certain bacteria to catabolize the many components of lignin and accumulate a chemical product of interest, such as pyruvate or wax esters (4, 5). To make such a scheme feasible in the laboratory, researchers need to modify aspects of microbial metabolism for biotechnology goals. For example, in nature, the

microbial degradation of lignin occurs very slowly by a consortium of microbes including aerobic and anaerobic fungi and bacteria (6). The net result in such an ecosystem is the production of new cells rather than the accumulation of a target chemical.

The advent of molecular, systems, and synthetic biology has ushered in exciting new possibilities for manipulating bacterial metabolism for biotechnology (7, 8). Various bacterial species, including *Acinetobacter baylyi* ADP1, have been studied for their ability to degrade a wide variety of lignin-derived aromatics (9). *A. baylyi* ADP1 is uniquely suited for the manipulation of aromatic compound metabolism because of its catabolic versatility, its facile genetic system, and newly developed genetic tools that facilitate laboratory evolution (10-12). My doctoral research contributes to the use of *Acinetobacter baylyi* ADP1 as an ideal host for engineering aromatic compound metabolism (11).

My focus has been to expand *A. baylyi's* ability to degrade lignin-derived aromatic compounds by introducing foreign pathways to increase the number of aromatics it can catabolize. I also seek to improve the efficiency of degrading multiple aromatic compounds in a mixture. The following literature review provides background material about the principles of lignin and aromatic compound degradation, the *A. baylyi* genetic system, and metabolic expansion using laboratory evolution. Additional background information is provided concerning proteins and processes that were revealed to contribute to the acquisition of new phenotypic traits in *A. baylyi* during my studies. For example, mutations affecting a phosphotransferase system, the GroE chaperone, peptidyl-prolyl isomerases, and the ClpAP protease provided selective benefit during metabolic studies described in this thesis.

Principles of aromatic compound and lignin degradation

Lignin, a major component of lignocellulosic biomass, is the second most abundant naturally occurring polymer on earth after cellulose (13). It functions to provide strength to plants, aids in water transport, and protects plants from decomposition (14). While cellulose has been well researched as a renewable resource for producing biofuels, lignin is typically separated from the polysaccharides and essentially discarded (8, 15, 16). Lignin is too valuable a renewable resource to be wasted. Not only is this polymer naturally abundant, but it has even greater energy content than cellulose due to its higher carbon to oxygen ratio. For the conversion of lignocellulose to biofuels to be profitable, compounds of value need to be produced from lignin. This process of lignin valorization is difficult because of the polymer's inherent chemical complexity.

Lignin is synthesized from *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, three aromatic components that differ from each other in their amount of methoxylation (**Figure 1.1**) (17). These units are connected by various C-O and C-C linkages involving the α, β, and/or γ carbons (14). The content and composition of lignin in different plant species varies widely depending on the ratios of the monomeric units and the type of linkages that connect them. Heterogeneity of lignin feedstocks can additionally result from the types of harsh pre-treatment methods used for separation and initial breakdown of this polymer (9). Common pre-treatment methods involve chemical, thermal, and biological processes that yield different complex mixtures that tend to be rich in aromatic compounds and organic acids (18). Since no two lignin streams are identical, further processing methods for the production of industrially relevant chemicals must be able to tolerate heterogeneity in lignin feedstocks (3).

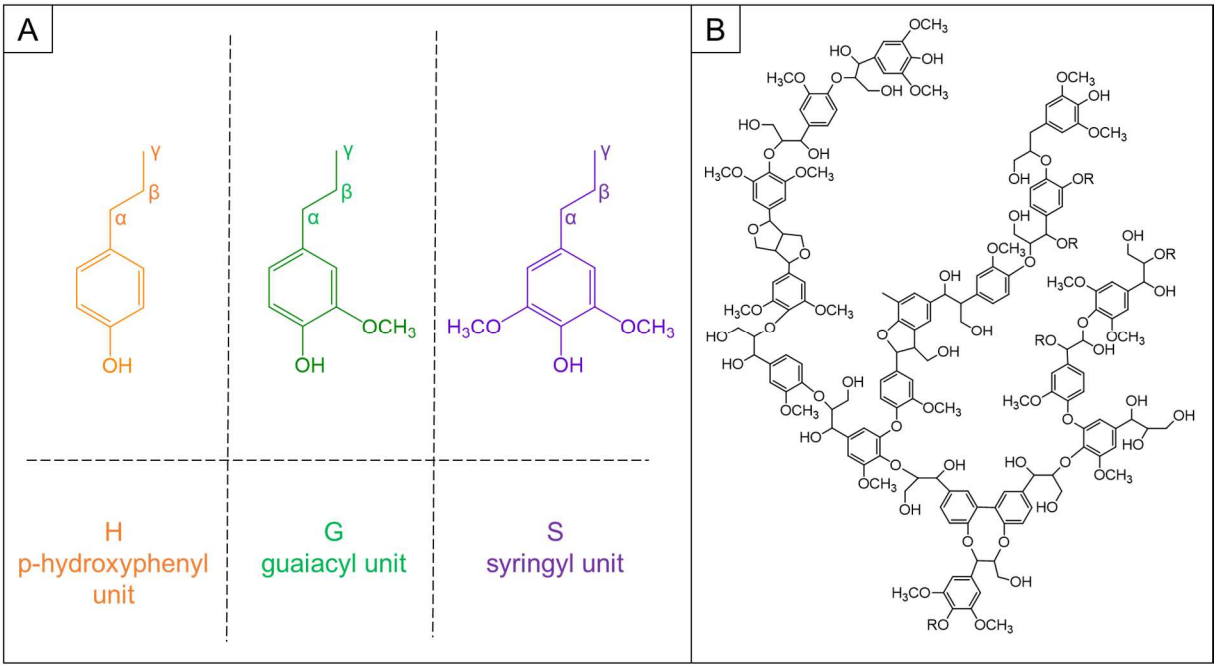


Figure 1.1: Monomeric units of lignin and lignin structure

Panel A: Line structures of the three monomeric units of lignin. In orange, a *p*-hydroxyphenyl unit (H) that arises from *p*-coumaryl alcohol. In green, a guaiacyl unit (G) that arises from coniferyl alcohol. In purple, a syringyl unit (S) that arises from sinapyl alcohol. The α, β, and/or γ carbons contribute to bonds with other monomers and with substituted oxygens. Panel B: Interconnected lignin structure adapted from (19).

One approach to lignin valorization takes advantage of the natural process of microbial lignin catabolism (20). In nature many fungi and bacteria utilize the aromatic components of lignin as carbon and energy sources channeling the carbon into intermediates of the tricarboxylic acid (TCA) cycle. These microbes have catabolic pathways that convert the various lignin-derived aromatics into a few conserved substrates which undergo aromatic ring cleavage, typically protocatechuate (PCA), catechol, gallate, or 3-O-methylgallate (21, 22). This process is often thought of as a biological funnel that channels many different aromatics into a few ring-cleavage

substrates (**Figure 1.2**) (8). In theory, a microbe could bring all lignin-derived aromatics into central carbon metabolism, remediating issues of lignin stream complexity. One limitation of this method is that no single microbe can catabolize all components of lignin. Another challenge is to redirect the carbon flow from central metabolism to a valuable product that can be extracted and sold.

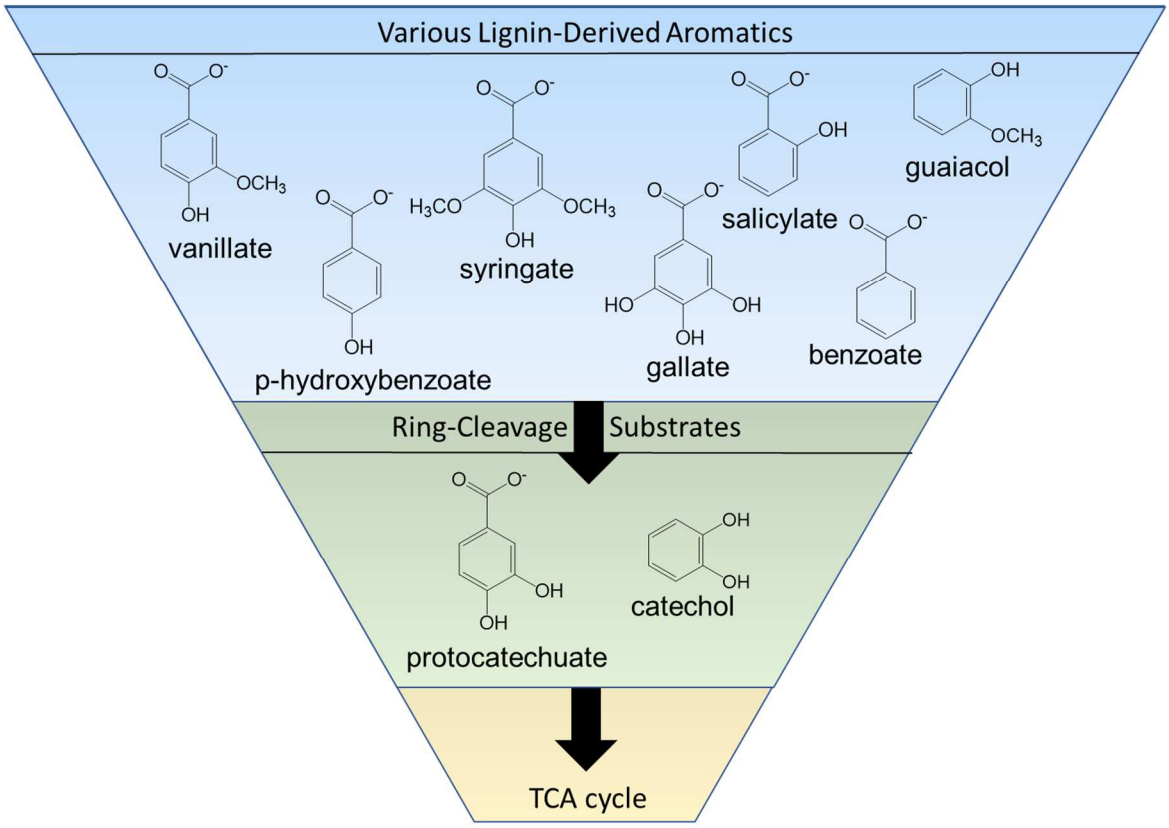


Figure 1.2: Biological catabolic funneling

Aromatic compound catabolism resembles a funnel, since many differently substituted aromatic compounds (upper blue region) are each converted into one of a very limited number of single ring-containing compounds that can be enzymatically cleaved (middle green region). Protocatechuate and catechol are the most common ring-cleavage substrates, and all aromatic

compounds degraded by *A. baylyi* ADP1 are converted to one of these compounds. After ring-cleavage, multi-step conserved catabolic pathways channel metabolite into the TCA cycle.

Salvachúa and co-authors sought to identify microbes suitable for lignin valorization by using an actual biomass feedstock and studying the depolymerization and catabolic activities of various bacteria (3). Five bacteria were identified as being especially promising: *Amycolatopsis sp*., *Pseudomonas putida* KT2440, *Rhodococcus jostii*, *Pseudomonas putida* mt-2, and *Acinetobacter baylyi* ADP1. To improve their use for lignin valorization, metabolic engineering can be used to increase the number of lignin-derived aromatics consumed and to enhance end-product formation.

P. putida KT2440 is one model organism for aromatic compound degradation that is amenable to genetic manipulation. This bacterium has been used to convert lignin-derived streams into pyruvate, adipic acid, muconic acid, and fifteen other metabolites that can be used as precursors to industrially relevant materials (5, 23-27). Although fewer studies have focused on *A. baylyi* ADP1, this bacterium naturally accumulates industrially relevant products such as wax esters and triacylglycerols (28, 29). Researchers have not only demonstrated but improved upon *A. baylyi*'s ability to produce wax esters from lignin rich streams (4, 30-32). Recently, Arvay and coauthors also engineered and demonstrated the production of mevalonate, an important chemical in cosmetics industries and a precursor to some polyesters, from lignin rich streams in *A. baylyi* as well (33).

To complement these advancements in using ADP1 to generate valuable chemicals, my doctoral research aims to increase the number of lignin-derived aromatic compounds that can be metabolized. The combination of increased consumption of lignin-derived compounds with

improved accumulation of end products is key to making this type of lignin valorization a reality. The powerful genetic system of ADP1 enables the chromosomal introduction of large fragments of DNA that can encode entire catabolic pathways. In the following sections of this review, I will discuss the unique abilities of *A. baylyi* ADP1 that make it a good platform organism for biotechnology as well as a good organism for undergraduate laboratory courses.

Acinetobacter baylyi ADP1: An ideal organism for synthetic biology and teaching

A. baylyi ADP1 is a strictly aerobic, nonpathogenic, nonmotile, gram-negative, gamma-proteobacterium. It was first isolated in the 1960's by Juni and Janik through growth on 2,3-butanediol and named *Acinetobacter calcoaceticus* BD4 (for butanediol) (34, 35). After subjecting the bacterium to ultraviolet radiation, an unencapsulated derivative was distinguished by its exceptional competence for natural transformation (36). This strain, BD413, was later called ADP1 and its species designation was reassigned as *baylyi* (37). ADP1 has a fully sequenced genome of 3.6 Mb, a G+C content of 40%, and simple nutritional requirements. It grows quickly and can utilize a wide variety of compounds for carbon and energy sources (38). The distribution of coding sequences (CDS) in the chromosome is asymmetric with about 60% on the leading strand and 40% on the lagging strand.

A. baylyi ADP1 readily takes up linear DNA from its environment through a unique DNA uptake system, which has components similar to type IV secretion systems. Single-stranded DNA is taken into the cell (39, 40). Transformation depends on Mg^{2+}, Mn^{2+} or Ca^{2+}, and it is inhibited by a metal chelator, EDTA (39). ADP1 competence increases during log growth phase and rapidly declines when cultures enter stationary phase (41). After DNA uptake, cells readily incorporate the DNA into the chromosome via homologous recombination. Chromosomal

incorporation of non-homologous DNA can occur, but at very low rates (42). When the target DNA is flanked by sequences identical to the ADP1 chromosome, allelic replacement occurs with high frequency that can vary based on the length of the sequence identity and on the genetic context (43). ADP1's natural transformation ability has been used to improve DNA cloning strategies that may be a bottleneck in high-throughput work flows if done in *Escherichia coli* (44). **Figure 1.3** shows allelic replacement in *A. baylyi* whereby chromosomal changes are easily engineered via transformation and homologous recombination using linear donor DNA.

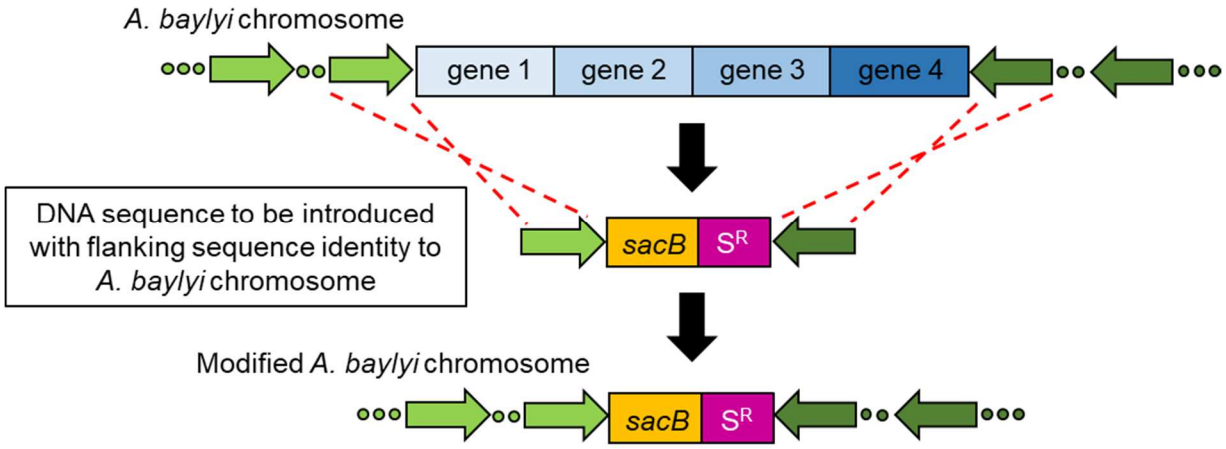


Figure 1.3: Allelic replacement in *A. baylyi* ADP1

Linear DNA can replace the corresponding region of the chromosome if flanking sequence identity enables recombination events to occur upstream and downstream of a target region. In this example, four native genes are removed and replaced with a cassette. The Streptomycin/Spectomycin resistance (S^R) gene can be selected. A second gene, encoding levan sucrase (*sacB*), enables subsequent counter-selection since the SacB enzyme is lethal in the presence of sucrose.

Genetic malleability may contribute to chromosomal organization in ADP1 as well as being important in bioengineering applications. It has been argued that the supraoperonic clustering of catabolic genes evolved under adaptive conditions where the chromosome can expand and contract significantly (43, 45). There are five islands of catabolic diversity (38), chromosomal regions with clustered genes responsible for ADP1's catabolic versatility (46, 47). This type of arrangement may reflect horizontal gene transfer because genes are more likely to be transferred together if they are neighbors than if they are in distant regions of the chromosome. Additionally, gene clustering could allow rapid environmental adaptation by increasing expression through chromosomal gene amplification (duplication and divergence) (48). Such transient changes are typically difficult to investigate. However, the natural transformation system of ADP1 facilitates the characterization of chromosomal rearrangements and enabled the Neidle lab to study spontaneous duplication events, some involving an insertion element, IS*1236* (43, 45, 49).

The natural transformation system of ADP1 also facilitated the development (or modification) of genetic tools to increase the efficiency of molecular biology approaches compared to what is possible in traditional model organisms, such as *E. coli*. For example, Súarez and coauthors recently developed a method for rapid deletion, modification, and additions to the *A. baylyi* genome, termed "Golden Transformation" as it uses linear DNA generated by Golden Gate Assembly to create these modifications (12). They also generated an ADP1 transposon insertion library and used Tn-seq to determine essential and non-essential genes, expanding on the previous *A. baylyi* ADP1 single gene knock-out collection (12, 46). Furthermore, after observing that IS*1236* elements often eliminated beneficial traits in *A. baylyi,* such as competency, during extending laboratory culturing, researchers in the Barrick lab created

a strain with all IS*1236* elements removed (50, 51). This new strain, ADP1-ISx, demonstrated reduced mutation rates and higher transformation efficiency, characteristics which could be beneficial for future experimentation and methodology development in *A. baylyi* (51).

During my doctoral studies, I collaborated with Bradley Biggs and other members of Dr. Keith Tyo's lab at Northwestern University to develop and verify a genetic toolkit for *A. baylyi* ADP1 (11). This work included the development of a promoter library and ribosome binding site variants which I tested by applying them to research on aromatic compound catabolism. Additionally, Biggs developed a marker-less, scar-less Cas-9 genomic integration method for *A. baylyi* ADP1 (11). Similarity, Súarez and coauthors have reprogramed *A. baylyi*'s native CRISPR/Cas locus to validate the removal of and prevent the reacquisition of genes from the chromosome (12). A recent manuscript by Luo and colleagues details yet another new method for laboratory evolution of *A. baylyi* ADP1 termed "Rapid Advantageous Mutation ScrEening and Selection" (RAMSES) (52). I am a co-author on this manuscript, which was positively reviewed and currently in revision, having contributed to the analysis of whole genome sequencing. Using this RAMSES method, *A. baylyi* evolved to tolerate and consume increasing concentrations of the aromatic compound ferulate by stepwise adaptation to high levels of this compound (52). These recent developments display *A. baylyi* ADP1's potential for synthetic biology and biotechnology applications.

The ease of genetic experimentation makes ADP1 an ideal bacterium for experimentation by undergraduate students. Much less is known about the metabolism and physiology of ADP1 compared to traditional model organisms. There are interesting and important scientific questions that can be addressed rapidly by relatively simple laboratory studies. On this basis, an experiential laboratory class was developed, by Drs. Ellen Neidle and Anna Karls, in which

students could conduct authentic research guided by professors and teaching assistants (TAs). Offered at UGA for the first time in 2012, Experimental Microbiology Lab (MIBO4600L) is a semester-long course in which students are characterizing the functions of some of the 44 predicted LysR-Type Transcriptional Regulators (LTTRs) in *A. baylyi* ADP1 (53). Of the 16 families of transcriptional regulatory proteins, LTTRs are the most common in prokaryotes (54). In *A. baylyi*, LTTRs comprise 44 of the 174 predicted transcriptional regulators (38, 53, 54). These regulators are predicted to participate in diverse metabolic functions, and several that control aromatic compound degradation in ADP1 have been the focus of research in the Neidle lab for many years, particularly BenM and CatM (55-63). These two regulators have overlapping functions controlling a complex regulon for benzoate and catechol degradation (63, 64). In 2008, an analysis of all possible LTTRs in ADP1 predicted functions based on genome context and comparative biology (53). In the past ten years, approximately 5-10 LTTRs have been studied by students in the course. The target genes were chosen based on the ability to generate testable hypothesis concerning their functions and on the chances that novel and important information could be obtained readily. Some LTTRs have been studied during several semesters to generate sufficient data for publication.

The first paper resulting from studies done in MIBO4600L was published in 2017 and included professors, TAs, and undergraduates as coauthors (65). The research described the role of an LTTR, MdcR, in malonate utilization by ADP1. I was a TA for MIBO4600L for two semesters, guided by Drs. Neidle and Hoover, during which we characterized two LTTRs, AalR (aspartate-ammonia lyase-regulator) and DarR (D-aspartate-responsive-regulator). These LTTRs, which control aspartate metabolism in ADP1, were studied during four different semesters of the course. In addition, I completed many experiments outside of class to complete the project.

These regulatory studies complement my other doctoral studies, as all of them address various aspects of metabolic and genetic regulation in ADP1. As the first author of a manuscript, I am preparing our work for publication that is described in **Chapter 3** of this thesis.

Aromatic compound catabolism: pathways and regulation in *A. baylyi* ADP1

A. baylyi and other *Acinetobacter* species degrade many naturally occurring aromatic compounds. Strain ADP1 has enzymes for the catabolism of vanillate, hydroxycinnamates (caffeate, coumarate, *p*-coumarate, and ferulate), benzoate, *p*-hydroxybenzoate, quinate, chlorogenate, salicylate esters, aryl esters, and anthranilate (38, 66-68). Each of these compounds is converted into protocatechuate or catechol and catabolized through the β-ketoadipate pathway (**Figure 1.4**) (69). This pathway is conserved in diverse eukaryotic and prokaryotic microbes, although there are some significant variations, especially in regulation.

As shown in Figure 1.4, there are two parallel pathway branches. In contrast to other bacteria, ADP1 uses isozymes to catalyze the same three final reactions of the pathway; one set of enzymes degrades metabolites generated from protocatechuate (PCA) cleavage and one set degrades the same metabolites if they are generated by the cleavage of catechol. The requirement for isozymes reflects differential regulation of the *cat* and *pca* genes encoding these enzymes. The *cat* and *pca* genes are located in different regions of the chromosome. The genes encoding the PCA branch are located in catabolic island IV as part of a large *dca-pca-qui-pob-hca* cluster (38). In contrast, genes encoding the catechol branch are located in catabolic island III as part of the *are-sal-ben-cat* cluster (38). Two LTTRs, CatM and BenM, control a complex regulon in which they play overlapping roles. Both regulators respond to *cis,cis*-muconate as an effector.

BenM additionally responds to benzoate, with synergistic effects resulting from the presence of both effectors. These LTTRs activate transcription of the *ben* and *cat* genes, and they repress transcription of the *pca* genes **(Figures 1.4 and 1.5: Panel A)** (53, 58, 59, 64, 70).

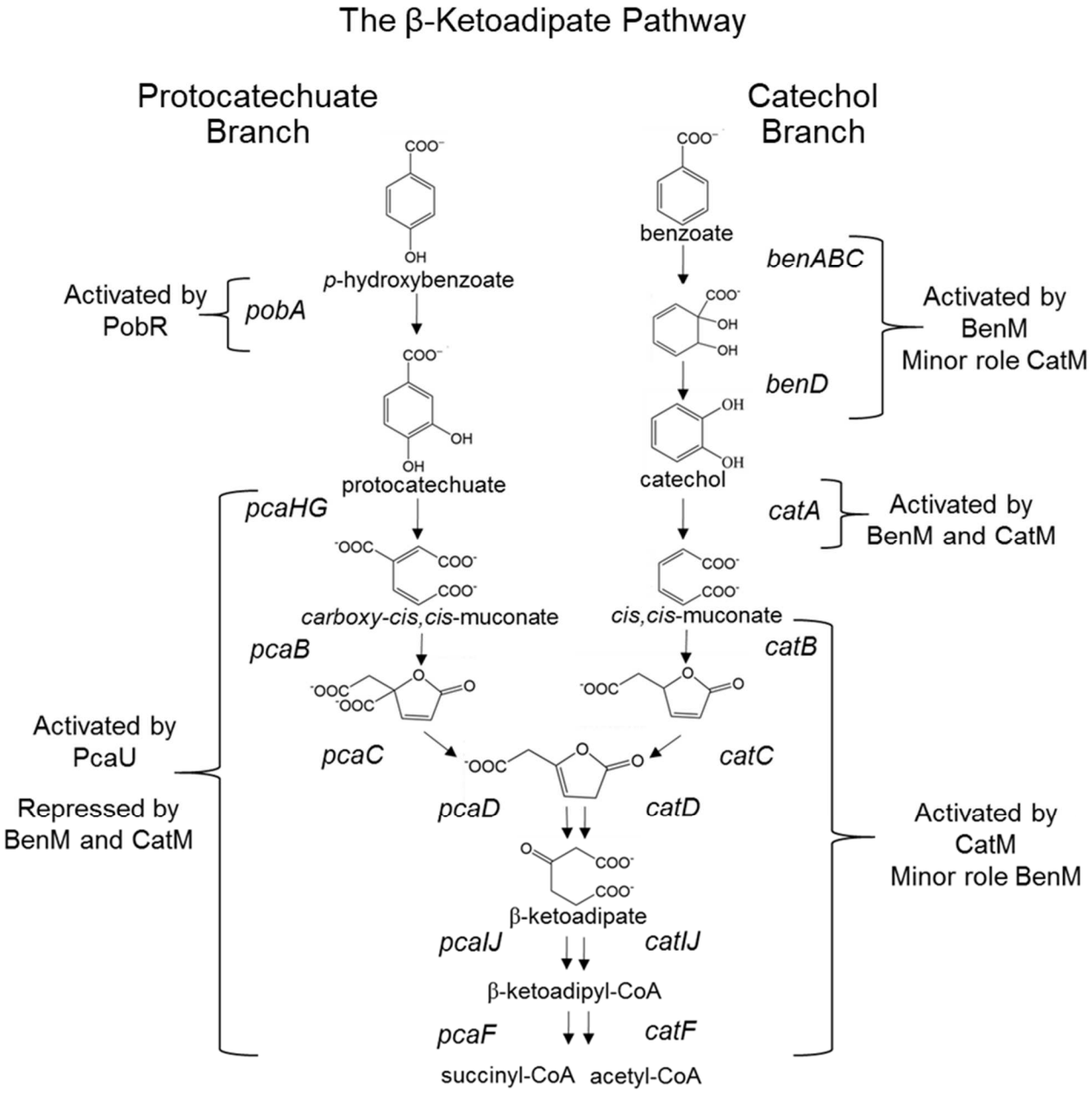


Figure 1.4: The β-Ketoadipate pathway of *A. baylyi* ADP1

Left side: Protocatechuate branch with corresponding transcriptional regulators indicated Right side: Catechol branch with corresponding transcriptional regulators indicated.

During growth on a mixture of benzoate and POB, ADP1 consumes almost all the benzoate before catabolizing POB (**Figure 1.5: Panel B)**. Previous studies demonstrated a role for BenM and CatM in this preferred use of benzoate, although one observation was puzzling. It appears that the LTTRs directly repress *pca* gene expression by binding to the promoter region of an operon encoding the PCA branch. However, such repression of the *pca* genes should not interfere with the conversion of POB to PCA, a reaction mediated by the PobA hydroxylase, which is encoded by a gene that is not in the *pca* operon. Nevertheless, such conversion failed to occur, indicating that the role of BenM and CatM might be complex. Inducer exclusion was proposed as a mechanism that might prevent PobA from hydroxylating POB. PobA expression requires this compound to enter the cell and serve as a transcriptional inducer. POB uptake is mediated by PcaK, a transporter encoded in the *pcaK* operon. Thus, repression of the *pca* operon by BenM and/or CatM might prevent POB degradation if its transport into the cell were blocked by the absence of PcaK. However, there are other known POB transport proteins in addition to PcaK, so it was not clear that *pcaK* repression could fully explain the preferential consumption of benzoate and the delayed consumption of POB when both growth substrates are available to ADP1.

During my doctoral studies, I contributed to the publication describing a new ADP1 tool kit by testing synthetic promoters (developed by Bradley Biggs) in investigations of the regulatory mechanisms underlying sequential aromatic compound consumption in ADP1 (11). These studies, published in 2020, are briefly summarized here. I sought to test aspects of the

hypothesis described above by expressing PcaK independently of the other genes in the *pcaK* operon. I deleted *pcaK* from its chromosomal position and integrated it elsewhere in the chromosome where it was expressed from constitutive promoters. When POB or PCA was provided as the sole carbon source, the engineered strain grew well. However, as shown in Figure 1.5 panel C, when benzoate was provided together with POB, the independent transcription of *pcaK* did not result in the uptake of POB and its conversion to PCA until benzoate had been consumed. These results indicate that the repression of *pcaK* is insufficient to account fully for preventing the conversion of POB to PCA.

To test additional aspects of the regulatory model, I mutated the *pca* promoter region to eliminate the repressor binding site for BenM and CatM while leaving intact a functional promoter for PCA and POB consumption. As shown in **Figure 1.5 Panel D**, eliminating the BenM/CatM binding site upstream the *pca* genes failed to abolish the hierarchical consumption pattern. These results indicate that repression of the *pca* operon by BenM and/or CatM is insufficient to account for the inhibition of POB degradation during benzoate catabolism.

My results raised questions about the roles of BenM and CatM in controlling the ADP1 preference for consuming benzoate before POB. Previous studies showed that muconate, a metabolite of benzoate degradation is required for the repression of POB utilization (71). Since BenM and CatM both respond to this effector compound, they seemed likely candidates to mediate the inhibition of POB consumption during growth on benzoate. Mutants were generated that degrade benzoate without BenM or CatM, and, in the absence of both LTTRs, benzoate and POB are consumed simultaneously (63). These results indicated that BenM and CatM play essential roles in the hierarchical consumption pattern. However, one feature of the mutants raised concerns about drawing firm conclusions. The mutants all had multiple chromosomal

copies of the *ben-cat* genetic region. It remained possible that increased gene dosage affected regulation in some unanticipated fashion.

I created new mutants that could grow on benzoate without BenM and CatM and with a single copy of the *ben-cat* chromosomal region. Normally the *ben* and *cat* genes require activation by BenM and CatM (29, 45). However, by putting these genes under the control of constitutive promoters, I was able to delete the BenM and CatM coding sequences from the chromosome while still allowing *A. baylyi* to grow on benzoate. This newly engineered ADP1-derived mutant, capable of growing on benzoate without BenM and CatM, displayed simultaneous consumption of benzoate and POB. Thus, in combination with previous studies, my results demonstrated that BenM and CatM are essential for the sequential consumption of benzoate and POB by ADP1 (**Figure 1.5 Panel E**). The role of BenM and CatM in multiple carbon source regulation may include their ability to repress the *pca* operon, but such repression must only be a portion of a more complex regulatory circuit, since my studies show that such regulation is insufficient to account for the observed catabolic patterns. The ability to alter the chromosome, relocate genes, and control transcription with a wide variety of promoter strengths now enables us to test regulatory models in novel ways that can be readily accomplished using our new tool set.

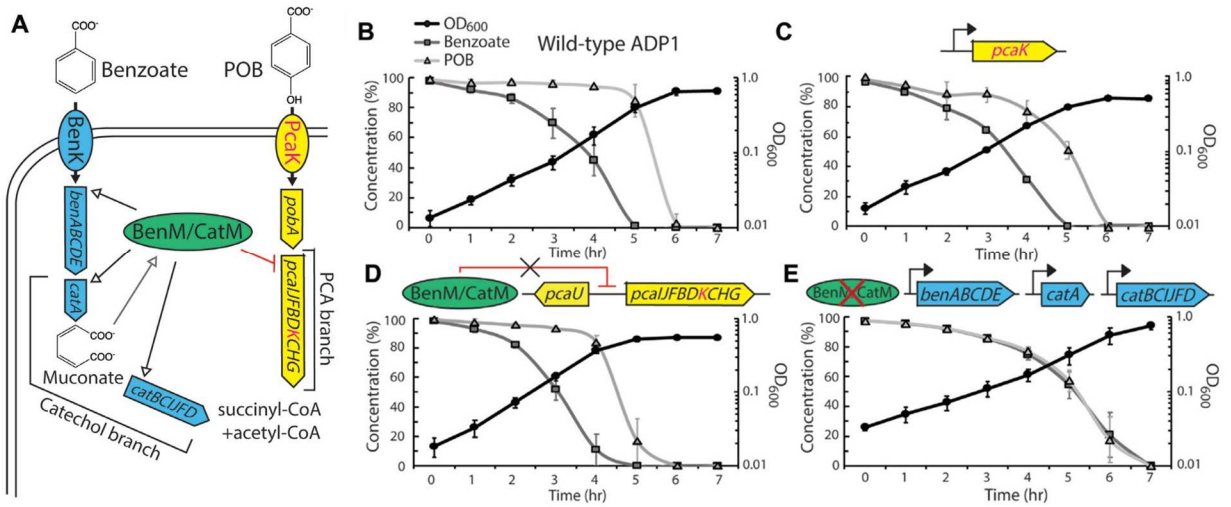


Figure 1.5: Testing the regulation of hierarchical benzoate and POB consumption

Panel A: Depiction of the regulation and metabolism of POB and benzoate by BenM and CatM. Regulators activate *ben* and *cat* genes, for benzoate and catechol catabolism, and repress the *pca* genes, for PCA catabolism. Consumption patterns of benzoate, grey squares, and POB, grey triangles, as measured by high performance liquid chromatography (HPLC) throughout growth are shown for: (Panel B) ADP1, (Panel C) an ADP1-derived strain where the transporter gene *pcaK* is transcribed constitutively, (Panel D) an ADP1-derived strain where the binding site for BenM/CatM upstream of the *pca* operon has been scrambled to prevent binding of these regulators, and (Panel E) an ADP1-derived strain without BenM and CatM and where the genes for benzoate degradation, which are normally activated by BenM and CatM, are under constitutive expression. OD_{595} (black circles) and samples for HPLC analysis were collected hourly during growth for each strain. Data shown are average of at least three biological replicates. Error bars represent standard deviation. (11)

While *A. baylyi* ADP1 catabolizes numerous aromatic compounds, there are some notable catabolic pathways it does not possess. For example, there exist other pathways for the

degradation of PCA and catechol besides the β-ketoadipate pathway. These other pathways involve the meta-cleavage of PCA and catechol, outside of the hydroxyl groups rather than between the hydroxyl groups as in the β-ketoadipate pathway (**Figure 2.1**) (72-74). Another catabolic pathway that is not present in ADP1 involves the degradation of derivatives of S-lignin. As shown in **Figure 1.1 Panel A**, there are three main monomeric units of lignin that differ in their number of substituents on the aromatic ring. ADP1 catabolizes many derivatives of H- and G-units, but it does not possess the catabolic genes for S-lignin units such as syringaldehyde, syringate, 3-O-methylgallate, or gallate (6). A canonical pathway for the degradation of these S-lignin-derived monomers is depicted in **Figure 2.1**. Additionally, recent studies characterized enzymes involved in the degradation of veratrate and isovanillate by *Comamonas testosteroni* BR6020 and *Pseudomonas putida* CSV86 (75, 76). While ADP1 cannot consume veratrate, it does catabolize an intermediate of veratrate catabolism, vanillate, utilizing a O-demethylase (VanAB) (66, 77). My doctoral research focused expanding *A. baylyi* ADP1's metabolic repertoire to include a PCA meta-cleavage pathway and a pathway for the catabolism of the veratrate and isovanillate; studies detailed in **Chapters 2 and 4** respectively.

Metabolic expansion using EASy laboratory evolution

The genetic malleability of *A. baylyi* can be exploited to alter its chromosome and introduce non-native catabolic pathways for new aromatic carbon sources. However, *A. baylyi* may not be able to utilize such foreign pathways optimally. Many issues could underlie ineffective utilization such as inappropriate transcriptional or translational expression levels, rare codon usage, promoter or ribosome binding site incompatibility, intermediate metabolite toxicity, and/or metabolic flux problems (78, 79). The EASy (Evolution by Amplification and

Synthetic biology) method of laboratory evolution developed in the Neidle lab can allow *A. baylyi* to overcome these barriers by increasing the amount of gene product produced to levels that enable growth. In this method, a precise duplication event and subsequent amplification of a specific region of the ADP1 chromosome is initiated (**Figure 1.6**) (80). In nature, the process of gene duplication and amplification occurs infrequently and randomly (81, 82). In contrast, the EASy method produces an immediate, designed amplification event, giving the bacterium the ability to adapt quickly to its environment. Changes in gene dosage serve as a rudimentary form of regulation. Under selective pressure, multiple gene copies can be maintained in a head-to-tail array.

This process can be employed with chromosomally inserted foreign genes and pathways to allow ADP1 to grow on new carbon and energy sources. With continuous selective pressure, the evolutionary process will occur as mutations accumulate within the amplicon or elsewhere in the genome. Because amplification is an unstable, dynamic process, when beneficial mutations occur that allow better growth on the desired carbon source, a new allele may be selected and retained on the chromosome. The whole process can occur in a relatively short time span, permitting the rapid adaptation of ADP1 to new carbon sources (80).

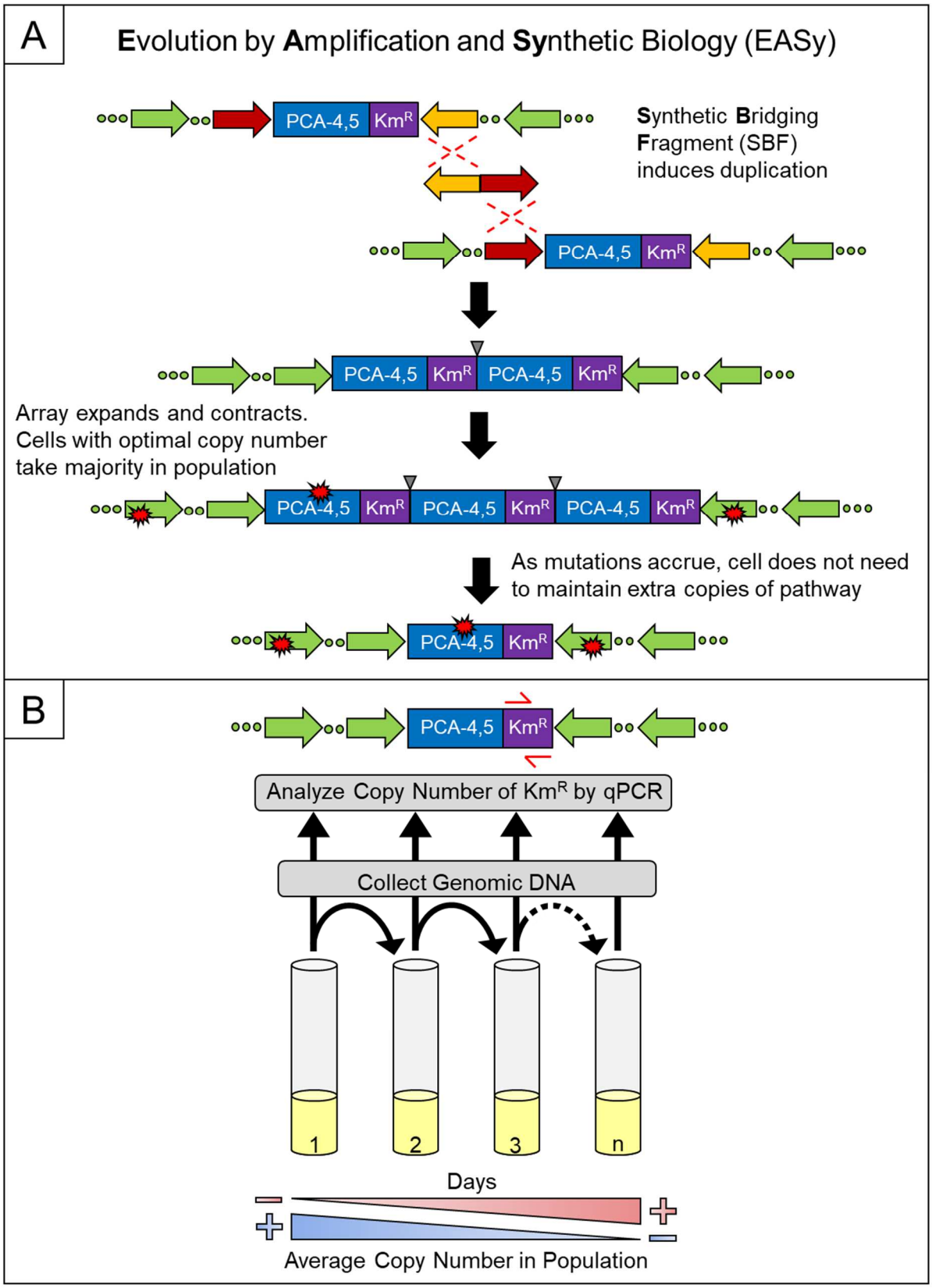


Figure 1.6 Evolution by amplification and synthetic biology (EASy)

Panel A: Schematic depiction of EASy. A synthetic bridging fragment is added to produce a targeted duplication of the desired region of the *A. baylyi* chromosome. Selective pressure chooses cells with the optimal number of copies to growth with the desired phenotype. Beneficial mutations (red stars) occur during continuous culturing which allow the cells to require fewer extra copies of the amplicon. Panel B: Depiction of the analysis of EASy evolving populations. As populations are cultured, genomic DNA is routinely collected, and the copy number of a gene in the amplicon is analyzed by qPCR (quantitative). Amplicon copy number can be followed and used as a proxy for the selection of beneficial mutations.

EASy has previously been used in evolutionary studies to allow ADP1 to grow on the aromatic compounds guaiacol and terephthalic acid as sole carbon sources following the introduction of genes originating in *Amycolatopsis sp.* ATCC 39116 and *Comamonas sp*. E6 respectively (80, 83). In the guaiacol experiments, genes encoding an O-demethylase enzyme, GcoAB, were introduced into the chromosome, and cells were evolved to grow on guaiacol as the sole carbon source by the EASy method. GcoAB converts guaiacol into catechol, which can then be catabolized by the native ADP1 β-ketoadipate pathway (84). Several evolved strains formed a fusion protein of GcoA with the native CatA, whose coding sequence was located upstream of the insertion site for *gcoAB*. Researchers then put this fusion protein into the chromosome of *Pseudomonas putida* KT2440 along with *gcoB* and discovered that it transferred the ability to utilize guaiacol as a carbon source in this organism as well. This study demonstrated that genes optimized through EASy in ADP1 can be utilized in other aromatic compound-degrading bacteria as well (80, 85). In my doctoral studies described in **Chapter 4**, EASy was employed in a similar fashion to allow *A. baylyi* to utilize veratrate and isovanillate as sole carbon sources through the introduction of genes from *Pseudomonas putida* CSV86.

Furthermore, my doctoral research demonstrated that EASy can be used to optimize growth for thermal tolerance, **Chapter 2**. Following the introduction of a predicted pathway for protocatechuate (PCA) degradation from *Pseudoalteromonas atlantica* T6c, *A. baylyi* was able to grow without any laboratory evolution on PCA or POB (*p*-hydroxybenzoate) as the sole carbon source. However, strains with the foreign pathway were not able to grow on PCA or POB at temperatures above 25 °C and had reduced growth rates compared to the native strain. *A. baylyi* is typically grown at 30 or 37 °C, so I sought to optimize growth through this foreign PCA degradation pathway by evolving it for thermal tolerance by the EASy method. After inducing a precise amplification of the foreign genes, I was able to select colonies that grew on PCA and POB at 37 °C. Eleven of these isolates were cultured continuously on minimal media with PCA or POB as the sole carbon source at 37 °C for about six months. All strains displayed a drastically reduced gene dosage at the end of this time period as detailed in **Chapter 2**. This result led us to believe that beneficial mutations occurred in the chromosomes that increased thermal tolerance in the evolved *A. baylyi* strains.

While various approaches were used to assess mutation significance, mutations were initially identified by whole genome re-sequencing (WGS). Individual colonies were isolated from the evolving populations at the week when they had the lowest copy number. Whole genome re-sequencing (WGS) was then used to assess genetic changes compared to the parent strain. WGS revealed a variety of mutations that might contribute to the acquisition of new traits, primarily thermal tolerance. In the following sections of this chapter, background information is provided relevant to some of the loci where mutations were discovered. As a first step in understanding how these mutations might impact the selected phenotype, I sought to understand previous characterization of these and related genes in ADP1 and other bacteria.

These loci included ACIAD_RS02085, which was annotated as *ptsP* indicating a possible role in a **p**hospho**t**ransferase **s**ystem, or PTS. The loci ACIAD_RS12795 and ACIAD_RS12800 annotated as *groEL* and *groES* respectively were included in spontaneous amplifications in multiple strains, suggesting that increasing their expression in the cell could be beneficial. Multiple loci encoding peptidylprolyl isomerase and chaperone domains contained mutations including ACIAD_RS02420, annotated as *tig* (trigger factor), ACIAD_RS05440, annotated as *slyD*, ACIAD_RS04740, annotated as *fbp*, and ACIAD_RS06505, annotated as *ppiD*. Lastly, included is a discussion of the ClpAP protease. A mutation was found in the *A. baylyi* locus ACIAD_RS06285, designated as *clpA*, in these EASy experiments. Similarly, *clpA* mutations were identified in *A. baylyi* teaching experiments involving aspartate metabolism in *A. baylyi*, as described in **Chapter 3**. Finding similar mutations in *A. baylyi* strains in multiple different studies, implies an important role for the ClpAP protease. Therefore, background information on this protease in provided to help define the role it could be playing in the selected phenotypes from evolutionary studies in *A. baylyi*.

Phosphotransferase systems (PTS) and metabolic regulation

Following laboratory evolution of *A. baylyi* growing through a foreign pathway, **Chapter 2**, mutations were discovered in ACIAD_RS02085, a gene annotated as *ptsP*. The PtsP, phosphoenolpyruvate-protein phosphotransferase, name was first used to indicate an *E. coli* homolog that differs from the first component, enzyme I (EI), in canonical phosphotransferase systems (PTS) used for bacterial sugar uptake (86). Subsequently, PtsP was shown to participate in a distinct type of three-component PTS, composed of proteins PtsN, PtsO, and PtsP. Both the canonical sugar-transport PTS and the three-component system, called PTS^{Ntr}, because of a

connection to nitrogen metabolism in some systems, are described in the following section. However, PtsP of ADP1 appears not to be involved in the same three-component system as its namesake, since there is no homolog of PtsN in ADP1. PtsP in ADP1 has not previously been studied, but the frequent, independent selection of mutations in *ptsP* in my studies indicates a previously undescribed role for a novel phosphotransferase system in *A. baylyi*.

Of the eleven evolved populations in my experiments described in **Chapter 2**, a single colony was selected from each and sent for whole genome re-sequencing. Seven of the eleven independently evolved isolates had a mutation in ACIAD_RS02085, and of these, six were unique mutations. Below is a table of these mutations (**Table 1.1**). Based on where they occur, we predict that all mutations would result in loss-of-function of PtsP, as illustrated by an alignment of the *A. baylyi* deduced protein product relative to a characterized *E. coli* PtsP homolog (**Figure 1.7**). These proteins have two domains. In the N-terminal region is a GAF domain, a conserved domain named for the first three protein families where it was discovered (87), InterPro entry IPR00318 (88). This domain may exert allosteric control over enzymatic activity and is often involved in phospho-transducing signaling pathways; it may bind cyclic nucleotides, such as cGMP, or other metabolites (89, 90). Adjacent to the GAF domain is a domain characteristic of the enzyme I-component of phosphotransferase systems, InterPro entry IPR006318 (88). This enzyme initiates a series of transfers by accepting the activated high-energy phosphate group from phosphoenolpyruvate (PEP).

Table 1.1: Summary of ACIAD_RS02085 *(ptsP)* mutations from laboratory evolution studies

Evolved Strain	Change	Nucleotide Position*	Effect on Protein
ACN2140	Δ1 bp	coding (1405/2295 nt)	C deletion in codon 469 causes a frameshift and a premature truncation after amino acid residue 498 Abolishes binding sites, metal binding sites, and an active site residue C680
ACN2429	Δ39 bp	coding (1069-1107/2295 nt)	In frame deletion of 13 amino acid residues after A354 Abolishes active site residue H367
ACN2432	C→T E340K (GAA →AAA)	452,510	Amino acid change that replaces a conserved amino acid (E) with a dissimilar residue (K)
ACN2433	Δ1 bp	coding (1405/2295 nt)	C deletion in codon 469 causes a frameshift and a premature truncation after amino acid residue 498 Abolishes binding sites, metal binding sites, and an active site residue, C680
ACN2434	+A	coding (2063/2295 nt)	insertion of A in codon 689 causes a frameshift and a premature truncation after amino acid 706 Frameshift changes protein sequence and binding/active site residues
ACN2437	Δ15 bp	coding (2032□2046/2295 nt)	5 amino acids removed by in-frame deletion after residue V677 including S678-I679-C680-G681-E682 Abolishes an active site residue, C680
ACN2438	T→C	intergenic (□111/□74)	Mutation in operator-promoter region of *rppH*/*ptsP* s may affect expression levels

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Figure 1.7: Protein sequence alignment of *A. baylyi* predicted PtsP to a characterized homolog from *E. coli*

The sequence of PtsP from ADP1 (Uniprot ID Q6FEW8) is aligned above that of an *E. coli* homolog (Uniprot ID P37177). Colored residues were identified in studies of the *E. coli* PtsP as active site (red), metal binding (pink), PEP binding site (green) residues, and GAF domain (yellow). Markings underneath each residue in the alignment indicate exact match (*), conserved replacement (:), or semiconserved replacement (.) (91).

As the basis of comparison for variations, the canonical glucose PTS is depicted in **Figure 1.8**, although ADP1 does not have this system. The sole sugar PTS in ADP1 mediates fructose uptake. Surprisingly, ADP1 cannot use fructose as a sole carbon source (38, 92). The EI enzyme of the fructose PTS is encoded by a gene distinct from *ptsP*. The EI components of sugar PTS do not have a GAF domain.

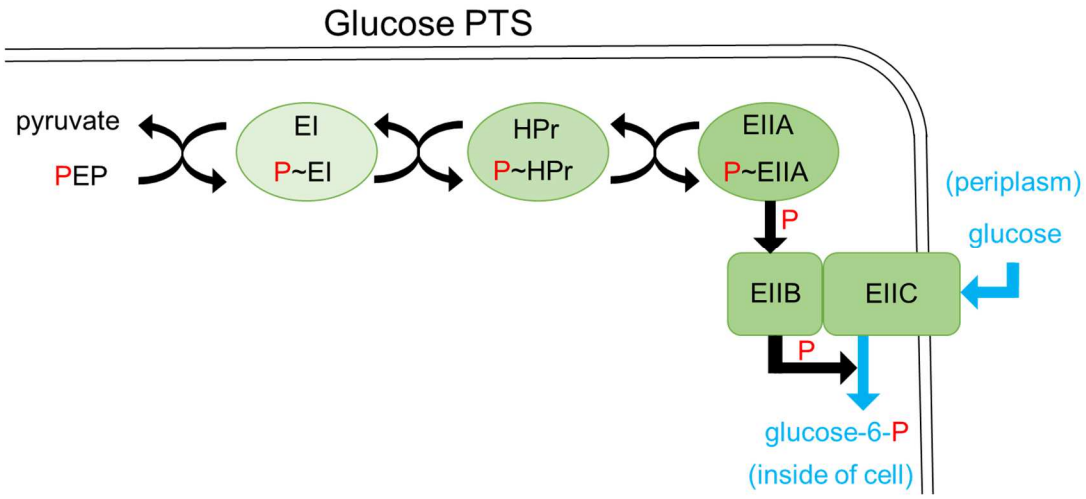


Figure 1.8: Canonical glucose PTS signaling cascade

As shown, a phosphate (red P) from PEP is transferred first to enzyme I (EI), and then to a protein containing a critical histidine residue (HPr). Next, the phosphate is transferred to enzyme II (EII) which has three components (A, B, and C) that may be separate polypeptides or may be different domains of protein fusions (93, 94). EIIB phosphorylates the sugar as it crosses the inner membrane into the cytoplasm via the transport protein (EIIC). The formation of glucose-6-phosphate sequesters the sugar inside the cell by preventing efflux, allowing it to accumulate against a gradient.

Bacteria (especially those in the *Enterobacteriaceae*, *Vibrionales* and *Firmicutes*), often have multiple PTS for the uptake and regulated consumption of different sugars. Consistent with

ADP1 being unable to consume many sugars, genome sequence analysis revealed only one sugar PTS. This fructose PTS, which was noted above, is composed of two proteins. A single polypeptide, FruA, contains domains corresponding to EI, HPr, and EIIA. FruB contains two domains corresponding to EIIB and EIIC (**Figure 1.9**). Although the glucose PTS was initially the best characterized, the fructose system has been proposed to be the evolutionary ancestor of all PTS (95).

A second type of bacterial PTS was discovered that is not associated with the uptake of sugar but appears to be involved in carbon versus nitrogen metabolism, as reviewed (96). This so-called nitrogen PTS (PTS^{Ntr}), first identified in *E. coli*, lacks EIIB and the transmembrane transport component (EIIC) (97). The genes encoding this abridged system were designated *ptsN* (encodes $EIIA^{Ntr}$), *ptsO* (encodes an HPr homolog, Npr), and *ptsP* (encodes EI^{Ntr}). In this system, it is unclear where the EII protein donates the activated phosphate. Moreover, the exact role of this system in nitrogen metabolism is controversial and may not be identical in all bacteria that carry homologs. Nevertheless, complex effects of sugar uptake systems and PTS^{Ntr} are evident in diverse processes including carbon catabolite repression (98), virulence (99-101), biofilm formation (102), and potassium homeostasis (103, 104).

The presence of a GAF domain in the EI^{Ntr} is notable and has been associated with fine-tuning metabolism and homeostasis. This regulatory domain can integrate signals from small molecules and nucleotides (second messengers) and effect allosteric control. In some cases, but not others, deletion of the GAF domain makes EI^{Ntr} unable to self-phosphorylate (105-107). The presence of a GAF domain as well as an EI-like domain encoded by ACIAD_RS02085 underlies its PtsP designation. However, unlike the three-component PTS^{Ntr} ADP1 has no gene resembling *ptsN* (encodes $EIIA^{Ntr}$). Thus, there no paralogs of the EII A, B, or C domains of FruA and

FruB. Searching for homologs of HPr revealed one gene, ACIAD_RS13825, which encodes an Hpr/Npr-like protein. I tentatively designated this gene *ptsO* to indicate a possible interaction of its encoded protein with PtsP, although there is no evidence that PtsO and PtsP participate together in a PTS. This gene is predicted to be co-transcribed as part of a large operon of unclear function (**Figure 1.9**).

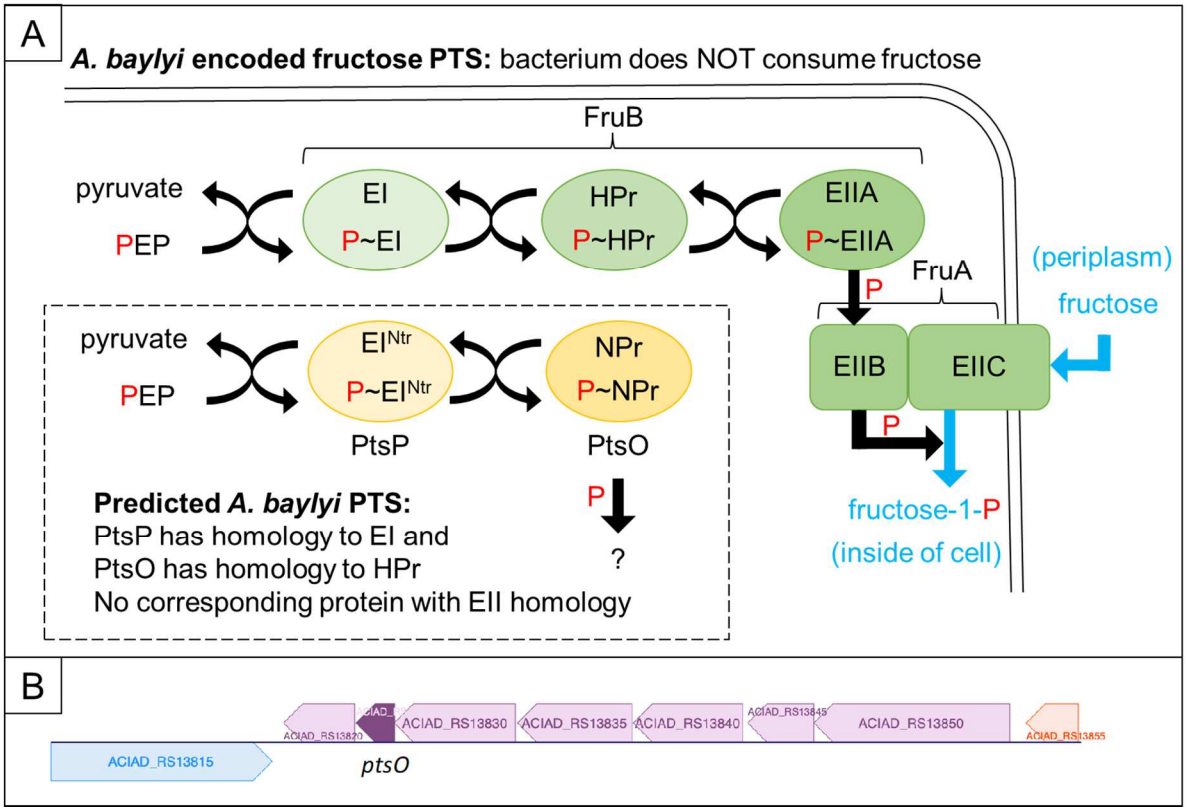


Figure 1.9: PTS encoded in *A. baylyi* ADP1

Panel A. Although ADP1 cannot grow on fructose, the FruAB PTS was shown to be capable of taking up fructose (92). The domains of FruA and FruB are discussed in the text. ADP1 encodes only two other PTS-like proteins, PtsP (a homolog of EI^{Ntr}), and PtsO (a homolog of Hpr/Npr). The function of PtsP of ADP1 is unknown, and its possible interaction with PtsO, the protein product of ACIAD_RS13825, is hypothetical. Panel B. the gene designated *ptsO* is predicted to

be transcribed as part of a large operon of unknown function, as shown in this image of ADP1 chromosomal organization captured from the Biocyc web site (108).

The possibility that a two-component PTS, involving PtsP and PtsO, is functional in ADP1 is supported by studies of a similar system in *Legionella pneumophila* (109). A partial PTS^{Ntr}, lacking EII, was shown in *L. pneumophila* to mediate a phosphorelay using *in vitro* and *in vivo* experiments. As in ADP1, it is unknown what may serve as a final phosphate acceptor. However, researchers postulated that the PTS might govern a complex regulatory circuit via signal integration involving two-component transcriptional regulators, which also carryout phosphotransfers. Several two-component transcriptional regulators were investigated, and a connection was demonstrated between the PtsP-PtsO cascade and a two-component transcriptional system, PmrA (a response regulator) and PmrB (a sensor histidine kinase). Despite experimental efforts, it was not possible to conclude if PtsO (the Npr component) could directly phosphorylate PmrA, yet PtsO (Npr) did activate the PmR regulon in a fashion that was dependent on this response regulator (109).

Based on the results in *L. pneumophila*, we are considering the possibility that PtsP in ADP1 exerts its metabolic effects in connection with two-component transcriptional regulators. *A. baylyi* has several such two-component transcriptional regulators (**Table 1.2**). Of special interest is the GacA/GacS two-component system. In ADP1 GacA/GacS regulates the transcription of non-coding RNAs that interact with CsrA, a small protein that modulates the translation of carbon-utilization genes. We have recently discovered that CsrA controls aromatic compound catabolism in ADP1 (Emily McIntyre, unpublished data). In ongoing aromatic-compound degrading studies using the EASy method, several *csrA* and *gacA/gacS* mutaations have been selected. Future studies are needed to investigate the role of PtsP in ADP1 and

whether there is any connection between this protein and GacA/GacS or any two-component transcriptional regulator in this bacterium.

Table 1.2: **Two-component transcriptional systems in *A. baylyi* ADP1**

Locus Tag	Gene	Description	NCBI-protein ID
ACIAD_RS06315	*glnL*	PAS domain-containing sensor histidine kinase	WP_011182272.1
ACIAD_RS06310	*glnG*	DNA-binding transcriptional regulator NtrC	WP_004925683.1
ACIAD_RS02855	*baeS*	two-component sensor histidine kinase	WP_004919832.1
ACIAD_RS02860	*baeR*	DNA-binding response regulator	WP_004919829.1
ACIAD_RS15330	*envZ*	two-component sensor histidine kinase	WP_004923640.1
ACIAD_RS15325	*ompR*	DNA-binding response regulator	WP_004923643.1
ACIAD_RS03330		two-component sensor histidine kinase	WP_004922536.1
ACIAD_RS03325		DNA-binding response regulator	WP_000076440.1
ACIAD_RS13430	*kdpD*	sensor histidine kinase KdpD	WP_004929597.1
ACIAD_RS13425	*kpdE*	DNA-binding response regulator	WP_004929596.1
ACIAD_RS10245		two-component sensor histidine kinase	WP_004927854.1
ACIAD_RS10250		DNA-binding response regulator	WP_004927857.1
ACIAD_RS01225	*gacS*	two-component sensor histidine kinase	WP_011182006.1
ACIAD_RS01230	*gacA*	DNA-binding response regulator	WP_004920720.1
ACIAD_RS01220		sigma-54-dependent Fis family transcriptional regulator	WP_004920725.1
ACIAD_RS07490		two-component sensor histidine kinase	WP_004924986.1
ACIAD_RS07495		DNA-binding response regulator	WP_004924983.1
ACIAD_RS13370		two-component sensor histidine kinase	WP_004929571.1
ACIAD_RS13375	*qseB*	DNA-binding response regulator PmrA	WP_004929572.1
ACIAD_RS06565		two-component sensor histidine kinase	WP_011182281.1
ACIAD_RS06570		DNA-binding response regulator	WP_004925537.1
ACIAD_RS16095	*phoR*	phosphate regulon sensor histidine kinase PhoR	WP_004923226.1
ACIAD_RS16090	*phoB*	DNA-binding response regulator	WP_004923228.1
ACIAD_RS11360		hybrid sensor histidine kinase/response regulator	WP_004928507.1

ACIAD_RS11355		DNA-binding response regulator	WP_004928503.1
ACIAD_RS15690		DNA-binding response regulator	WP_004923453.1
ACIAD_RS02715		two-component sensor histidine kinase	WP_004919910.1
ACIAD_RS13890	*barA*	hybrid sensor histidine kinase/response regulator	WP_004924459.1
ACIAD_RS14390		histidine kinase	WP_004924206.1

* Regulator pairs are shaded together in table

Essential protein chaperone GroE

During laboratory evolution of *A. baylyi,* strains with increased thermal tolerance were found to have mutations in loci ACIAD_RS12795, annotated as *groEL*, and ACIAD_RS12800, annotated as *groES*, as detailed in **Chapter 2**. These two genes, *groEL* and *groES* together encode the protein chaperone GroE. The L and S designations indicate the large and small subunits of the protein complex. Because a selective pressure for increased thermal tolerance was used in my evolution studies, changes to protein chaperones appear to affect the thermal stability of proteins within my evolved cells. **Table 1.3** outlines a mutation discovered in *groES* as well as two independently evolved chromosomal amplifications encompassing *groEL* and *groES*. One of these amplifications appears to cause a chromosomal duplication partway through the *groEL* coding sequence and extending past the *groES* coding sequence. The second amplification, which appears to have three or four chromosomal copies, encompasses a larger chromosomal region starting in a gene upstream of *groEL* and ending in the gene downstream *groES*. In *A. baylyi* ADP1 *groEL* and *groES* are essential, but they have not been studied outside of homology searches (92). To better understand the role these *groEL/groES* changes might be having in my

evolved strains, I will summarize the GroE chaperone mechanism as well as previous studies implicating the GroE chaperone in metabolite toxicity tolerance.

Table 1.3: Summary of *groES* and *groEL* changes from laboratory evolution studies

Evolved strain	Gene locus ID (name)	Position*	Predicted description	Mutation type	DNA change	Protein effect
ACN2434	ACIAD_RS12800 *(groES)*	2,776,999	GroES	substituition	A→G	*97Q

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Evolved strain	*A. baylyi* loci involved in amplification of *groEL* and *groES*
ACN2437	ACIAD_RS12795 (*groEL*) and intergenic region between ACIAD_RS12800 (*groES*) and ACIAD_RS12805
ACN2438	ACIAD_RS12785 and ACIAD_RS12805 (*groEL* and *groES* are between these loci within the amplicon)

Exactly how a protein forms from a string of amino acids to a fully folded, active protein with tertiary structure is one of the great questions in biology. Anfinsen's dogma in molecular biology, earning a Nobel prize in chemistry in 1973, posits that a protein's tertiary structure is defined by its primary amino acid sequence (110, 111). Small differences in Gibbs free energy drives such protein folding and many *in vitro* experiments corroborate this theory (112). However, the demonstrated self-folding of proteins such as bovine pancreatic RNase is too slow to be effective *in vivo* (113). Additionally, many proteins are prone to aggregation or misfolding (114). Thus, numerous proteins require the help of molecular chaperones to hasten and/or accomplish folding into their active conformations (115). One of the best studied molecular chaperones, and the focus of this literature review section, is GroE (116-122).

The molecular chaperone subcategory of chaperonins is ubiquitous across all domains of life (116). Chaperonins are ATP-dependent molecular chaperones composed of heat shock protein (Hsp) 60 kD subunits that oligomerized into ring structures (123). Hsp proteins were named for their role in the cellular stress response to heat. GroEL, growth essential large, and

GroES, growth essential small, are the two protein subunits of the essential bacterial chaperonin GroE (116). Characteristic of chaperonins, GroEL subunits oligomerize into two seven-membered rings, with a double barrel shaped quaternary structure, each barrel of which can accommodate proteins up to 57 kilodaltons is size (**Figure 1.9 Panel A**) (124). GroES subunits oligomerize into a single seven-member ring that acts as a cap to the GroEL ring once a protein has entered the cavity for folding. Current understanding of the GroE mechanism is depicted in **Figure 1.9 Panel B** (125).

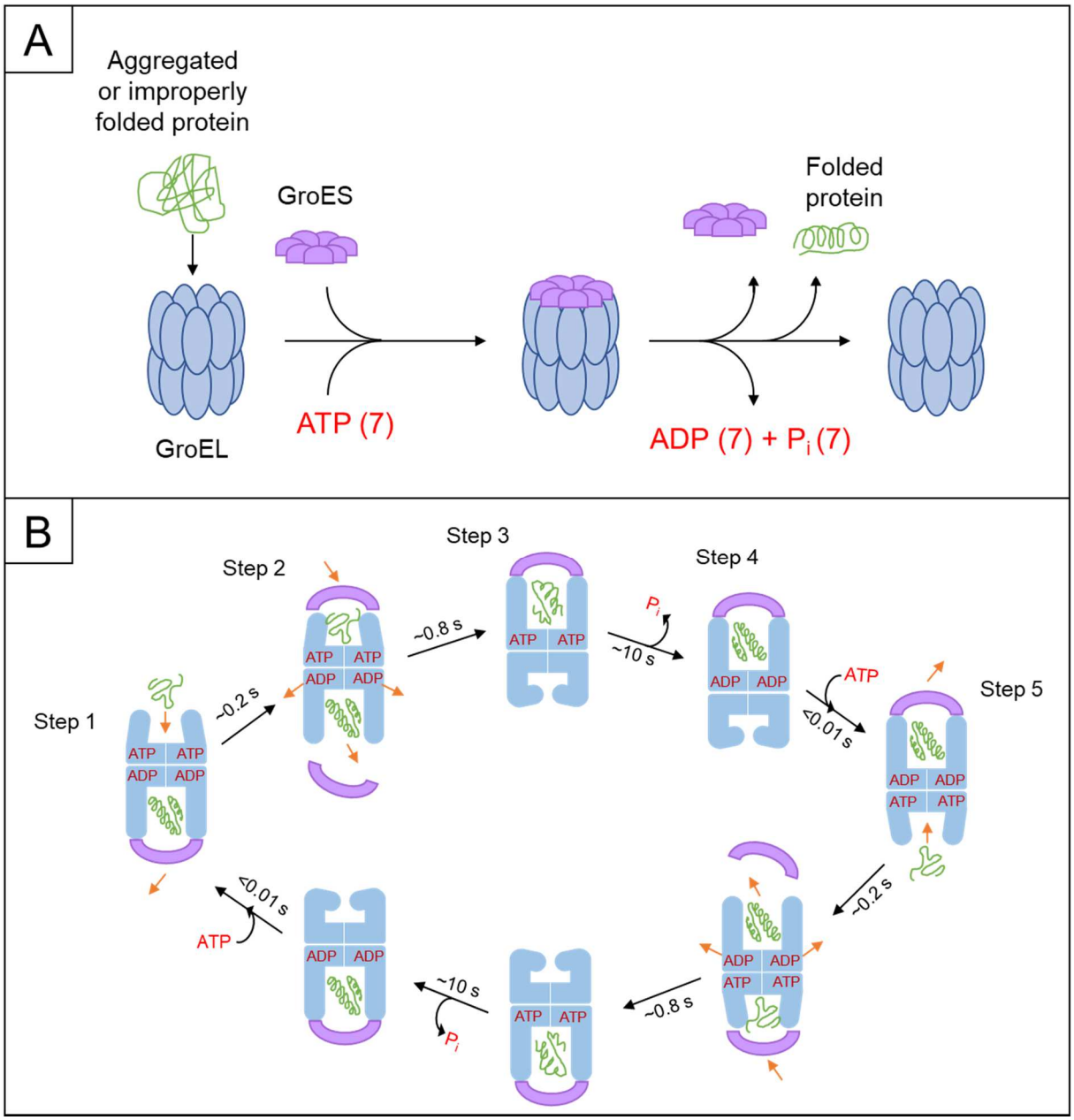


Figure 1.10: Proposed mechanism of action for GroE cycle

Blue rings/barrels depict GroEL quaternary structure, purple ring depicts GroES quaternary structure, and green lines represent a protein in various states of folding. Panel A: Depiction of all components involved in GroE assisted protein folding. Each GroEL subunit can utilize 1 ATP, for a total of 7 ATP (one per GroEL subunit) per cycle. Panel B: Cycle of GroE assisted protein folding starts at position 1 and circles around to position 5 where the cycle restarts with

the bottom GroEL actively participating in protein folding. Approximate times for each step to occur are reported next to arrows; ATP hydrolysis is rate limiting. Adapted from (125).

Briefly, binding of ATP to one GroEL ring, top ring **Figure 1.10 Panel B**, causes a conformational change extending the barrel structure of GroEL (**Step 1**). The GroEL conformational change also releases GroES attached to the bottom GroEL ring (**Step 1**). Exposed hydrophobic residues on a cytosolic protein can then bind to hydrophobic patches on the edge of GroEL allowing entrance into the folding cavity. GroES then binds to GroEL sequestering the protein inside the cavity (**Step 2**). The binding of GroES to the top GroEL induces the bottom GroEL heptamer to release its bound ADP and protein (**Steps 2 and 3**). The GroEL-bound ATP is then hydrolyzed to ADP while protein folding occurs inside the GroE cavity (**Step 4**). It is hypothesized that two mechanisms assist in protein folding inside GroE. First, the protein is sequestered from other components of the cytosol, preventing interruptions and/or interactions that could prevent correct folding in the cytosol (119). Second, hydrophobic patches inside the GroEL cavity may stabilize certain, necessary transition states for the folding protein that are not thermodynamically favored in the open cytosol (119). After ATP hydrolysis, new ATP can bind to the bottom GroEL ring and the process cycles through the second GroEL ring (**Step 5**). Proteins often go through multiple folding cycles with GroE to achieve their final tertiary structure (126).

While this is the current, understanding of the GroE mechanism, some bacteria may undergo variations of this mechanism (127-129). Additionally, there is still debate as to whether or not a protein needs to be fully confined within the GroE complex to be assisted in folding

(114). Because GroE has not been studied in *A. baylyi* ADP1, the exact GroE mechanism occurring in this strain is unclear.

Recently, overexpression of *groEL* and *groES* have been reported to increase toxicity tolerance in a genome-reduced *Pseudomonas putida* strain growing on thermochemical wastewater samples, a potential microbial feedstock containing many toxic substrates (130). Similarly, overexpression of GroE encoding genes in *E. coli* increased cell viability when grown on several relatively toxic metabolites (131). In both studies, researchers hypothesize that increased levels of GroE helped cells overcome various metabolite toxicities by rescuing unfolded, damaged, and/or unstable proteins. Furthermore, Laun et al. demonstrated that heterologous expression of GroE components from solvent-tolerant or thermotolerant bacteria can help other bacteria overcome various environmental stressors (132). Specifically, when overexpressed in *E. coli*, the *Psuedomonas putida* (solvent-tolerant) GroE improved *E. coli's* ethanol-tolerance and thermal-tolerance (132). These studies have contributed to a rich literature confirming GroE as important in various bacterial stress-response mechanisms and whose overexpression increases cellular tolerance to many types of stresses (122, 133). It has been noted that engineering post-translational machineries, such as GroE, is may be necessary to enhance toxicity tolerance in many bacteria (130, 133).

Peptidyl-prolyl isomerases (PPIases)

Another type of post-translational modification system in which many mutations were discovered during my laboratory evolution studies include peptidyl-prolyl isomerases (PPIases). PPIases are proteins that convert proline residues within a peptide chain between the *cis* and *trans* conformers (**Figure 1.11**) (134). Proline isomerization is often necessary for proteins to

achieve their final active conformations, and it is often the rate limiting step during protein maturation (134). Eight ADP1-derived isolates from eleven independently evolved populations had a mutation in a predicted PPIase. These mutations, listed in **Table 1.4**, appear to prevent functioning of at least some of the PPIases based on the type of mutation found. The mutations include two complete deletions of predicted PPIases, a premature truncation, an amino acid replacement, and a frequently occurring small deletion. Furthermore, these mutations do not just occur in one gene but in four, ACIAD_RS02420 (*tig*), ACIAD_RS04740 (*fbp*), ACIAD_RS05440 (*slyD*), and ACIAD_RS06505 (*ppiD*). While *A. baylyi* ADP1 encodes a number of predicted PPIases listed in **Table 1.5**, only one PPIase, ACIAD_RS16500 (*rotA/ppiA*), has been characterized (135). Therefore, it is unclear what role these coding sequences might be playing in my *A. baylyi* strains, but previous studies may help shed some light on to their possible effects in my evolved strains.

Table 1.4: Summary of FKBP-type (FK506 binding protein) peptidyl-prolyl isomerase mutations from evolution studies

Evolved Strain	Gene	Change	Effect
ACN2140	ACIAD_RS02420 (*tig*)	Δ2 bp	2 bp deletion AA328-329 causing frameshift and premature truncation after AA333
ACN2140	ACIAD_RS04740 (*fbp*)	Chromosomal region deletion	N/A
ACN2429	ACIAD_RS05440 (*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	Small deletion
ACN2430	ACIAD_RS05440 (*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	Small deletion
ACN2431	ACIAD_RS05440 (*slyD*)	C→T	E46K; AA replacement
ACN2432	ACIAD_RS05440 (*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	Small deletion
ACN2433	ACIAD_RS05440 (*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	Small deletion
ACN2434	ACIAD_RS05440 (*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	Small deletion

ACN2714	ACIAD_RS06505 (*ppiD*)	Chromosomal region deletion	N/A

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Table 1.5: Peptidyl-prolyl Isomerases in *A. baylyi* ADP1

Locus Tag	Gene	Type (Based on sequence similarity)	NCBI Protein ID
ACIAD_RS02420	*tig*	FKBP (Named for ability to bind human immunosuppressant FK506) (FK506 binding protein)	WP_004920071.1
ACIAD_RS05440	*slyD*		WP_004926184.1
ACIAD_RS04740	*fbp*		WP_004921748.1
ACIAD_RS00315	*fkpA*		WP_004930852.1
ACIAD_RS00110	*fkpB*		WP_004930977.1
ACIAD_RS00310	*fklB*		WP_004930855.1
ACIAD_RS16500	*rotA/ppiA*	Cyclophilin (Named after ability to bind cyclosporin A)	WP_004930172.1
ACIAD_RS08845	*ppiB*		WP_004927153.1
ACIAD_RS03240	*ppiC*	Parvulin (Named after first identified PPIase in *E. coli* not inhibited by FK506 or cyclosporin A)	WP_004922592.1
ACIAD_RS06505	*ppiD*		WP_004925574.1
ACIAD_RS10850	*surA*		WP_004928216.1

* Highlighting indicates loci in which mutations arose during adaptive evolution

Peptide bonds have two forms, *cis* (0°) or *trans* (180°) (**Figure 1.10**). Because of steric hindrance, the *trans* form is greatly favored in nearly all peptide bonds. However, proline is unique among the proteogenic amino acids because its side chain circularizes about the nitrogen involved in peptide bonding, N-alkylation. Because of the stability offered to the peptide bond from the N-alkylation, the *cis* peptide bond is much more stable than for other amino acids and

thus occurs more frequently in peptide chains. Studies of short peptides demonstrate the stability of Xaa-Pro (Xaa is used here as a stand in for any amino acid) *cis* conformers in peptide bonds as they make up around 30% of all Xaa-Pro bonds (136, 137). In protein structures, prolines are typically found located in β-turns and loops, and the specific conformation of the proline bond defines the structure and thus type of turn (138, 139). Therefore, cells need to be able to interconvert between the *cis* and *trans* forms of the proline peptide bonds for proper protein folding and functioning. Because the spontaneous isomerization is very slow, typically on the order of minutes, as it involves rotation around a partial double bond, cells have evolved specific peptidylprolyl isomerases to carry out this reaction (140). During protein maturation, this step is often rate limiting (141).

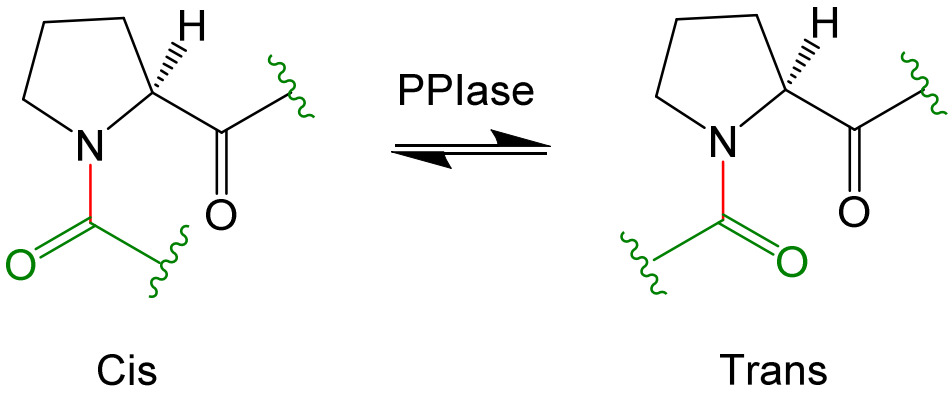


Figure 1.11: *cis-trans* isomers of proline peptide bond

Black stick figure represents proline in a peptide chain. Green stick figure represents the C terminus of the amino acid in the peptide chain before proline. Green squiggle lines represent ongoing peptide chain not depicted. Peptide bond about which the *cis-trans* conformation is oriented is red.

There are currently three major categories of PPIases characterized by sequence homology, FK506-binding proteins (FKBPs), cyclophilins (inhibited by cyclosporin A), and parvulin (not inhibited by FK506 or cyclosporin A) (142). This review will focus on FKBP-type PPIases since most of the mutations discovered during my studies occurred in these types of PPIases. Other reviews summarize cyclophilin-type and parvulin-type PPIases (143, 144). FKBP-type PPIases, InterPro entry IPR001179, are present in all kingdoms of life, but the first characterized was the human FKBP12 that bound FK506, an immunosuppressive macrolide lactone (88, 145). FKBPs often also contain protein chaperone domains that assist in protein folding (146).

In my laboratory evolution experiments, many strains developed a 9 bp deletion in the DNA encoding an insert-in-flap (IF) chaperone domain of the predicted *A. baylyi* SlyD. Some prokaryotic FKBPs, including as SlyD, have an IF chaperone domain instead of the flap loop domain common to most other FKBPs (147). IF domains are predicted to function by binding to refolding proteins to act as a chaperone during PPIase activity (148). This domain is reported to demonstrate 100 fold higher activity compared to flap loop chaperon domains, and in one study, it increased the chaperone activity of FKBP12 200 fold when inserted in place of the native flap loop domain (148, 149). The evolved *A. baylyi slyD* allele has not been characterized, but it is interesting that multiple strains evolved the same deletion in this IF chaperone domain. It therefore appears that removing the IF chaperone domain from SlyD provides some benefit to the evolved *A. baylyi* cells under the selected conditions.

A truncation of the predicted FKBP-type PPIase trigger factor also evolved in my laboratory evolution studies. While the *slyD* mutations were found in many evolved strains, a mutation affecting trigger factor was only present in one. Trigger factor, encoded by the gene *tig*,

is the only currently known PPIase/chaperone in prokaryotes to associate with the ribosome to assist in folding as nascent protein chains are synthesized (150). The *E. coli* trigger factor associates with the ribosome in a 1:1 ratio by binding the L23 ribosomal subunit (151, 152). Trigger factor has 3 domains. The N-terminus (residues 1-112) binds the L23 protein of the ribosome and contributes to the chaperone activity of the protein, the middle domain (residues 150-246) contains the PPIase, and the C-terminal domain (residues 247-432) functions as a molecular chaperone (**Figure 1.12**) (150).

Figure 1.12 shows an alignment of the characterized *E. coli* K12 trigger factor and the predicted *A. baylyi* ADP1 trigger factor protein sequence. The evolved *tig* allele from my studies truncates the protein after amino acid 333, downstream of the PPIase domain (indicated in yellow) and within the chaperone domain in the black box (**Figure 1.12)**. Similar to the evolved *slyD* allele, the evolved *tig* allele appears to be altering (in this case deleting) the chaperone domain of this PPIase. While my studies in **Chapter 2** indicate that the *tig* mutation, in combination with another evolved mutation, is sufficient to confer the evolved phenotype, future studies are needed to elucidate how these alleles are functioning in *A. baylyi*.

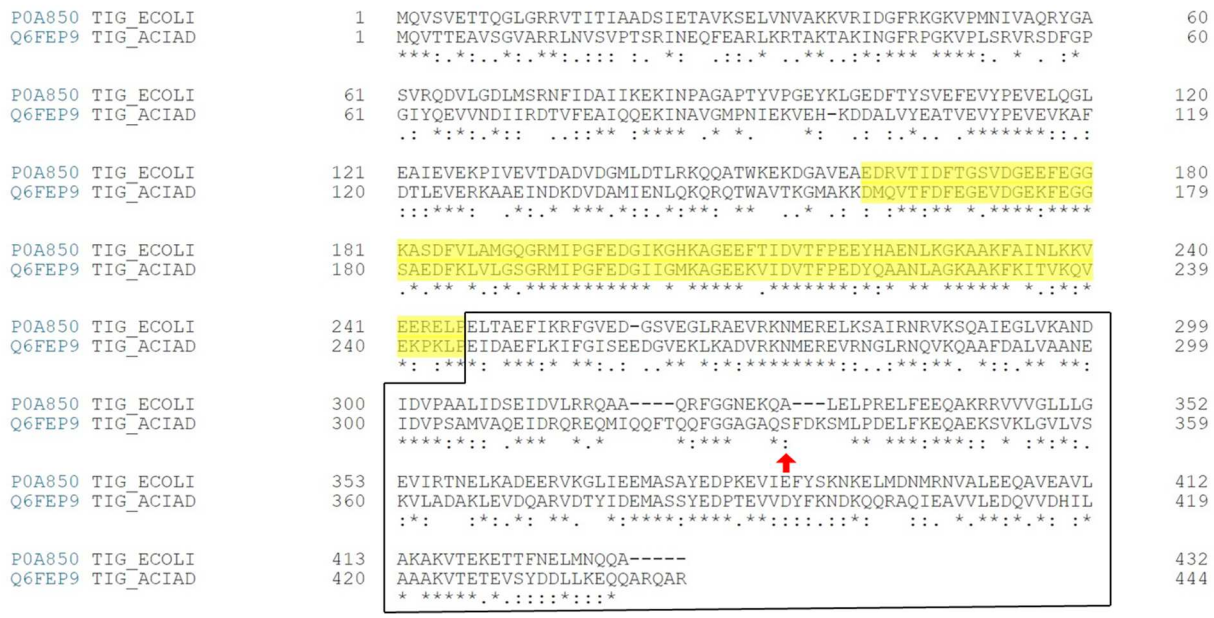


Figure 1.12: Alignment of *E. coli* K12 trigger factor to *A. baylyi* ADP1 predicted trigger factor

Alignment of *A. baylyi* predicted trigger factor (UniProt ID Q6FEP9) and *E. coli* trigger factor (UniProt ID P0A850) (91). Residues highlighted in yellow indicate the PPIase domain. Chaperone domain corresponds to all residues occurring after the PPIase domain within black box. Red arrow shows where *A. baylyi* evolved allele truncates the protein. Markings underneath each residue in the alignment indicate exact match (*), conserved substitution (:), or semiconserved substitution (.) (91).

ClpA: a molecular chaperone and component of the ClpAP protease

During my laboratory evolution studies in **Chapter 2**, one evolved ADP1-derived strain acquired a mutation in the loci ACIAD_RS06285, annotated as *clpA*. Even though this locus only acquired one mutation out of all eleven independently evolved *A. baylyi* isolates, it was noticeable because mutations in this locus appeared multiple times in my studies of aspartate regulation in ADP1, **Chapter 3**. ADP1-derived strains missing the transcriptional regulator AalR

lose the ability to grow on aspartate as a sole carbon source. However, suppressor mutants appear after multiple days of incubation. When sequenced, many of these suppressors had acquired mutations in the locus annotated as *clpA* that are sufficient to restore growth on aspartate in the strains missing AalR.

ClpA is the translocase/unfoldase chaperone component of the ATP-dependent ClpAP protease (153-156). ClpA, often with the help of substrate modulators, recognizes, unfolds, and translocates proteins to the ClpP protease for degradation (157, 158). ClpA is a major player in the regulation of protein turnover in bacterial cells. The mutations discovered in my studies, shown in **Table 1.6**, all appear to cause pre-mature truncations of *clpA*. In the AalR studies, we demonstrated that the effects were caused by loss of a functional ClpA, thereby presumably preventing the targeted degradation of certain proteins in the cell. To improve understanding of this process, an overview of ClpA is provided here.

Table 1.6: *clpA* mutations from multiple adaptive evolution studies in *A. baylyi* ADP1

Strain	Genotype	Growth selection	Mutation in *clpA*	Position of Truncation (natively a 758 AA long sequence)
ACN2435	Δ*pcaUIJFBDCHG*::*lac-pcaK52106* ; ACIAD_RS07700:*ligRKUJIABC52119*; ΩK^R *52119*; *SBF52132*	POB^+	ΔT (nucleotide 1736 of 2277)	Alternative codon at AA 579 with 6 AA alternative tail then stop
ACN2180	Δ*aalR*	L-asp^+	11 nucleotide repeat (nucleotides 1138-1148)	Alternative codon at AA 383 with 622 AA alternative tail then stop
ACN2917	Δ*aalR*	L-asp^+	G -> T (82^{nd} nucleotide of 2277)	AA 28
ACN2926	*sacB*:K^R interrupting *aalR*	L-asp^+ and D-asp^+	GG insertion (duplication of nucleotides 1600 and 1601 of CDS)	Alternative codon at AA 534 with 14 AA alternative tail then stop
ACN2927	*sacB*:K^R interrupting *aalR*	L-asp^+ and D-asp^+	ΔC (nucleotide 899 of 2277)	Alternative codon at AA 300 with 5 AA alternative tail then stop
ACN2928	*sacB*:K^R interrupting *aalR*	L-asp^+ and D-asp^+	T insertion (after 560^{th} nucleotide of 2277)	Alternative codon at AA 191 with 4 AA alternative tail then stop

ClpAP is a member of the AAA+ family of proteases, ATPases Associated with diverse cellular Activities, that typically consist of an AAA+ unfoldase and a self-compartmentalized peptidase (159-161). Each of the six subunits making up the ClpA unfoldase ring contain two phylogenetically distinct AAA+ domains, D1 and D2 (**Figure 1.13**). D1 is from the extended AAA group and D2 is from the protease, chelatase, transcriptional activators, and transport (PACTT) group (159). Because each of the different AAA+ domains function in a large number of cellular processes, clades are not differentiated by function but rather by the presence or absence of distinct structural elements near the AAA+ domain (159). In studies of the *E. coli* ClpA, it has been suggested that D1 acts primarily as the regulatory domain, recognizing proteins to be degraded, and only assists D2 in its role as the major ATP-powered unfoldase and translocase (162-164).

Figure 1.13 depicts the current understanding of ClpAP proteolysis in which ClpA recognizes and unfolds a target protein. It then translocates the protein directly into the ClpP protease channel (161, 165, 166). ClpP is a serine protease, meaning serine acts as the molecular scissors to cleave the peptide chain (166). ClpA may directly recognize a degron on a protein, a tag marking it for degradation, or protein recognition can be mediated by an adaptor such as ClpS (not shown in figure) (158, 167). ClpS, a homolog of which appears to be present in ADP1, binds to the exposed N-terminus of ClpA and targets proteins with an N terminal bulky hydrophobic reside for degradation (leucine, phenylalanine, tyrosine, or tryptophan) (167).

Neither ClpA, nor ClpP, has been characterized in ADP1, but one previous study explored evolved *clpA* mutants in ADP1 (50, 168). In this study, mutant *clpA* alleles were

selected during laboratory evolution as ADP1 was propagated on LB for 1,000 generations and then analyzed for chromosomal changes. Researchers found that strains that developed mutations in *clpA* also had statistically significant higher rates of IS*1236*-mediated events than other strains (50). Since my selected strains lack ClpA, certain proteins normally targeted by ClpA appear to evade degradation and are therefore present longer in the *A. baylyi* cells. Because I have not selected strains with mutated *clpS* alleles, it could be possible that the proteins recognized do not have the ClpS degron, but rather some other degron directly recognized by ClpA.

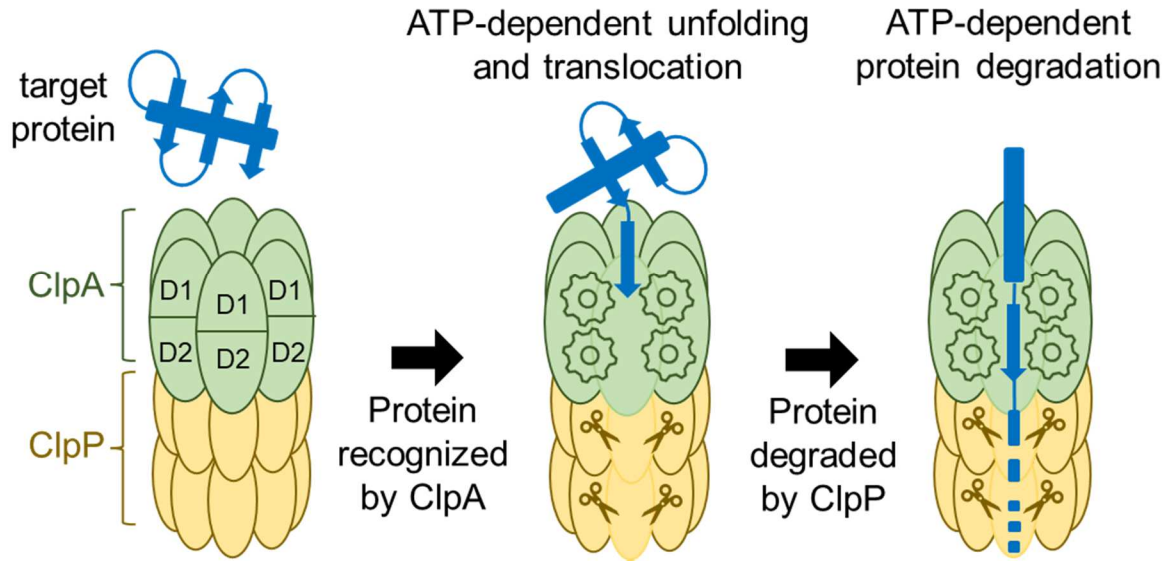


Figure 1.13: ClpAP mechanism of action

Green ring represents ClpA hexamer with D1 and D2 ATP-binding domains which coordinate a ratchet-like movement to translocate and unfold the target protein (gears within ClpA). Yellow rings represent ClpP heptamers which cleave the protein fed into the ClpP ring by ClpA (scissors within ClpP).

Dissertation overview

The remaining chapters of this dissertation will focus on findings from the research I have done during my doctoral studies that are not yet published. Papers on which I am a co-author are cited or discussed as relevant (11, 52, 169). **Chapter 2** will summarize the laboratory evolution of a foreign protocatechuate degradation pathway inserted into the *A. baylyi* chromosome. This project demonstrated that large foreign pathways can be heterologously expressed on the *A. baylyi* chromosome and subsequently evolved by the EASy laboratory evolution method. These studies revealed a possible PTS regulatory system in *A. baylyi* and implicated multiple protein chaperones as important in thermal tolerance.

Chapter 3 summarizes the characterization of aspartate metabolism in *A. baylyi* ADP1. The impacts of two transcriptional regulators on L-aspartate and D-aspartate were investigated as well as multiple aspartate transporters. This work was done in collaboration with many instructors, TAs, and students of the undergraduate course MIBO4600L, of which I was the TA for two semesters.

Chapter 4 will cover my efforts to insert and express a pathway for the degradation of the aromatic compounds veratrate and isovanillate in *A. baylyi* ADP1. **Chapter 5** will summarize key information learned throughout my studies and detail future directions for the research described here. A notable future direction is the continuation of my work in developing genetic tools for use in *A. baylyi* ADP1. One tool seeks to discover new genetic pathways from environmental DNA samples by taking advantage of the unique genetic system of ADP1 to take up and incorporate large segments of environmental DNA directly into the ADP1 chromosome. Another method being developed involves the expression a restriction enzyme in ADP1 to digest DNA transformed into the cells for integration into multiple separate regions of the ADP1

chromosome. It is hoped that collectively this work will increase the success of using ADP1 as model bacterium for understanding fundamental biology and for achieving biotechnology goals, such as lignin valorization.

Abbreviations

Amp	ampicillin
Ben	benzoate
Cat	catechol
EASy	evolution by amplification and synthetic biology
HPLC	high performance liquid chromatography
Ivan	isovanillate
Kan	kanamycin
LB	lysogeny broth
MM	minimal media
OD	optical density
PCA	protocatechuate
PCA-2,3	protocatechuate 2,3-cleavage pathway
PCA-3,4	protocatechuate3,4-cleavage pathway
PCA-4,5	protocatechuate 4,5-cleavage pathway
PCR	polymerase chain reaction
POB	para-hydroxybenzoate
PPIase	peptidyl-prolyl isomerase
PTS	phosphotransferase system

qPCR	quantitative polymerase chain reaction
SBF	synthetic bridging fragment
St/Sr	spectinomycin and streptomycin
Ver	veratrate
WGS	whole genome sequencing
WT	wild type

References

1. Dagley S. 1978. Determinants of biodegradability. Quarterly Reviews of Biophysics 11:577-602.
2. Fuchs G, Boll M, Heider J. 2011. Microbial degradation of aromatic compounds—from one strategy to four. Nature Reviews Microbiology 9:803.
3. Salvachúa D, Karp EM, Nimlos CT, Vardon DR, Beckham GT. 2015. Towards lignin consolidated bioprocessing: simultaneous lignin depolymerization and product generation by bacteria. Green Chemistry 17:4951-4967.
4. Salmela M, Lehtinen T, Efimova E, Santala S, Santala V. 2019. Alkane and wax ester production from lignin-related aromatic compounds. Biotechnology and Bioengineering 116:1934-1945.
5. Johnson CW, Beckham GT. 2015. Aromatic catabolic pathway selection for optimal production of pyruvate and lactate from lignin. Metabolic Engineering 28:240-247.
6. Kamimura N, Sakamoto S, Mitsuda N, Masai E, Kajita S. 2019. Advances in microbial lignin degradation and its applications. Current Opinion in Biotechnology 56:179-186.

7. Chen Y, Banerjee D, Mukhopadhyay A, Petzold CJ. 2020. Systems and synthetic biology tools for advanced bioproduction hosts. Current Opinion in Biotechnology 64:101-109.

8. Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT. 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proceedings of the National Academy of Sciences 111:12013-12018.

9. Beckham GT, Johnson CW, Karp EM, Salvachúa D, Vardon DR. 2016. Opportunities and challenges in biological lignin valorization. Current Opinion in Biotechnology 42:40-53.

10. Elliott KT, Neidle EL. 2011. Acinetobacter baylyi ADP1: transforming the choice of model organism. IUBMB life 63:1075-1080.

11. Biggs BW, Bedore SR, Arvay E, Huang S, Subramanian H, McIntyre EA, Duscent-Maitland CV, Neidle EL, Tyo KEJ. 2020. Development of a genetic toolset for the highly engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. Nucleic Acids Research doi:10.1093/nar/gkaa167.

12. Suarez GA, Dugan KR, Renda BA, Leonard SP, Gangavarapu LS, Barrick JE. 2020. Rapid and assured genetic engineering methods applied to *Acinetobacter baylyi* ADP1 genome streamlining. Nucleic Acids Research doi:10.1093/nar/gkaa204.

13. Schlesinger W. 1977. Carbon balance in terrestrial detritus, Annu. Re.

14. Boerjan W, Ralph J, Baucher M. 2003. Lignin biosynthesis. Annual Review of Plant Biology 54:519-546.

15. Chundawat SP, Beckham GT, Himmel ME, Dale BE. 2011. Deconstruction of lignocellulosic biomass to fuels and chemicals. Annual Review of Chemical and Biomolecular Engineering 2:121-145.

16. Kaur R, Kaur P. 2021. Chemical valorization of cellulose from lignocellulosic biomass: a step towards sustainable development. Cellulose Chemistry and Technology 55:207-222.

17. Sannigrahi P, Pu Y, Ragauskas A. 2010. Cellulosic biorefineries—unleashing lignin opportunities. Current Opinion in Environmental Sustainability 2:383-393.

18. Chio C, Sain M, Qin W. 2019. Lignin utilization: a review of lignin depolymerization from various aspects. Renewable and Sustainable Energy Reviews 107:232-249.

19. Prieur B, Meub M, Wittemann M, Klein R, Bellayer S, Fontaine G, Bourbigot S. 2017. Phosphorylation of lignin: characterization and investigation of the thermal decomposition. RSC Advances 7:16866-16877.

20. Kamimura N, Takahashi K, Mori K, Araki T, Fujita M, Higuchi Y, Masai E. 2017. Bacterial catabolism of lignin□derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. Environmental Microbiology Reports 9:679-705.

21. Bugg TD, Rahmanpour R. 2015. Enzymatic conversion of lignin into renewable chemicals. Current Opinion in Chemical Biology 29:10-17.

22. Seaton SC, Neidle EL. 2018. Using aerobic pathways for aromatic compound degradation to engineer lignin metabolism. Lignin Valorization: Emerging Approaches 19:252.

23. Vardon DR, Franden MA, Johnson CW, Karp EM, Guarnieri MT, Linger JG, Salm MJ, Strathmann TJ, Beckham GT. 2015. Adipic acid production from lignin. Energy & Environmental Science 8:617-628.

24. Johnson CW, Salvachúa D, Khanna P, Smith H, Peterson DJ, Beckham GT. 2016. Enhancing muconic acid production from glucose and lignin-derived aromatic compounds via increased protocatechuate decarboxylase activity. Metabolic Engineering Communications 3:111-119.

25. Salvachúa D, Johnson CW, Singer CA, Rohrer H, Peterson DJ, Black BA, Knapp A, Beckham GT. 2018. Bioprocess development for muconic acid production from aromatic compounds and lignin. Green Chemistry 20:5007-5019.

26. Bentley GJ, Narayanan N, Jha RK, Salvachúa D, Elmore JR, Peabody GL, Black BA, Ramirez K, De Capite A, Michener WE. 2020. Engineering glucose metabolism for enhanced muconic acid production in *Pseudomonas putida* KT2440. Metabolic Engineering.

27. Johnson CW, Salvachúa D, Rorrer NA, Black BA, Vardon DR, John PCS, Cleveland NS, Dominick G, Elmore JR, Grundl N. 2019. Innovative chemicals and materials from bacterial aromatic catabolic pathways. Joule 3:1523-1537.

28. Alvarez H, Steinbüchel A. 2002. Triacylglycerols in prokaryotic microorganisms. Applied Microbiology and Biotechnology 60:367-376.

29. Kalscheuer R, Steinbüchel A. 2003. A novel bifunctional wax ester synthase/acyl-CoA: diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in *Acinetobacter calcoaceticus* ADP1. Journal of Biological Chemistry 278:8075-8082.

30. Santala S, Efimova E, Santala V. 2018. Dynamic decoupling of biomass and wax ester biosynthesis in *Acinetobacter baylyi* by an autonomously regulated switch. Metabolic Engineering Communications 7:e00078.

31. Santala S, Efimova E, Koskinen P, Karp MT, Santala V. 2014. Rewiring the wax ester production pathway of *Acinetobacter baylyi* ADP1. ACS Synthetic Biology 3:145-151.

32. Santala S, Santala V, Liu N, Stephanopoulos G. 2021. Partitioning metabolism between growth and product synthesis for coordinated production of wax esters in *Acinetobacter baylyi* ADP1. Biotechnology and Bioengineering 118:2283-2292.

33. Arvay E, Biggs BW, Guerrero L, Jiang V, Tyo K. 2021. Engineering A*cinetobacter baylyi* ADP1 for mevalonate production from lignin monomers. Metabolic Engineering Communications:e00173.

34. Baumann P. 1968. Isolation of Acinetobacter from soil and water. Journal of Bacteriology 96:39-42.

35. Baumann P, Doudoroff M, Stanier R. 1968. A study of the Moraxella group II. Oxidative-negative species (genus Acinetobacter). Journal of Bacteriology 95:1520-1541.

36. Juni E, Janik A. 1969. Transformation of *Acinetobacter calco-aceticus* (Bacterium anitratum). Journal of Bacteriology 98:281-288.

37. Vaneechoutte M, Dijkshoorn L, Tjernberg I, Elaichouni A, de Vos P, Claeys G, Verschraegen G. 1995. Identification of Acinetobacter genomic species by amplified ribosomal DNA restriction analysis. Journal of Clinical Microbiology 33:11-15.

38. Barbe V, Vallenet D, Fonknechten N, Kreimeyer A, Oztas S, Labarre L, Cruveiller S, Robert C, Duprat S, Wincker P. 2004. Unique features revealed by the genome sequence of *Acinetobacter sp.* ADP1, a versatile and naturally transformation competent bacterium. Nucleic Acids Research 32:5766-5779.

39. Palmen R, Vosman B, Buijsman P, Breek CK, Hellingwerf KJ. 1993. Physiological characterization of natural transformation in *Acinetobacter calcoaceticus*. Microbiology 139:295-305.

40. Lorenz MG, Wackernagel W. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiological Reviews 58:563-602.

41. Palmen R, Buijsman P, Hellingwerf KJ. 1994. Physiological regulation of competence induction for natural transformation in *Acinetobacter calcoaceticus*. Archives of Microbiology 162:344-351.

42. de Vries J, Meier P, Wackernagel W. 2001. The natural transformation of the soil bacteria *Pseudomonas stutzeri* and *Acinetobacter sp.* by transgenic plant DNA strictly depends on homologous sequences in the recipient cells. FEMS Microbiology Letters 195:211-215.

43. Seaton SC, Elliott KT, Cuff LE, Laniohan NS, Patel PR, Neidle EL. 2012. Genome‐wide selection for increased copy number in *Acinetobacter baylyi* ADP1: locus and context‐dependent variation in gene amplification. Molecular Microbiology 83:520-535.

44. Jiang X, Palazzotto E, Wybraniec E, Munro LJ, Zhang H, Kell DB, Weber T, Lee SY. 2020. Automating cloning by natural transformation. ACS Synthetic Biology 9:3228-3235.

45. Reams AB, Neidle EL. 2004. Gene amplification involves site-specific short homology-independent illegitimate recombination in *Acinetobacter sp*. strain ADP1. Journal of molecular biology 338:643-656.

46. De Berardinis V, Vallenet D, Castelli V, Besnard M, Pinet A, Cruaud C, Samair S, Lechaplais C, Gyapay G, Richez C. 2008. A complete collection of single□gene deletion mutants of *Acinetobacter baylyi* ADP1. Molecular Systems Biology 4:174.

47. Gerischer U. 2008. Acinetobacter Molecular Microbiology. Caister Academic Press.

48. Reams AB, Neidle EL. 2004. Selection for gene clustering by tandem duplication. Annual Reviews Microbiology 58:119-142.

49. Reams AB, Neidle EL. 2003. Genome plasticity in Acinetobacter: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. Molecular Microbiology 47:1291-1304.

50. Renda BA, Dasgupta A, Leon D, Barrick JE. 2015. Genome instability mediates the loss of key traits by *Acinetobacter baylyi* ADP1 during laboratory evolution. Journal of Bacteriology 197:872-81.

51. Suarez GA, Renda BA, Dasgupta A, Barrick JE. 2017. Reduced Mutation Rate and Increased Transformability of Transposon-Free *Acinetobacter baylyi* ADP1-ISx. Applied Environmental Microbiology83.

52. Jin Luo EM, Stacy Bedore, Ville Santala, Ellen Neidle, and Suvi Santala. 2021. Characterization of highly ferulate-tolerant *Acinetobacter baylyi* ADP1 isolates by a rapid reverse-engineering method. Applied and Environmental Microbiology.

53. Craven SH, Ezezika OC, Momany C, Neidle EL. 2008. LysR homologs in Acinetobacter: insights into a diverse and prevalent family of transcriptional regulators. Norfolk: Caister Academic Press.

54. Pareja E, Pareja-Tobes P, Manrique M, Pareja-Tobes E, Bonal J, Tobes R. 2006. ExtraTrain: a database of Extragenic regions and Transcriptional information in prokaryotic organisms. BMC Microbiology 6:1-10.

55. Alanazi AM, Neidle EL, Momany C. 2013. The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N11-A sequence motif by LysR-type transcriptional regulators. Acta Crystallographica Section D: Biological Crystallography 69:1995-2007.

56. Collier LS, Gaines GL, Neidle EL. 1998. Regulation of benzoate degradation in *Acinetobacter sp.* strain ADP1 by BenM, a LysR-type transcriptional activator. Journal of Bacteriology 180:2493-2501.

57. Clark T, Haddad S, Neidle E, Momany C. 2004. Crystallization of the effector-binding domains of BenM and CatM, LysR-type transcriptional regulators from *Acinetobacter sp.* ADP1. Acta Crystallographica Section D: Biological Crystallography 60:105-108.

58. Ezezika OC, Haddad S, Clark TJ, Neidle EL, Momany C. 2007. Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. Journal of Molecular Biology 367:616-629.

59. Bundy BM, Collier LS, Hoover TR, Neidle EL. 2002. Synergistic transcriptional activation by one regulatory protein in response to two metabolites. Proceedings of the National Academy of Sciences 99:7693-7698.

60. Neidle E, Hartnett C, Ornston L. 1989. Characterization of *Acinetobacter calcoaceticus* catM, a repressor gene homologous in sequence to transcriptional activator genes. Journal of Bacteriology 171:5410-5421.

61. Ezezika OC, Collier-Hyams LS, Dale HA, Burk AC, Neidle EL. 2006. CatM regulation of the benABCDE operon: functional divergence of two LysR-type paralogs in *Acinetobacter baylyi* ADP1. Applied and Environmental Microbiology 72:1749-1758.

62. Romero-Arroyo CE, Schell MA, Gaines G, Neidle EL. 1995. catM encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. Journal of Bacteriology 177:5891-5898.

63. Brzostowicz PC, Reams AB, Clark TJ, Neidle EL. 2003. Transcriptional cross-regulation of the catechol and protocatechuate branches of the β-ketoadipate pathway contributes to carbon source-dependent expression of the *Acinetobacter sp.* strain ADP1 pobA gene. Applied and Environmental Microbiology 69:1598-1606.

64. Ezezika OC, Collier-Hyams LS, Dale HA, Burk AC, Neidle EL. 2006. CatM regulation of the benABCDE operon: functional divergence of two LysR-type paralogs in *Acinetobacter baylyi* ADP1. Applied Environmental Microbiology72:1749-1758.

65. Stoudenmire JL, Schmidt AL, Tumen-Velasquez MP, Elliott KT, Laniohan NS, Whitley SW, Galloway NR, Nune M, West M, Momany C. 2017. Malonate degradation in *Acinetobacter baylyi* ADP1: Operon organization and regulation by MdcR. Microbiology 163:789.

66. Segura A, Bünz PV, D'Argenio DA, Ornston LN. 1999. Genetic Analysis of a Chromosomal Region Containing *vanA* and *vanB*, Genes Required for Conversion of Either Ferulate or Vanillate to Protocatechuate inAcinetobacter. Journal of Bacteriology 181:3494-3504.

67. Jones RM, Pagmantidis V, Williams PA. 2000. sal genes determining the catabolism of salicylate esters are part of a supraoperonic cluster of catabolic genes in *Acinetobacter sp.* strain ADP1. Journal of Bacteriology 182:2018-2025.

68. Bundy BM, Campbell AL, Neidle EL. 1998. Similarities between the antABC-encoded anthranilate dioxygenase and the benABC-encoded benzoate dioxygenase of *Acinetobacter sp.* strain ADP1. Journal of Bacteriology 180:4466-4474.

69. Harwood CS, Parales RE. 1996. The β-ketoadipate pathway and the biology of self-identity. Annual Review of Microbiology 50:553-590.

70. Craven SH, Ezezika OC, Haddad S, Hall RA, Momany C, Neidle EL. 2009. Inducer responses of BenM, a LysR□type transcriptional regulator from *Acinetobacter baylyi* ADP1. Molecular Microbiology 72:881-894.

71. Gaines Gr, Smith L, Neidle EL. 1996. Novel nuclear magnetic resonance spectroscopy methods demonstrate preferential carbon source utilization by *Acinetobacter calcoaceticus*. Journal of Bacteriology 178:6833-6841.

72. Kamimura N, Masai E. 2014. The protocatechuate 4, 5-cleavage pathway: overview and new findings. Biodegradative Bacteria:207-226.

73. Kasai D, Fujinami T, Abe T, Mase K, Katayama Y, Fukuda M, Masai E. 2009. Uncovering the protocatechuate 2, 3-cleavage pathway genes. Journal of Bacteriology 191:6758-6768.

74. Sala□trepat JM, Evans WC. 1971. The meta Cleavage of Catechol by Azotobacter Species: 4□Oxalocrotonate Pathway. European Journal of Biochemistry 20:400-413.

75. Mohan K, Phale PS. 2017. Carbon source-dependent inducible metabolism of veratryl alcohol and ferulic acid in *Pseudomonas putida* CSV86. Applied and Environmental Microbiology 83:e03326-16.

76. Providenti MA, O'Brien JM, Ruff Jr, Cook AM, Lambert IB. 2006. Metabolism of isovanillate, vanillate, and veratrate by *Comamonas testosteroni* strain BR6020. Journal of Bacteriology 188:3862-3869.

77. Morawski B, Segura A, Ornston LN. 2000. Repression of Acinetobacter vanillate demethylase synthesis by VanR, a member of the GntR family of transcriptional regulators. FEMS Microbiology Letters 187:65-68.

78. Wenzel SC, Müller R. 2005. Recent developments towards the heterologous expression of complex bacterial natural product biosynthetic pathways. Current Opinion in Biotechnology 16:594-606.

79. Gustafsson C, Govindarajan S, Minshull J. 2004. Codon bias and heterologous protein expression. Trends in Biotechnology 22:346-353.

80. Tumen-Velasquez M, Johnson CW, Ahmed A, Dominick G, Fulk EM, Khanna P, Lee SA, Schmidt AL, Linger JG, Eiteman MA. 2018. Accelerating pathway evolution by increasing the gene dosage of chromosomal segments. Proceedings of the National Academy of Sciences 115:7105-7110.

81. Copley SD. 2020. Evolution of new enzymes by gene duplication and divergence. The FEBS Journal 287:1262-1283.

82. Taylor JS, Raes J. 2004. Duplication and divergence: the evolution of new genes and old ideas. Annual Review Genetics 38:615-643.

83. Pardo I, Jha RK, Bermel RE, Bratti F, Gaddis M, McIntyre E, Michener W, Neidle EL, Dale T, Beckham GT. 2020. Gene amplification, laboratory evolution, and biosensor screening reveal MucK as a terephthalic acid transporter in *Acinetobacter baylyi* ADP1. Metabolic Engineering 62:260-274.

84. Dardas A, Gal D, Barrelle M, Sauret-Ignazi G, Sterjiades R, Pelmont J. 1985. The demethylation of guaiacol by a new bacterial cytochrome P-450. Archives of Biochemistry and Biophysics 236:585-592.

85. Mallinson SJ, Machovina MM, Silveira RL, Garcia-Borràs M, Gallup N, Johnson CW, Allen MD, Skaf MS, Crowley MF, Neidle EL. 2018. A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. Nature Communications 9:1-12.

86. Görke B, Stülke J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nature Reviews Microbiology 6:613-624.

87. Martinez SE, Heikaus CC, Beavo JA. 2010. Cyclic Nucleotide-Binding GAF Domains in Phosphodiesterases and Adenylyl Cyclases, p 1531-1536, Handbook of Cell Signaling. Elsevier.

88. Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S. 2021. The InterPro protein families and domains database: 20 years on. Nucleic Acids Research 49:D344-D354.

89. Aravind L, Ponting CP. 1997. The GAF domain: an evolutionary link between diverse phototransducing proteins. Trends in Biochemical Sciences 22:458-459.

90. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA, Kustu S. 2000. Nitrogen regulatory protein C-controlled genes of Escherichia coli:

scavenging as a defense against nitrogen limitation. Proceedings of the National Academy of Sciences 97:14674-14679.

91. Pundir S, Martin MJ, O'Donovan C, Consortium U. 2016. UniProt tools. Current Protocols in Bioinformatics 53:1.29. 1-1.29. 15.

92. Durot M, Le Fèvre F, de Berardinis V, Kreimeyer A, Vallenet D, Combe C, Smidtas S, Salanoubat M, Weissenbach J, Schachter V. 2008. Iterative reconstruction of a global metabolic model of *Acinetobacter baylyi* ADP1 using high-throughput growth phenotype and gene essentiality data. BMC Systems Biology 2:85.

93. Deutscher J, Aké FMD, Derkaoui M, Zébré AC, Cao TN, Bouraoui H, Kentache T, Mokhtari A, Milohanic E, Joyet P. 2014. The bacterial phosphoenolpyruvate: carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. Microbiology and Molecular Biology Reviews 78:231-256.

94. Deutscher J, Francke C, Postma PW. 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiology and Molecular Biology Reviews 70:939-1031.

95. Saier Jr M, Hvorup R, Barabote R. 2005. Evolution of the bacterial phosphotransferase system: from carriers and enzymes to group translocators. Portland Press Ltd.

96. Pflüger-Grau K, Görke B. 2010. Regulatory roles of the bacterial nitrogen-related phosphotransferase system. Trends in Microbiology 18:205-214.

97. Powell BS, Inada T, Nakamura Y, Michotey V, Cui X, Reizer A, Saier MH, Reizer J. 1995. Novel Proteins of the Phosphotransferase System Encoded within the rpoN Operon of Escherichia coli: Enzyme IIANtr affects growth on organic nitrogen and the

conditional lethality of an erats mutant (∗). Journal of Biological Chemistry 270:4822-4839.

98. Del Castillo T, Ramos JL. 2007. Simultaneous catabolite repression between glucose and toluene metabolism in *Pseudomonas putida* is channeled through different signaling pathways. Journal of Bacteriology 189:6602-6610.

99. Choi J, Shin D, Yoon H, Kim J, Lee C-R, Kim M, Seok Y-J, Ryu S. 2010. Salmonella pathogenicity island 2 expression negatively controlled by EIIANtr–SsrB interaction is required for Salmonella virulence. Proceedings of the National Academy of Sciences 107:20506-20511.

100. Dozot M, Poncet S, Nicolas C, Copin R, Bouraoui H, Mazé A, Deutscher J, De Bolle X, Letesson J-J. 2010. Functional characterization of the incomplete phosphotransferase system (PTS) of the intracellular pathogen Brucella melitensis. PLoS One 5.

101. Higa F, Edelstein PH. 2001. Potential virulence role of the Legionella pneumophila ptsP ortholog. Infection and Immunity 69:4782-9.

102. Cabeen MT, Leiman SA, Losick R. 2016. Colony□morphology screening uncovers a role for the *Pseudomonas aeruginosa* nitrogen□related phosphotransferase system in biofilm formation. Molecular Microbiology 99:557-570.

103. Lee C-R, Cho S-H, Yoon M-J, Peterkofsky A, Seok Y-J. 2007. *Escherichia coli* enzyme IIANtr regulates the K+ transporter TrkA. Proceedings of the National Academy of Sciences 104:4124-4129.

104. Lee CR, Cho SH, Kim HJ, Kim M, Peterkofsky A, Seok YJ. 2010. Potassium mediates *Escherichia coli* enzyme IIANtr□dependent regulation of sigma factor selectivity. Molecular Microbiology 78:1468-1483.

105. Sánchez-Cañizares C, Prell J, Pini F, Rutten P, Kraxner K, Wynands B, Karunakaran R, Poole PS. 2020. Global control of bacterial nitrogen and carbon metabolism by a PTSNtr-regulated switch. Proceedings of the National Academy of Sciences 117:10234-10245.

106. Lee CR, Park YH, Kim M, Kim YR, Park S, Peterkofsky A, Seok YJ. 2013. Reciprocal regulation of the autophosphorylation of enzyme INtr by glutamine and α□ketoglutarate in *Escherichia coli*. Molecular Microbiology 88:473-485.

107. Untiet V, Karunakaran R, Krämer M, Poole P, Priefer U, Prell J. 2013. ABC transport is inactivated by the PTSNtr under potassium limitation in *Rhizobium leguminosarum* 3841. PLoS One 8:e64682.

108. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM, Krummenacker M, Midford PE, Ong Q. 2019. The BioCyc collection of microbial genomes and metabolic pathways. Briefings in Bioinformatics 20:1085-1093.

109. Speiser Y, Zusman T, Pasechnek A, Segal G. 2017. The *Legionella pneumophila* incomplete phosphotransferase system is required for optimal intracellular growth and maximal expression of PmrA-regulated effectors. Infection and Immunity 85:e00121-17.

110. Anfinsen CB. 1973. Principles that govern the folding of protein chains. Science 181:223-230.

111. Anfinsen C, Scheraga H. 1975. Experimental and theoretical aspects of protein folding. Advances in Protein Chemistry 29:205-300.

112. Thirumalai D, O'Brien EP, Morrison G, Hyeon C. 2010. Theoretical perspectives on protein folding. Annual Review of Biophysics 39:159-183.

113. Goldberger RF, Epstein CJ, Anfinsen CB. 1963. Acceleration of reactivation of reduced bovine pancreatic ribonuclease by a microsomal system from rat liver. Journal of Biological Chemistry 238:628-635.

114. Lin Z, Rye HS. 2006. GroEL-mediated protein folding: making the impossible, possible. Critical Reviews in Biochemistry and Molecular Biology 41:211-239.

115. Ruddon RW, Bedows E. 1997. Assisted protein folding. Journal of Biological Chemistry 272:3125-3128.

116. Ishii N. 2017. GroEL and the GroEL-GroES complex. Macromolecular Protein Complexes:483-504.

117. Saibil HR, Fenton WA, Clare DK, Horwich AL. 2013. Structure and allostery of the chaperonin GroEL. Journal of Molecular Biology 425:1476-1487.

118. Madan D, Lin Z, Rye HS. 2008. Triggering protein folding within the GroEL-GroES complex. Journal of Biological Chemistry 283:32003-32013.

119. Brinker A, Pfeifer G, Kerner MJ, Naylor DJ, Hartl FU, Hayer-Hartl M. 2001. Dual function of protein confinement in chaperonin-assisted protein folding. Cell 107:223-233.

120. Mayhew M, da Silva AC, Martin J, Erdjument-Bromage H, Tempst P, Hartl FU. 1996. Protein folding in the central cavity of the GroEL–GroES chaperonin complex. Nature 379:420-426.

121. Weissman JS, Hohl CM, Kovalenko O, Kashi Y, Chen S, Braig K, Saibil HR, Fenton WA, Norwich AL. 1995. Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. Cell 83:577-587.

122. Fayet O, Ziegelhoffer T, Georgopoulos C. 1989. The groES and groEL heat shock gene products of Escherichia coli are essential for bacterial growth at all temperatures. Journal of Bacteriology 171:1379-1385.

123. Yébenes H, Mesa P, Muñoz IG, Montoya G, Valpuesta JM. 2011. Chaperonins: two rings for folding. Trends in Biochemical Sciences 36:424-432.

124. Sakikawa C, Taguchi H, Makino Y, Yoshida M. 1999. On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. Journal of Biological Chemistry 274:21251-21256.

125. Horwich AL. 2011. Protein folding in the cell: an inside story. Nature medicine 17:1211-1216.

126. Nieba-Axmann SE, Ottiger M, WuÈthrich K, PluÈckthun A. 1997. Multiple cycles of global unfolding of GroEL-bound cyclophilin A evidenced by NMR. Journal of Molecular Biology 271:803-818.

127. Ishii N, Taguchi H, Sumi M, Yoshida M. 1992. Structure of holo-chaperonin studied with electron microscopy Oligomeric cpn10 on top of two layers of cpn60 rings with two stripes each. FEBS Letters 299:169-174.

128. Sumi M, Taguchi H, Yokoyama K, Ishii N, Yoshida M. 1992. Identification and characterization of a chaperonin from Paracoccus denitrificans. Life Sciences Advanced Technologies 11:179-182.

129. Todd MJ, Walke S, Lorimer G, Truscott K, Scopes RK. 1995. The Single-Ring *Thermoanaerobacter brockii* Chaperonin 60 (Tbr-EL7) Dimerizes to Tbr-EL14. cntdot. Tbr-ES7 under Protein Folding Conditions. Biochemistry 34:14932-14941.

130. Jayakody LN, Johnson CW, Whitham JM, Giannone RJ, Black BA, Cleveland NS, Klingeman DM, Michener WE, Olstad JL, Vardon DR. 2018. Thermochemical wastewater valorization via enhanced microbial toxicity tolerance. Energy & Environmental Science 11:1625-1638.

131. Zingaro KA, Papoutsakis ET. 2013. GroESL overexpression imparts *Escherichia coli* tolerance to i-, n-, and 2-butanol, 1, 2, 4-butanetriol and ethanol with complex and unpredictable patterns. Metabolic Engineering 15:196-205.

132. Luan G, Dong H, Zhang T, Lin Z, Zhang Y, Li Y, Cai Z. 2014. Engineering cellular robustness of microbes by introducing the GroESL chaperonins from extremophilic bacteria. Journal of Biotechnology 178:38-40.

133. Borchert AJ, Henson WR, Beckham GT. 2022. Challenges and opportunities in biological funneling of heterogeneous and toxic substrates beyond lignin. Current Opinion in Biotechnology 73:1-13.

134. Schmid FX. 1993. Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. Annual Review of Biophysics and Biomolecular Structure 22:123-143.

135. Kok RG, Christoffels VM, Vosman B, Hellingwerf KJ. 1994. A gene of *Acinetobacter calcoaceticus* BD413 encodes a periplasmic peptidyl-prolyl cis-trans isomerase of the cyclophilin sub-class that is not essential for growth. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression 1219:601-606.

136. Cheng H, Bovey F. 1977. Cis□trans equilibrium and kinetic studies of acetyl□L□proline and glycyl□L□proline. Biopolymers: Original Research on Biomolecules 16:1465-1472.

137. Grathwohl C, Wüthrich K. 1981. NMR studies of the rates of proline cis–trans isomerization in oligopeptides. Biopolymers: Original Research on Biomolecules 20:2623-2633.

138. Stewart DE, Sarkar A, Wampler JE. 1990. Occurrence and role ofcis peptide bonds in protein structures. Journal of Molecular Biology 214:253-260.

139. MacArthur MW, Thornton JM. 1991. Influence of proline residues on protein conformation. Journal of Molecular Biology 218:397-412.

140. Schmidpeter PA, Schmid FX. 2015. Prolyl isomerization and its catalysis in protein folding and protein function. Journal of Molecular Biology 427:1609-1631.

141. Brandts JF, Halvorson HR, Brennan M. 1975. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. Biochemistry 14:4953-4963.

142. Göthel SF, Marahiel M. 1999. Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. Cellular and Molecular Life Sciences CMLS 55:423-436.

143. Wang P, Heitman J. 2005. The cyclophilins. Genome Biology 6:1-6.

144. Barik S. 2006. Immunophilins: for the love of proteins. Cellular and Molecular Life Sciences CMLS 63:2889-2900.

145. Harding MW, Galat A, Uehling DE, Schreiber SL. 1989. A receptor for the immuno-suppressant FK506 is a cis–trans peptidyl-prolyl isomerase. Nature 341:758-760.

146. Quistgaard EM, Weininger U, Ural-Blimke Y, Modig K, Nordlund P, Akke M, Löw C. 2016. Molecular insights into substrate recognition and catalytic mechanism of the chaperone and FKBP peptidyl-prolyl isomerase SlyD. BMC Biology 14:1-25.

147. Kovermann M, Schmid FX, Balbach J. 2013. Molecular function of the prolyl cis/trans isomerase and metallochaperone SlyD. Biological Chemistry 394:965-975.

148. Knappe TA, Eckert B, Schaarschmidt P, Scholz C, Schmid FX. 2007. Insertion of a chaperone domain converts FKBP12 into a powerful catalyst of protein folding. Journal of Molecular Biology 368:1458-1468.

149. Jakob RP, Zoldák G, Aumüller T, Schmid FX. 2009. Chaperone domains convert prolyl isomerases into generic catalysts of protein folding. Proceedings of the National Academy of Sciences 106:20282-20287.

150. Ferbitz L, Maier T, Patzelt H, Bukau B, Deuerling E, Ban N. 2004. Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. Nature 431:590-596.

151. Crooke E, Guthrie B, Lecker S, Lill R, Wickner W. 1988. ProOmpA is stabilized for membrane translocation by either purified E. coli trigger factor or canine signal recognition particle. Cell 54:1003-1011.

152. Kramer G, Rauch T, Rist W, Vorderwülbecke S, Patzelt H, Schulze-Specking A, Ban N, Deuerling E, Bukau B. 2002. L23 protein functions as a chaperone docking site on the ribosome. Nature 419:171-174.

153. Reid BG, Fenton WA, Horwich AL, Weber-Ban EU. 2001. ClpA mediates directional translocation of substrate proteins into the ClpP protease. Proceedings of the National Academy of Sciences 98:3768-3772.

154. Hoskins JR, Singh SK, Maurizi MR, Wickner S. 2000. Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP. Proceedings of the National Academy of Sciences 97:8892-8897.

155. Hoskins JR, Pak M, Maurizi MR, Wickner S. 1998. The role of the ClpA chaperone in proteolysis by ClpAP. Proceedings of the National Academy of Sciences 95:12135-12140.

156. Wickner S, Gottesman S, Skowyra D, Hoskins J, McKenney K, Maurizi MR. 1994. A molecular chaperone, ClpA, functions like DnaK and DnaJ. Proceedings of the National Academy of Sciences 91:12218-12222.

157. Hoskins JR, Kim S-Y, Wickner S. 2000. Substrate recognition by the ClpA chaperone component of ClpAP protease. Journal of Biological Chemistry 275:35361-35367.

158. Torres-Delgado A, Kotamarthi HC, Sauer RT, Baker TA. 2020. The intrinsically disordered N-terminal extension of the ClpS adaptor reprograms its partner AAA+ ClpAP protease. Journal of Molecular Biology 432:4908-4921.

159. Snider J, Thibault G, Houry WA. 2008. The AAA+ superfamily of functionally diverse proteins. Genome biology 9:1-8.

160. Baker TA, Sauer RT. 2012. ClpXP, an ATP-powered unfolding and protein-degradation machine. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1823:15-28.

161. Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. Annual Review of Biochemistry 80:587-612.

162. Kress W, Mutschler H, Weber-Ban E. 2009. Both ATPase domains of ClpA are critical for processing of stable protein structures. Journal of Biological Chemistry 284:31441-31452.

163. Baytshtok V, Baker TA, Sauer RT. 2015. Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proceedings of the National Academy of Sciences 112:5377-5382.

164. Kotamarthi HC, Sauer RT, Baker TA. 2020. The non-dominant AAA+ ring in the ClpAP protease functions as an anti-stalling motor to accelerate protein unfolding and translocation. Cell Reports 30:2644-2654. e3.

165. Ogura T, Wilkinson AJ. 2001. AAA+ superfamily ATPases: common structure–diverse function. Genes to Cells 6:575-597.

166. Olivares AO, Baker TA, Sauer RT. 2016. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nature Reviews Microbiology 14:33-44.

167. Dougan DA, Reid BG, Horwich AL, Bukau B. 2002. ClpS, a substrate modulator of the ClpAP machine. Molecular cell 9:673-683.

168. Renda BA. 2016. Laboratory evolution and natural transformation in *Acinetobacter baylyi* ADP1.

169. Singh A, Bedore SR, Sharma NK, Lee SA, Eiteman MA, Neidle EL. 2019. Removal of aromatic inhibitors produced from lignocellulosic hydrolysates by *Acinetobacter baylyi* ADP1 with formation of ethanol by *Kluyveromyces marxianus*. Biotechnology for Biofuels 12:91.

CHAPTER 2

EVOLUTION OF THERMOADAPTATION BY TARGETED GENE AMPLIFICATION

Introduction

Aromatic compounds are prevalent in nature and are synthesized by plants, which incorporate the aromatic components into a heterogenous polymer, lignin. Due to the chemical stability of the benzene ring, aromatic compounds resist degradation (1). In nature, bacteria play a key role in degrading aromatic compounds derived from lignin as part of the carbon cycle because, unlike most higher organisms, they can capture the oxidative power of molecular oxygen by using metal-dependent dioxygenases to cleave the benzene ring (2, 3). Current research focuses on taking advantage of this natural process of microbial lignin degradation for biotechnology to convert lignin into industrially relevant products (3-6). These biological methods of lignin valorization seek to engineer bacterial pathways to convert lignin-derived aromatic compounds into valuable chemicals.

Acinetobacter baylyi ADP1 was tested in a set of microbial hosts for the metabolism of a biomass-derived, lignin-enriched feedstock (7). ADP1 was among the bacteria demonstrating greatest potential for use in lignin bioprocessing that combines simultaneous depolymerization and product generation (7). Specifically, ADP1 can convert the carbon and energy extracted from lignin into wax esters, which are economically valuable storage compounds (8-10). Recent developments in synthetic biology tools for ADP1 also (11) make it an attractive target for biotechnology due to the ease of genetic modification and laboratory evolution (12-15).

There are many challenges to microbial lignin valorization strategies. To date, no single organism can degrade all lignin-derived aromatic compounds. Furthermore, initial processing of lignin typically generates complex mixtures of many aromatic and non-aromatic compounds. When multiple compounds are present, degradative pathways are often tightly regulated to constrain sequential consumption of individual components (16-21). ADP1 converts each of the aromatic compounds it degrades into either of two substrates of ring-cleavage enzymes, catechol or protocatechuate (PCA). These compounds are then degraded through parallel branches of the well-studied β-ketoadipate pathway (22). In ADP1, aromatic compounds degraded through the catechol branch tend to be consumed before those degraded through the PCA branch (12, 23-25). This regulation likely evolved in environments where the hierarchical consumption of varied substrates is beneficial. For biotechnology purposes, it may be necessary to modify metabolic pathways of ADP1 and/or override some of the natural regulatory features to maximize the rapid and simultaneous conversion of diverse lignin-derived substrates to desired end products.

In this study, we sought to modify the PCA degradation pathway of ADP1, which involves *ortho* cleavage of the aromatic ring between the hydroxyl substituents (3,4-cleavage, **Figure 2.1**) (22). Other bacteria degrade PCA using *meta*-cleavage pathways that involve cleavage of the aromatic ring on either side of the hydroxyl groups (2,3- or 4,5-cleavage, **Figure 2.1**) (26, 27). As a first step in expanding the metabolic capabilities of ADP1, we replaced the PCA 3,4-cleavage pathway with a non-native PCA-4,5 cleavage pathway using DNA from a marine bacterium, *Pseudoalteromonas atlantica* T6c. There were several reasons for this experimental strategy. First, we wanted to establish methodology for the integration and functional adaptation of foreign DNA in ADP1. The facile genetic system of ADP1 allows large segments of foreign DNA to be incorporated directly into the ADP1 chromosome to broaden the

range of compounds catabolized. Once integrated, this foreign DNA can also be adapted by a method that mimics natural evolution by duplication and divergence. This method, termed EASy (**E**volution by **A**mplification and **Sy**nthetic Biology), has previously been used in *A. baylyi* to evolve and optimize one or two metabolic steps to allow the consumption of non-native aromatic compounds, such as guaiacol and terephthalic acid (28, 29).

Our current studies of the PCA-4,5 pathway represent the most challenging chromosomal integration efforts to date, involving 6 enzymatic steps. During aromatic compound degradation, the accumulation of toxic intermediates can make it difficult to express multiple foreign genes at appropriate levels. Necessary information for the rational design of metabolic flux may not exist or even be predictable. However, the EASy method selects cells that grow rapidly with a target growth substrate. The copy number of the foreign genes can readily increase or decrease, thereby providing a rudimentary method to vary expression levels. Gene amplification also provides increased DNA for mutations that can enable the selection of better-functioning alleles. After strains are selected that can use the PCA-4,5 pathway, it should be possible to expand degradative capabilities modularly. For example, as shown in **Figure 2.1**, the addition of another pathway module could allow syringaldehyde and other S-lignin derivatives to be funneled into central metabolism via a common metabolite (4).

Another reason for altering the PCA degradative pathway in *A. baylyi* is to engineer different metabolic schemes. Thee three PCA degradation pathways use different redox balance and yield different products entering central metabolism (**Figure 2.1**). There could be advantages to each pathway, depending on the circumstances. Lastly, it is possible that by establishing utilization of PCA via the non-native 4,5-pathway there will be perturbations in the regulation of

multiple carbon source utilization that may improve the simultaneous consumption of different compounds and/or reveal the molecular basis of novel regulatory mechanisms.

As described in this report, the initial chromosomal integration of the foreign DNA enabled slow PCA consumption, but not at temperatures above 25 °C. ADP1 is typically grown at 30 °C or 37 °C, depending on the application (12, 13). After applying the EASy method, strains were isolated that grew more rapidly on PCA using the 4,5-cleavage pathway than they did initially. Furthermore, whole genome sequencing revealed mutations that emerged during laboratory evolution. By exploiting the natural transformability of ADP1 to test the functional significance of mutations (30), genetic changes were identified that improve thermotolerance to allow growth with the foreign pathway at increased temperatures (up to 30 °C).

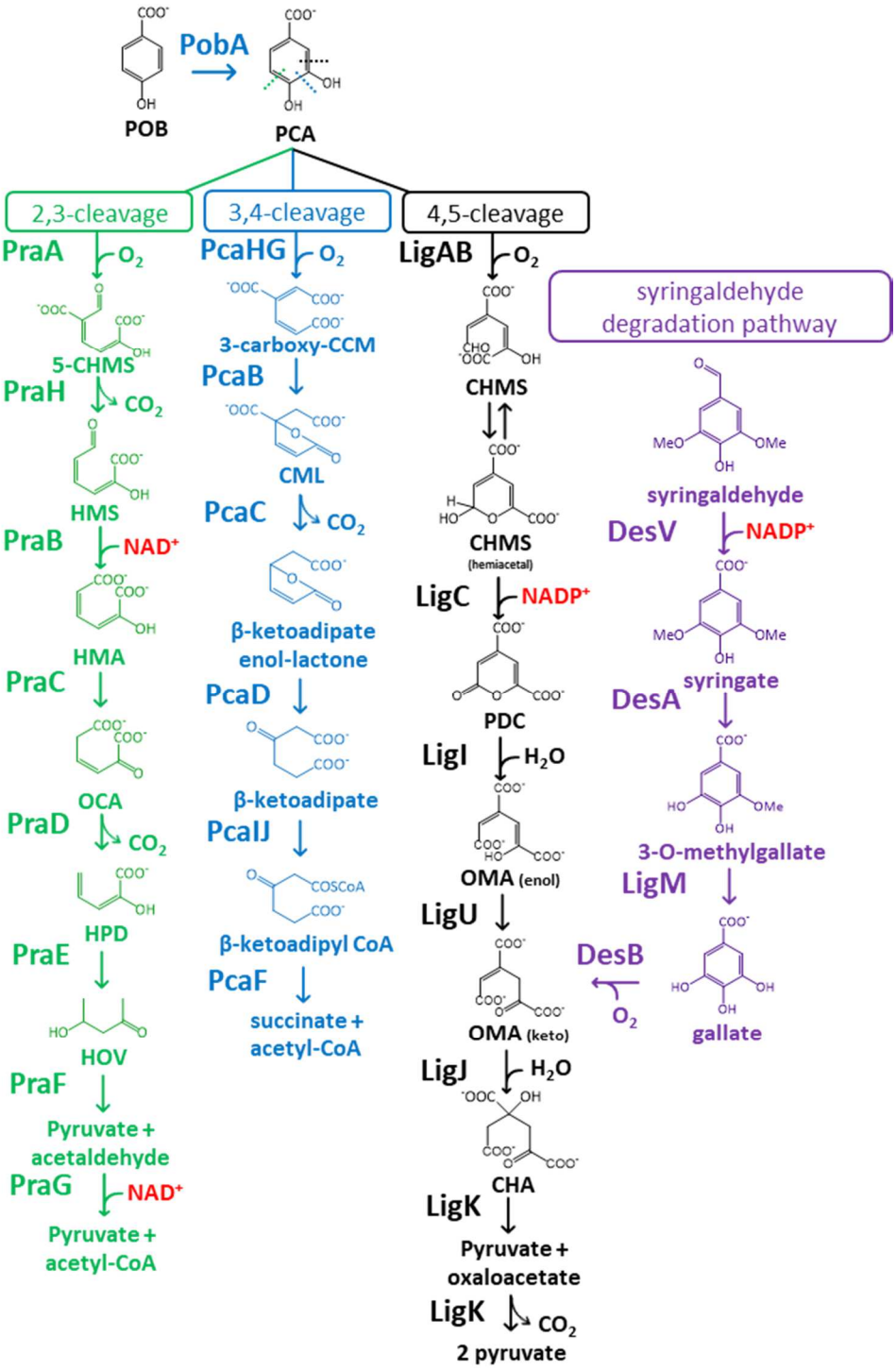
Results

Heterologous expression of *P. atlantica* T6c predicted PCA-4,5 degradation pathway

To allow ADP1 to utilize a foreign PCA degradation pathway, ADP1's native PCA-3,4 pathway was first removed from strains. Genes encoding the PCA-3,4 pathway exist in ADP1 as a single *pca* operon with an IclR-type transcriptional regulator, *pcaU*, encoded divergently on the chromosome (**Figure 2.2**) (31). Two strains were constructed missing some or all the *pca* genes. The first strain, ACN462, had only the genes encoding the native PCA dioxygenase, *pcaHG,* removed from the chromosome, with the rest of the operon left intact (32). The strain was unable to grow on PCA. It also failed to use compounds funneled through PCA, such as *p*-hydroxybenzoate (POB), as the sole carbon source (33). While ACN462 did not catabolize PCA, it should be able to transport PCA since the native PCA/POB transporter, *pcaK*, remained functional and natively expressed. This genetic background would thus avoid potential issues

with transport of PCA and POB. ACN462 was later found to contain a spontaneous deletion of a 90 kbp segment from a historically unstable region of the ADP1 chromosome (**Figure 2.2 Panel B**) (34-36), as discussed in greater detail later. The second strain, ACN2106, had the entire *pca* operon removed. The PCA/POB transporter *pcaK* was put back in the chromosome under the control of a synthetic promoter. Because this strain does not have any of the native *pca* genes, there should be no interference from the native PCA-3,4 pathway or its regulator, PcaU, during growth using the foreign pathway.

Furthermore, ADP1 encodes an additional transporter of PCA and POB, VanK. VanK and PcaK are both members of the major facilitator superfamily (MFS) of transporters and have been shown to transport multiple aromatic compounds including PCA and POB (37). The gene encoding VanK is not located near the *pca* operon in the chromosome and should, consequently, not be affected by manipulations to the *pca* operon. Therefore, in ADP1-derived strains missing all or part of the native PCA-3,4 cleavage pathway, the aromatic compounds PCA and POB should still be transported into the cell.



2,3-*meta* cleavage	3,4-*ortho* cleavage	4,5-*meta* cleavage	**Pathway**
2 NAD^+	none	1 $NADP^+$	**Reducing power used**
2 (C3)	1 (C4) + 1 (C2)	1 (C3) + 1 (C2)	**Carbons entering central metabolism per PCA molecule**
2	1	1	**CO_2 generated**

Figure 2.1: Alternative PCA degradation pathways

Three bacterial PCA degradation pathways: 2,3-cleavage (green), 3,4- cleavage (blue), and 4,5-cleavage (black). Each PCA pathway leads to the production of different products feeding into the tricarboxylic acid (TCA) cycle. The pathways also utilize different reducing power, depicted in red. In purple is the potential modular expansion for introducing a pathway for the catabolism of syringaldehyde, modeled from that of *Sphingobium sp*. SYK6, which feeds into an intermediate of the PCA-4,5 pathway. Abbreviations: 5-CHMS (5-carboxy-2-hydroxymuconate-6-semialdehyde), HMS (2-hydroxymuconate-6-semialdehyde), HMA (2-hydroxymuconate), OCA (4-oxalocrotonate), HPD (2-hydroxypenta-2,4-dienoate), HOV (4-hydroxy-2-oxovalerate), CCM (*cis,cis*-muconate), CML (γ-carboxymuconolactone), CHMS (4-carboxy-2-hydroxymuconate-6-semialdehyde), PDC (2-pyrone-4,6-dicarboxylate), CHM (4-carboxy-2-hydroxymuconate), and OMA (4-oxalomesaconate)

A PCA-4,5 catabolic pathway from *Pseudoalteromonas atlantica* T6c was chosen for this study. It had previously been predicted by sequence homology to have a PCA-4,5 pathway (38, 39). The pathway from this organism was chosen over those that had already been characterized in *Sphingobium sp*. SYK6 and *Comamonas testosteroni* CNB-1 because these strains have a high G+C content (63.4% and 61.5% respectively), whereas *A. baylyi* and *P. atlantica* have similar low G+C content (40.4% and 42.8% respectively) (27, 39). Additionally, the *Sphingobium sp*. PCA-4,5 pathway is split up in two separate operons. The *P. atlantica* pathway, on the other hand, may exist in a single operon with only catabolic pathway genes and a gene encoding a LysR-type transcriptional regulator (LTTR) transcribed divergently from its pathway genes.

Although transcription has not been studied, the neighboring genes are closely grouped together. We predict that this LTTR controls expression of the *P. atlantica* catabolic genes because this chromosomal organization is characteristic of LTTRs and a similar regulator has been characterized in controlling expression of the *Sphingobium sp*. PCA-4,5 catabolic genes (40, 41). Consequently, for simplicity and for the hope of optimal heterologous expression, the *P. atlantica* pathway was preferred as the source of DNA to integrate into the ADP1 chromosome.

We used two ADP1 chromosomal locations for incorporation of the foreign PCA-4,5 genes from *P. atlantica*, referred to hereafter as *lig* genes following their well-researched homologs in *Sphingobium sp*. SYK6 (27). One location is within one of five catabolic islands in the ADP1 chromosome that contain large clusters of catabolic genes (42). Catabolic island IV was selected as a location for the insertion of the *lig* genes because it is where the native *pca* operon is located as well as genes encoding enzymes for the degradation of other aromatics that are catabolized through PCA. Specifically, the *lig* genes were put in between the convergently transcribed *quiA*, encoding quinate/shikimate dehydrogenase, and *pobS*, encoding a regulatory protein. The second location for genomic integration of the foreign DNA is upstream catabolic island IV in between two non-essential hypothetical genes, ACIAD_RS07700 and ACIAD_RS07705 (43). This location was chosen so as to be separate from other clusters of aromatic compound catabolism genes. To increase chances for success and to learn more about chromosomal manipulation and evolution, multiple strains were constructed in which the *lig* genes were incorporated in different genomic locations. In some strains the *lig* genes were orientated on the chromosome to be replicated on the leading DNA strand, and in other strains these genes would be replicated on the lagging strand. In this fashion, we could test whether

there was a strand based effect on heterologous gene expression for this foreign catabolic pathway (44).

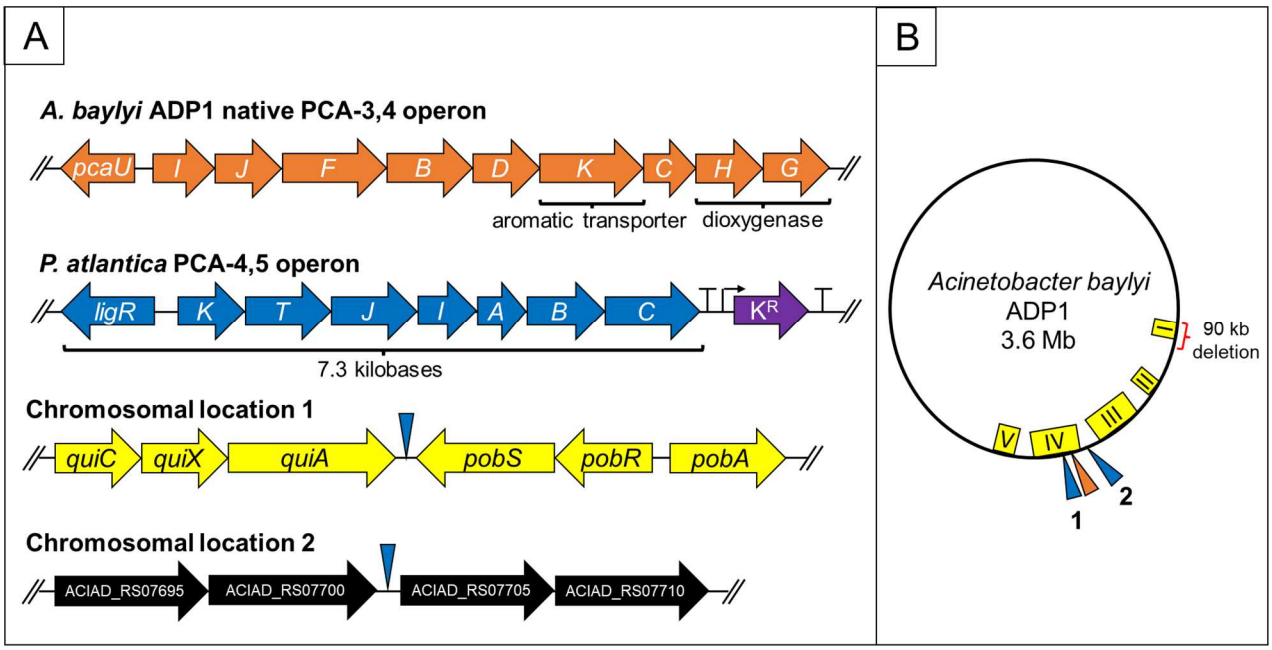


Figure 2.2: Genetic constructs and chromosomal locations

Panel A: In orange is the *A. baylyi* native *pca* operon and upstream regulator coding sequence, *pcaU*. Blue is the *P. atlantica* predicted PCA-4,5 pathway construct used in this study. An ΩK^R cassette, genetic construct conferring kanamycin resistance flanked by transcription and translation termination signals (45, 46), was inserted downstream of *ligC* to select for chromosomal integration and gene amplification. In yellow and black are the two locations on the ADP1 chromosome where the PCA-4,5 pathway was inserted in different strains. One is within the catabolic island IV, that also houses the native *pca* genes. The other is in a location upstream of catabolic island IV in between two hypothetical genes. Panel B: Depiction of where on the ADP1 chromosome the catabolic islands are (yellow), the PCA-4,5 operon was inserted

(blue), location of the native *pca* genes (orange), and location of the 90 kbp deletion including catabolic island I.

The PCA-4,5 pathway was integrated into the chromosome, along with a downstream cassette encoding drug resistance (ΩK^R) that facilitates genomic manipulation. The pathway allowed all the resulting ADP1-derived strains (ACN1910, ACN2118, ACN2119, ACN2120, and ACN2121) to grow on PCA or POB as the sole carbon source. However, growth was only observed at temperatures up to 25 °C. The doubling times of strains on POB at 25 °C are reported in **Table 2.1**. All strains with the foreign pathway grew slower than the wild-type (WT) strain with the native PCA-3,4 pathway. The WT strain doubled in about 72 min. The other strains utilizing the foreign PCA-4,5 pathway had doubling times approximately 12 to 54 min longer than ADP1 using the native PCA-3,4 pathway. There was no correlation between the doubling time and the orientation or location in the chromosome of the *lig* genes. We next sought to improve growth of the mutants through the foreign PCA-4,5 pathway. Because ADP1 grows optimally at 37 °C and the goal is ultimately to use these strains for biotechnology, we sought to use laboratory evolution to increase their thermotolerance while utilizing the foreign PCA-4,5 pathway.

Strain	*pca* genes deleted	90 kbp deletion (Yes/No)	Replication strand orientation for PCA catabolic genes	*lig* gene location	Doubling time (Hours)	Standard deviation
ADP1 (WT)	N/A	No	Leading (*pca*)	N/A	1.26	0.04
ACN1910	*pcaHG*	Yes	Lagging (*lig*)	1	1.96	0.14
ACN2118	*pcaUIJFBDCHG*	No	Leading (*lig*)	1	1.51	0.13
ACN2119	*pcaUIJFBDCHG*	No	Lagging (*lig*)	2	1.47	0.07
ACN2120	*pcaUIJFBDCHG*	No	Leading (*lig*)	1	1.51	0.07
ACN2121	*pcaUIJFBDCHG*	No	Lagging (*lig*)	2	2.11	0.41

Table 2.1: Doubling times of ADP1-derived strains on POB at 25 °C

All strains were grown on POB at 25 °C. All were done in biological triplicate. Doubling time was calculated as Doubling time = ln2 / growth rate, with growth rate calculated by the line of best fit (log scale) during log growth phase with R values ≥ 0.98.

Laboratory evolution by chromosomal gene amplification to increase thermotolerance

To increase the growth temperature of the ADP1-derived strains expressing the PCA-4,5 pathway, the EASy method of laboratory evolution was used (28, 29). This method increases the gene dosage of a specific chromosomal region, which in turn allows higher levels of the gene products to be expressed. If such a change is beneficial under selective conditions, the cells can overcome a growth barrier. In this study, cells were selected to grow via use of the foreign PCA-4,5 pathway at temperatures above 25 °C. The method involves a precise chromosomal duplication that is initiated by the natural transformation of a recipient strain using, as the donor DNA, a synthetic bridging fragment (SBF) that serves as a platform for homologous

recombination. Evolution of the duplicated populations can occur as cultures undergo daily serial transfers (**Figure 2.3 Panel A**).

For these studies of the PCA-4,5 pathway, Linearized pBAC1415 served as the SBF for recipient strains with the *lig* genes in location 1 (ACN1910, ACN2118, and ACN2120). For recipient strains with the *lig* genes in location 2 (ACN2119, and ACN2121), linearized pBAC1414 served as the SBF. After transformation with the SBF, successful gene duplication and subsequent amplification events were selected by growth on plates containing high levels of kanamycin (Km, 1 mg/ml). This high level of resistance is conferred by multiple copies of the ΩK^R cassette that is co-amplified with the *lig* genes. Thus, Km selection generates a starting population with multiple copies of the PCA 4,5 pathway. Isolates from each parent strain, shown in **Figure 2.3 Panel B**, were obtained as individual colonies growing on high levels of Km. Because of the instability caused by having multiple chromosomal copies of the same region, we considered these isolates to be heterogenous populations. Several such populations were chosen for long-term culturing. Three of these were derived from ACN2118, ACN2119, and ACN2121 and one population was derived from ACN1910 and ACN2120 (**Figure 2.3 Panel B**).

After obtaining isolates that grew on high concentrations of Km, the drug selection was removed. Cells were next challenged to grow on plates containing a sole carbon source consumed through the foreign pathway, PCA or POB, at 37 °C. Cells that grew under these conditions were moved to 5 ml liquid culture. Selective pressure for growth at 37 °C on PCA or POB was then maintained as populations were sub-cultured daily. For the first few days of culturing, the inoculum was diluted 10-fold during serial transfer. Cells were subcultured only once per day in test tubes with 5 ml of fresh medium. This culturing method limited the total number of daily generations achieved as cells consumed the growth substrates and entered

stationary phase. Over time, as beneficial mutations accrued, it became possible to transfer a smaller inoculum, which enabled an increased number of daily generations. This change was implemented gradually during the first 21 days of serial transfers until 1000-fold dilutions were carried out daily. Populations were cultured for a total of 115 days corresponding to approximately 1800 generations.

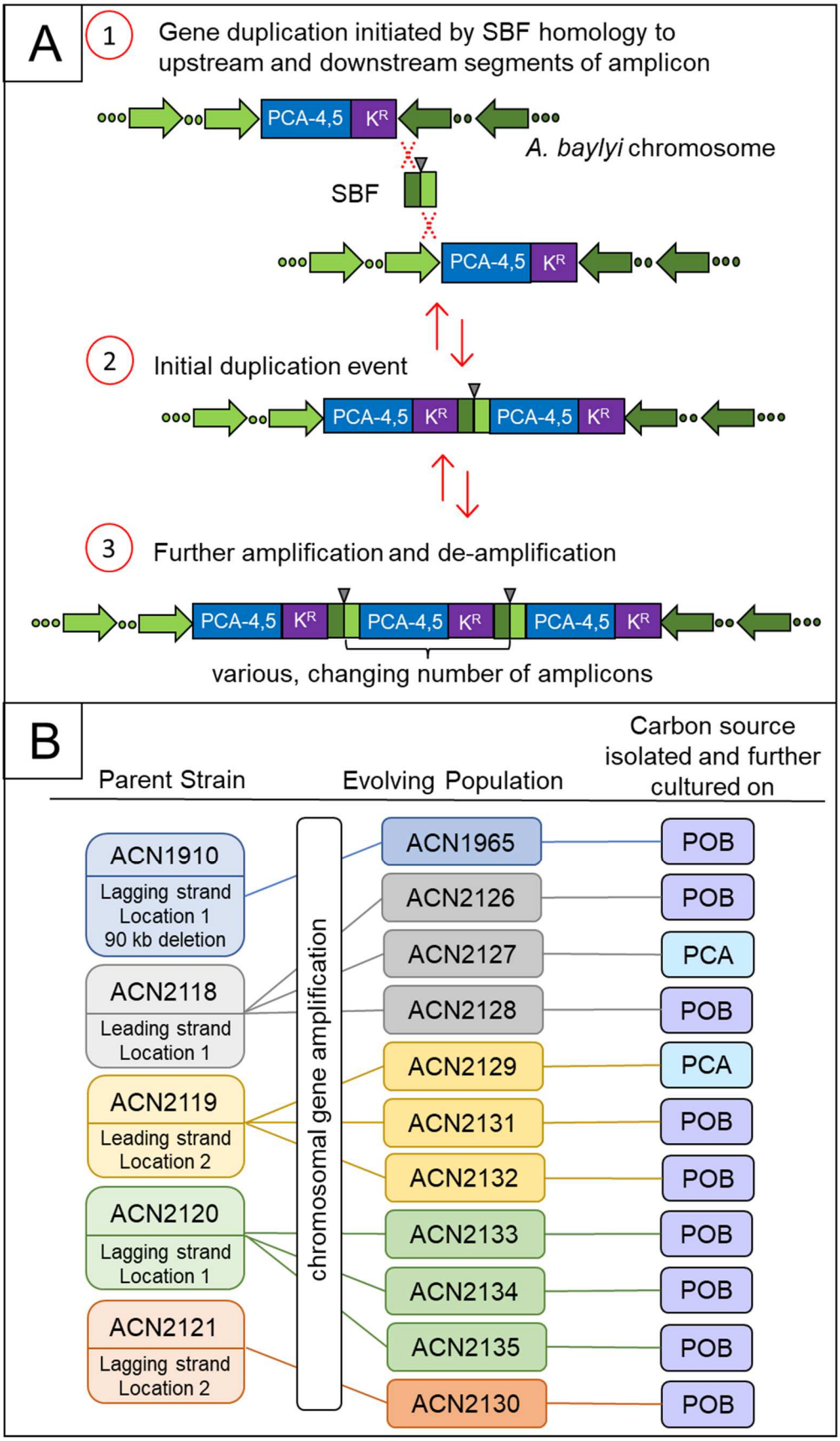

A
1
Gene duplication initiated by SBF homology to upstream and downstream segments of amplicon
PCA-4,5
K^R
A. baylyi chromosome
SBF
2
Initial duplication event
3
Further amplification and de-amplification
various, changing number of amplicons
B
Parent Strain
Evolving Population
Carbon source isolated and further cultured on
ACN1910
Lagging strand Location 1 90 kb deletion
ACN2118
Leading strand Location 1
ACN2119
Leading strand Location 2
ACN2120
Lagging strand Location 1
ACN2121
Lagging strand Location 2
chromosomal gene amplification
ACN1965
ACN2126
ACN2127
ACN2128
ACN2129
ACN2131
ACN2132
ACN2133
ACN2134
ACN2135
ACN2130
POB
POB
PCA
POB
PCA
POB
POB
POB
POB
POB

Figure 2.3 Laboratory evolution by EASy

Panel A: Schematic of EASy method. A precise tandem duplication is initiated by transformation with the synthetic bridging fragment (SBF). The duplication provides additional sequence identity for further amplification by homologous recombination. These amplification events occur stochastically and are unstable. The number of amplicons can expand and contract as selected during evolution. Panel B: Names and origin of EASy evolving populations. The parent strain designation indicates a mutant with one copy of the PCA-4,5 pathway integrated in the chromosome in one of two locations, oriented in one of two directions. As noted in the text, ACN1910 was derived from a Δ*pcaGH* strain that was later found to have a 90-kbp deletion in a different region known to undergo frequent, spontaneous deletions (36). After transformation with the SBF and selection for a tandem array in the chromosome of the region carrying the foreign DNA, the population used for laboratory evolution was given a new designation. For three of the parent strains, multiple lineages were maintained in parallel experiments after amplification.

Using qPCR to follow adaptation of evolving populations

As beneficial mutations spontaneously arise during long-term culturing, the cells can lose unnecessary copies of the amplicon (47). If these mutations enhance growth, cells carrying them will outcompete other cells, and these mutants become enriched in the evolving population. During the course of the experiment, the average number of amplicons per chromosome in the population can be followed by quantitative (q) PCR. Copy number of a PCR fragment from the amplicon is measured relative to the copy number of a fragment generated from the *rpoA* gene, RNA polymerase subunit A, because it is assumed to be in single copy in the *A. baylyi*

chromosome. In this experiment, a specific PCR fragment from the coding sequence of the Km-resistance gene was used to assess copy number (**Figure 2.4**). The copy number of this small fragment in the amplicon, designated F87, was used as an approximation of the gene dosage for the amplified region, since typically the tandem array increases or decreases by integral amplicon units (48). However, it is possible that chromosomal rearrangements result in changes that are not detected by this qPCR method.

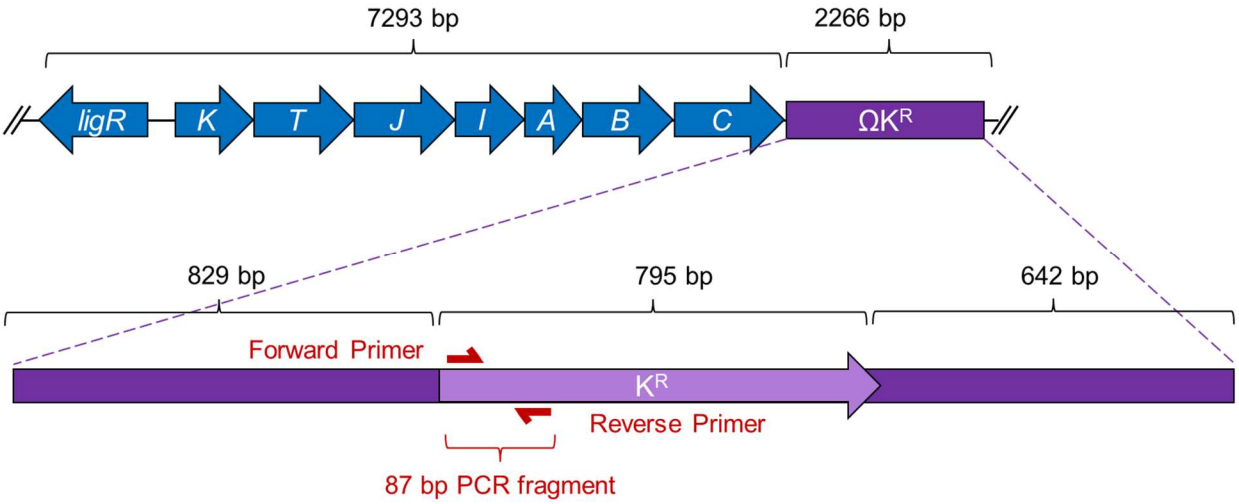


Figure 2.4: DNA region of the amplicon measured by qPCR

Depicted here is the 87 bp fragment, nucleotides 7-94, in the K^R coding sequence (light purple arrow), used to quantify copy number. Gene dosage of this fragment (F87) provides an estimate of the copy number of chromosomal amplicons during long term growth experiments. Copy number was assessed relative to a fragment from *rpoA*, encoding RNA polymerase subunit A, presumed to exist in single copy in the chromosome, as previously described (28).

Genomic DNA was extracted from samples of the whole population taken at various time points, and the qPCR method was used to assess gene dosage with four technical replicates conducted using the same DNA sample as template. The first genomic DNA preparation analyzed by qPCR came from the population grown on a plate containing PCA or POB as the sole carbon source at 37 °C. This sample corresponds to the initial population growing at a higher temperature under selective conditions that require *lig*-gene functionality. As the populations underwent long-term culturing in liquid media, samples were taken, and genomic DNA was isolated once per week for 16 weeks. Copy number therefore reflects an average of the entire population for each sample.

After initial selection for growth on plates with PCA or POB at 37 °C populations had an average F87 copy number between 27 and 172 (purple bars **Figure 2.5 Panel A**). While the initial copy number varied, gene dosage rapidly decreased for all populations (**Figure 2.5 Panel B**). Copy number remained stable for each strain between weeks 6 and 16, with average copy numbers between 4 and 13 (blue bars **Figure 2.5 Panel A**). The decrease in copy number suggested that cells within each population accumulated beneficial mutations that enabled growth at a higher temperature than the corresponding parent strain. Single colonies were isolated from populations for further examination. Colonies were isolated from the week in which each population had the lowest average copy number of the fragment as evaluated by qPCR.

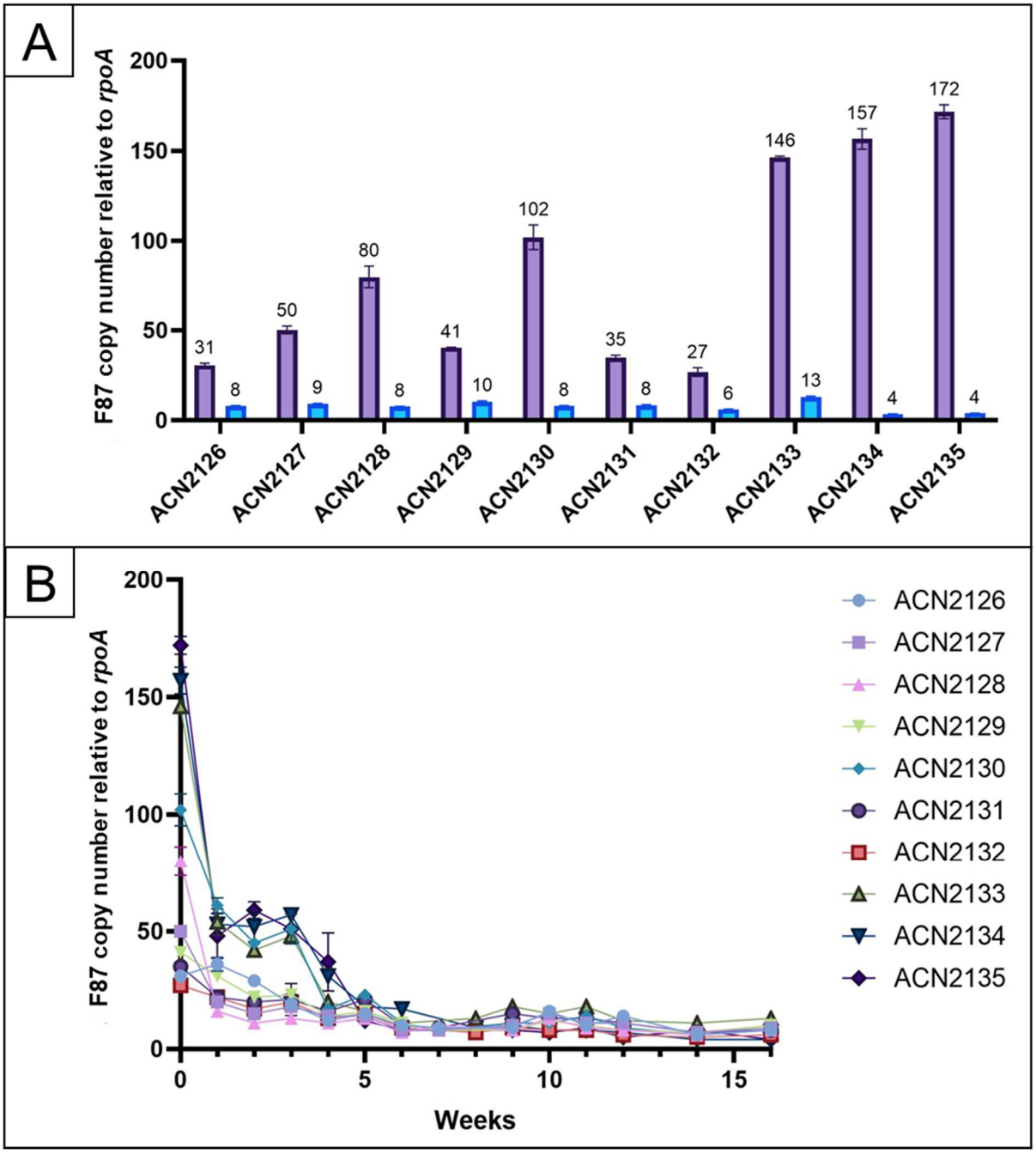


Figure 2.5: Copy number change over time of *A. baylyi* populations

Data here represent the average copy number of the amplicon as measured by F87 (fragment in the K^R gene serves as a proxy for amplicon copy number) relative to *rpoA* during long-term growth experiments. Panel A: Purple bars represent F87 copy number of the population after the first exposure to selection (growth on PCA or POB at 37 °C). Blue bars represent final copy

number at day 115 of long-term culturing. Panel B: Weekly copy number change for all evolving populations. All populations reduced F87 copy number over time and seemed to stabilize around week 6. Error bars indicate standard deviation of four technical replicates.

Although the F87 copy number is used to approximate the amplicon gene dosage, chromosomal rearrangements or other mutations can occur that result in the qPCR measurement not reflecting the level of *lig*-gene amplification. Such a rearrangement likely occurred in the population derived from strain ACN1910, named ACN1965. The F87 qPCR indicated an average amplicon copy number of 1 for this population after initial selection for growth on POB at 37 °C. This result was unexpected since the populations only had a single day of exposure to POB at 37 °C, and it did not seem possible that multiple *lig*-gene copies could be lost this quickly. A single colony was isolated from this population at week 5, designated ACN2041, and used for further analysis.

Analyzing chromosomal changes that accumulate during long-term culturing

The reduced copy number of F87 suggested that mutations elsewhere in the chromosome could be contributing to thermotolerance utilizing the *lig* genes at temperatures above 25 °C. I thus developed a genetic approach to test this hypothesis. All remaining copies of the *lig* genes, encoding the PCA-4,5 pathway, in strains that underwent long-term culturing experiments were removed from the chromosome and replaced with a single un-mutated copy of the *lig* genes (**Figure 2.5**). Because this approach would leave strains with just a single copy of the original *lig* genes, any thermotolerant growth phenotype observed should be the result of mutations elsewhere in the chromosomes.

Steps in this method are depicted in **Figure 2.6**. First, the multiple amplicon copies, possibly carrying adaptive mutations (grey stars **Figure 2.6 Panel A**), were removed from the chromosome and replaced with a *sacB*-S^R selectable/counter-selectable marker. Allelic replacement of the amplicon array with the cassette was selected by resistance to streptomycin (Sm) and spectinomycin (Sp) conferred by S^R. Because these strains also lost all copies of the *lig* genes and ΩK^R, they could not grow in the presence of Km or by using PCA or POB as carbon sources. These strains were then used as the recipients for a transformation using linear donor DNA with a single copy of the unmutated original *lig*-gene pathway. In this case, the allelic replacement method used counterselection with the *sacB* gene, encoding levan sucrase. Since SacB expression is lethal in the presence of sucrose, this compound was added to the medium to select transformants in which *sacB* was removed by homologous recombination that re-introduced the *lig*-genes into the chromosome (**Figure 2.6 Panel B**).

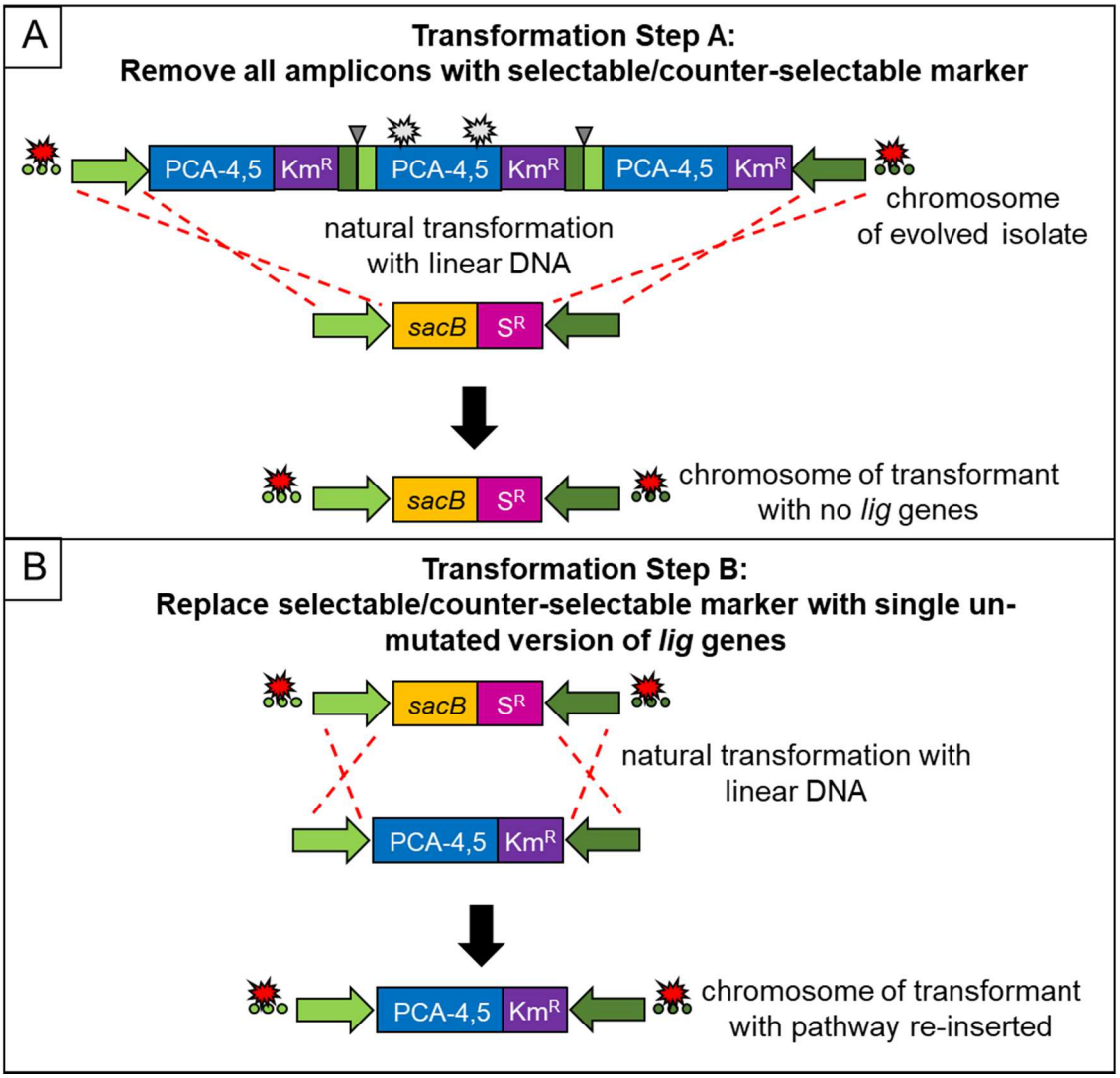


Figure 2.6: Genetic strategy to replace multiple copies of *lig* genes with an un-mutated copy

Multiple amplicons within the chromosomes of strains that underwent long-term culturing experiments were first replaced by homologous recombination (red dotted Xs) with a selectable/counter-selectable marker shown as the orange and pink boxes (A). Successful transformants from step A then underwent a subsequent transformation and homologous recombination (red dotted Xs) to replace the selectable/counter-selectable marker with an unmutated copy of the *lig* genes (B).

This methodology was successful for six isolates from the evolved populations. For these six isolates, all amplicons were replaced with a single, un-mutated copy of the *lig* genes (**Figure 2.7 Panel B**). These six strains were transferred to medium with POB as the sole carbon source, and five of the strains grew on POB as the sole carbon source at 30 °C, a temperature that was higher than the maximal growth temperature of the parent strains, which did not grow on POB or PCA above 25 °C. The strains that grew at 30 °C were renamed in case new mutations were selected during the transition to these final growth conditions (ACN2711, ACN2712, ACN2713, ACN2714, and ACN2718) (**Figure 2.7**). Although the long-term growth experiments were carried out at 37 °C, none of the transformants constructed as shown in Fig. 2.7 grew at 37 °C. This result indicates that mutations accrued in the chromosomes of some of the evolved strains are contributing to the thermotolerant phenotype of these strains. Nevertheless, multiple copies of the *lig* genes still appear to be necessary for strains to grow at 37 °C. In one case, replacement of the amplicons with a an unmutated copy of the pathways did not enable growth on POB at temperatures above 25 °C. Any chromosomal mutations present in this strain (ACN2538) do not appear to be sufficient to increase the temperature at which it can grow using the foreign pathway. Chromosomal mutations from all evolved isolates were identified through whole genome re-sequencing (WGS).

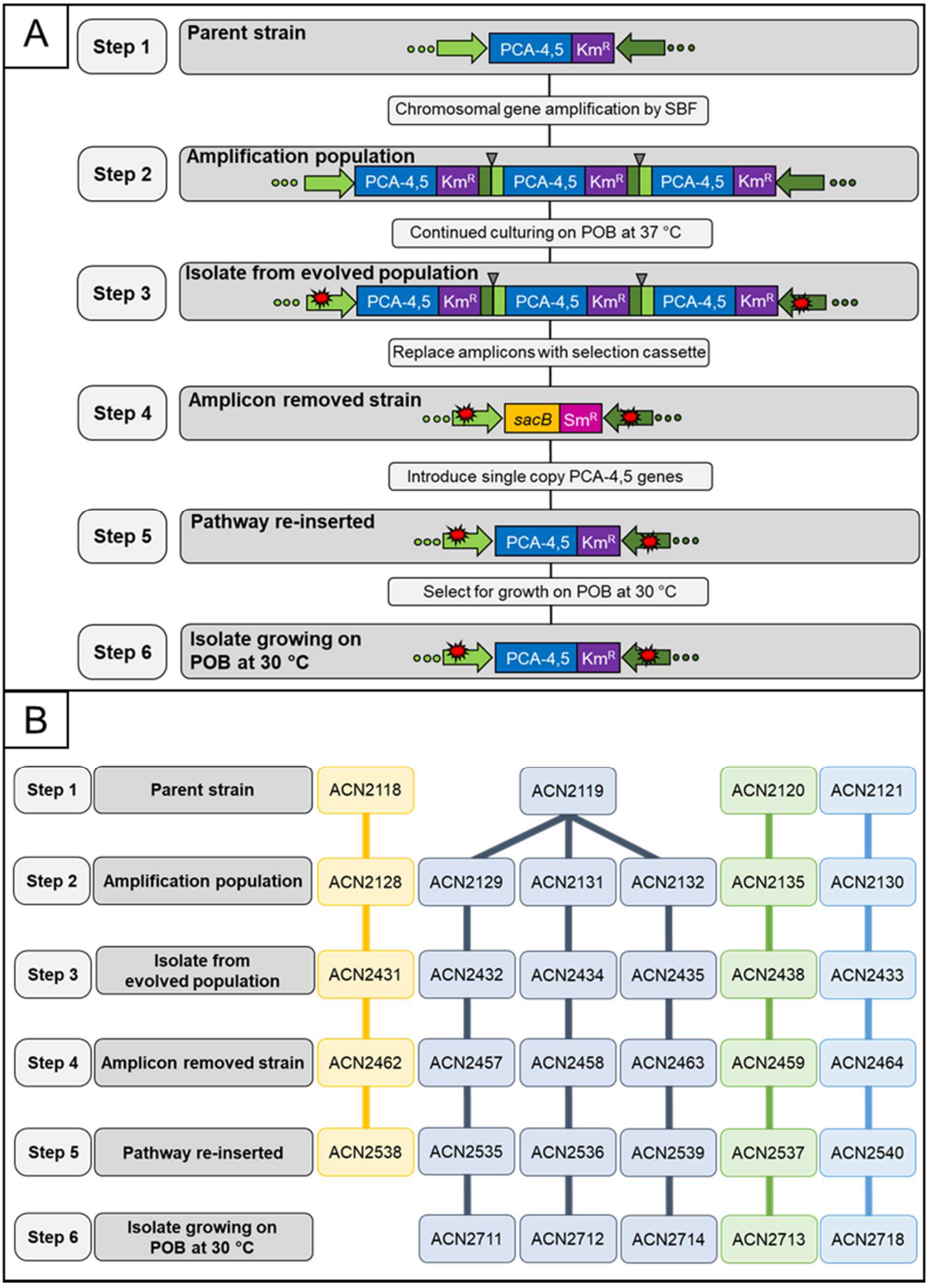

A
Step 1
Parent strain
PCA-4,5
KmR
Chromosomal gene amplification by SBF
Step 2
Amplification population
Continued culturing on POB at 37 °C
Step 3
Isolate from evolved population
Replace amplicons with selection cassette
Step 4
Amplicon removed strain
sacB
SmR
Introduce single copy PCA-4,5 genes
Step 5
Pathway re-inserted
Select for growth on POB at 30 °C
Step 6
Isolate growing on POB at 30 °C
B
Step 1
Parent strain
ACN2118
ACN2119
ACN2120
ACN2121
Step 2
Amplification population
ACN2128
ACN2129
ACN2131
ACN2132
ACN2135
ACN2130
Step 3
Isolate from evolved population
ACN2431
ACN2432
ACN2434
ACN2435
ACN2438
ACN2433
Step 4
Amplicon removed strain
ACN2462
ACN2457
ACN2458
ACN2463
ACN2459
ACN2464
Step 5
Pathway re-inserted
ACN2538
ACN2535
ACN2536
ACN2539
ACN2537
ACN2540
Step 6
Isolate growing on POB at 30 °C
ACN2711
ACN2712
ACN2714
ACN2713
ACN2718

Figure 2.7: Genetic method to analyze chromosomal mutation in evolved strains

Panel A: Depiction of *A. baylyi* chromosomes with the foreign PCA-4,5 pathway at various stages of experimental manipulation (hypothetical chromosomal mutations designated as red stars). Panel B: Progression of strain construction resulting in strains with a single copy of the PCA-4,5 pathway that can grow at temperatures up to 30 °C (step 6).

Whole genome re-sequencing of *A. baylyi* isolates growing through the PCA-4.5 pathway above 25 °C

All isolates resulting from EASy as well as the strains constructed with an unmutated pathway (**step 6, Figure 2.7**) underwent Illumina whole genome re-sequencing (WGS) to uncover any mutations that accrued during long-term culturing. WGS reads were mapped to the *A. baylyi* reference genome NCBI entry NC_005966. WGS analysis was done through the Breseq pipeline (49). Mutations found in each strain are reported in **Tables 2.2 and 2.4**. Larger chromosomal rearrangements are reported by their novel junction in **Table 2.3**. The frequency (or percentage) of reads that display the new junction can give insights about each specific rearrangement. For example, if a region of the chromosome were duplicated, the frequency of reads containing this junction would be 50%. Therefore, this frequency is included in **Table 2.3**. New IS*1236* insertion element events are also reported in **Table 2.3**.

Only a single mutation was found within the foreign *lig* genes (Patl_3889 in ACN2435). This result is likely due in part to all amplicon reads mapping to the same sequence in the reference genome. While no single mutation or mutated gene was found in all 11 evolved strains and 5 pathway re-inserted strains, there were a few common mutations that occurred in multiple lineages. The loci where mutations were often observed are discussed in the following sections.

Mutations selected in *ptsP*

Seven of the eleven strains that underwent the EASy method acquired mutations in ACIAD_RS02085 (new alleles summarized in **Table 2.5**). This locus was annotated as *ptsP* (**P**hospho**t**ransferase **s**ystem- PTS). In the introductory chapter of this thesis, I provide information about PtsP and its homologs. Briefly, this protein corresponds to the first enzyme in phosphorelay systems, and it is predicted to accept a high-energy phosphate group from phosphoenolpyruvate (PEP). Based on homology, PtsP from *A. baylyi* ADP1 has a GAF domain that is likely to be involved in signaling circuits using compounds such as metabolites involved in carbon/nitrogen metabolism (50). Homologs of this protein have two partners in a three-protein transfer system that does not involve transport of a sugar or other substrate into the cell. Rather, related PTS systems have typically been designated PTS^{Ntr} to indicate some role in nitrogen sensing, although this name has also been used for PTS systems with no known role in nitrogen metabolism. *A. baylyi* ADP1 does not encode a counterpart to the third protein in such systems (PtsN), although there is a homolog to another component (PtsO). The *ptsP* gene, ACIAD_RS02085, appears to be involved in an undescribed PTS in *A. baylyi*. It is unclear how many (and which) proteins participate with PtsP in a phospho-transfer relay. The final phosphate acceptor in this system is also unknown. PtsP does not appear to participate in the only canonical PTS in *A. baylyi*, which is involved in fructose uptake (43).

The frequency of independent *ptsP* mutations being selected in this investigation hints at a role in thermotolerance when utilizing the PCA-4,5 pathway. Mutations to *ptsP* either prematurely truncate the encoded protein, or cause changes to important residues based on a sequence alignment to the characterized *E. coli* PtsP (**Figure 1.7**). The selected alleles therefore

seem to be eliminating the function of PtsP in the *A. baylyi* strains. Experiments detailed later tested one of these alleles to determine its effect. Similarly, a complete *ptsP* deletion was also constructed and used to determine whether the loss of *ptsP* could increase the temperature of growth for aromatic compound degradation through the PCA 4,5-pathway.

Mutations were selected in genes predicted to encode peptidyl-prolyl isomerases (PPIases)

Eight strains independently acquired mutations in genes encoding four different predicted PPIases (**Table 2.6**). These enzymes, which are described in greater detail in chapter 1 of this thesis, are involved in protein maturation by converting proline bonds between *cis* and *trans* conformers (51). They often also contain chaperone domains to aid in the folding of their target protein (52). Changes found in PPIases in these experiments included small deletions, amino acid replacements, and large chromosomal deletions. Mutations were discovered in ACIAD_RS02420 (*tig*) and ACIAD_RS04740 (*fbp*) in ACN2041, ACIAD_RS05440 (*slyD*) in ACN2429, ACN2430, ACN2431, ACN2432, ACN2433, and ACN2434, and ACIAD_RS06505 (*ppiD*) in the amplicon replacement strain ACN2714 (**Table 2.6**). Because most of these mutated alleles caused partial or complete deletions of the PPIase-encoding gene, it seems likely that these alleles eliminate the function of each PPIase. Experiments were used to determine growth effects at 30 ° of these mutations, which affect PPIases, as detailed later.

Mutations were selected in genes encoding the GroE chaperone

Spontaneous amplifications in two evolved strains, ACN2437 and ACN2438, included ACIAD_RS12795, annotated as *groEL*, and ACIAD_RS12800, annotated as *groES*. Additionally, a mutation causing an elongation of *groES* was found in another evolved isolate,

ACN2434. *groEL* and *groES* encode the two subunits of the protein chaperone GroE. GroE, an essential chaperone, contributes to protein maturation and stability by assisting in the folding of many bacterial proteins (43, 53). More information about this chaperone is included in **Chapter 1**. Because increased gene dosage of *groEL* and *groES* was selected independently in different strains, their contribution to the growth phenotype at 30 °C most likely depends on their function and involves elevated protein expression.

The independent occurrence of mutations in the same or similar loci provides clues about genes that contribute to thermotolerance during growing through the PCA-4,5 pathway. To test whether these effects could be demonstrated, I reconstructed some mutated alleles or made deletions of specific coding sequences and tested the ability of the resulting engineered strains to grow on POB at 30 °C. Combinations of key mutations were also tested.

Table 2.2 Mutations in PCA-4,5 EASy evolved strains

Evolved strain	Gene locus ID (name)	Position*	Predicted description	Mutation type	DNA change	Protein effect
ACN2041	ACIAD_RS02085 *(ptsP)*	452,122	phosphoenolpyruvate protein phosphotransferase	deletion	Δ1 bp	Frame shift
	ACIAD_RS02420 *(tig)*	521,040 – 521-041	trigger factor	deletion	Δ2 bp	Frame shift
ACN2429	ACIAD_RS01760 *(phaAB)*	368,764	monovalent cation/H+ antiporter subunit A	Insertion	+G	Frame shift
	ACIAD_RS02030 *(rne)*	438,393	ribonuclease E (RNase E)	substitution	T→C	I345V
	ACIAD_RS02085 *(ptsP)*	452,421 – 452,459	phosphoenolpyruvate protein phosphotransferase	deletion	Δ39 bp	13 AA deletion
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP type peptidylprolyl isomerase	deletion	(GATTTCAAC)2→1	3 AA deletion
	ACIAD_RS07105	1,544,371	LysR family transcriptional regulator	substitution	C→T	R200Q
	ACIAD_RS12810 ← / → ACIAD_RS12815 (*pckG*)	2,779,459	hypothetical protein/phosphoenolpyruvate carboxykinase	deletion	(T)8→7	N/A
ACN2430	ACIAD_RS01760 *(phaAB)*	369,373	monovalent cation/H+ antiporter subunit A	insertion	+T	Frame shift
	ACIAD_RS02155 *(infA)*	466,491	translation initiation factor IF 1	substitution	T→G	E9A

	ACIAD_RS02155 (*infA*) ← / → ACIAD_RS 02160	466,561	translation initiation factor IF□1/AraC family transcriptional regulator	deletion	(T)8→7	N/A
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP□type peptidylprolyl isomerase	deletion	(GATTTCAA C)2→1	3 AA deletion
	ACIAD_RS10375	2,236,309 – 2,236,321	lipoprotein	deletion	Δ13 bp	Frame shift
	ACIAD_RS10460	2,259,196 – 2,259,201	malic enzyme	insertion	(GGTGAA)1 →2	2 AA insertion
ACN2431	ACIAD_RS05440 *(slyD)*	1,180,864	FKBP□type peptidylprolyl isomerase	substitution	C→T	E46K
ACN2432	ACIAD_RS04520 *(vanK)*	969,446 – 969,451	MFS superfamily vanillate transporter	deletion	Δ6 bp	2 AA deletion
	[Patl 3887 (LigI)]– [ACIAD_RS07925 (*pobA)*]	1,726,806	Multiple genes	deletion	Δ47,063 bp	N/A
	ACIAD_RS12280 ← / → ACIAD_RS12285	2,659,323	hypothetical protein/hypothetical protein	deletion	(T)9→8	N/A
ACN2433	ACIAD_RS01750 *(phaD)*	366,378	monovalent cation/H+ antiporter subunit D	insertion	(C)6→7	Frame shift
	ACIAD_RS02085 *(ptsP)*	452,122	phosphoenolpyruvate □protein phosphotransferase	deletion	Δ1 bp	Frame shift
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP□type peptidylprolyl isomerase	deletion	(GATTTCAA C)2→1	3 AA deletion
	ACIAD_RS07750 (*mucK)* ← / → ACIAD_R S07755 (*caiB)*	1,687,308	major facilitator superfamily cis,cis□muconate transporter/L□carnitin e dehydrogenase	substitution	G→T	N/A
	ACIAD_RS14555 *(rpmC)*	3,130,364 – 3,130,378	50S ribosomal protein L29	insertion	(GGTTTTAC GAGCAAT)1 →2	5 AA deletion
ACN2434	ACIAD_RS02085 *(ptsP)*	451,463	phosphoenolpyruvate □protein phosphotransferase	insertion	+A	Frame shift
	ACIAD_RS02190 *(gltP)*	474,662	glutamate/aspartate:pr oton symporter	substitution	C→T	Q124*
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP□type peptidylprolyl isomerase	deletion	(GATTTCAA C)2→1	3 AA deletion
	ACIAD_RS07750 *(muck)*	1,686,451	major facilitator superfamily cis,cis□muconate transporter	substitution	T→C	Y243C
	ACIAD_RS11075 *(ppsA)*	2,389,497	phosphoenolpyruvate synthase	substitution	G→A	R515H
	ACIAD_RS12800 *(groES)*	2,776,999	GroES	substitution	A→G	*97Q
	ACIAD_RS14725 *(suhB)*	3,161,940 – 3,162,011	inositol monophosphatase	deletion	Δ72 bp	24 AA deletion

	ACIAD_RS15195 *(comM)*	3,266,721 – 3,266,804	membrane protein (ComM)	deletion	Δ84 bp	28 AA deletion
ACN2435	ACIAD_RS06285 *(clpA)*	1,359,837	ATP□binding protease component	deletion	Δ1 bp	Frame shift
	Patl_3889 *(ligJ)*	N/A	2-keto-4-carboxy-3-hexenedioate hydratase	substitution	A→G	D311G
	ACIAD_RS10375	2,236,391 – 2,236,421	lipoprotein	insertion	(GTGCCTGC ATCATCTT TAAATAGC GATAATG)1 →2	Frame shift
ACN2436	ACIADtRNAThr_7	296,778	tRNA□Thr	deletion	Δ1 bp	Frame shift
	ACIAD_RS02155 *(infA)* ← / → ACIAD_RS02160	466,561	translation initiation factor IF□1/AraC family transcriptional regulator	deletion	(T)8→7	N/A
	ACIAD_RS11045 *(hemK)*	2,381,348	methyl transferase	insertion	+C	Frame shift
	ACIAD_RS12815 *(pckG)*	2,780,282	phosphoenolpyruvate carboxykinase	substitution	A→G	E238G
	ACIAD_RS12985 *(sucA)*	2,817,587	2□oxoglutarate dehydrogenase E1	substitution	G→A	A914V
ACN2437	ACIAD_RS01760 *(phaAB)*	369,749	monovalent cation/H+ antiporter subunit A	substitution	C→T	G175D
	ACIAD_RS02085 *(ptsP)*	451,482 – 451,496	phosphoenolpyruvate □protein phosphotransferase	deletion	Δ15 bp	5 AA deletion
	ACIAD_RS02155 *(infA)*	466,342	translation initiation factor IF□1	substitution	T→G	T59P
	ACIAD_RS03845 *(wax□dgaT)*	820,521	bifunctional wax ester synthase/acyl□CoA; diacylglycerol acyltransferase	insertion	(TTCATC)2 →3	2 AA insertion
	ACIAD_RS15195 *(comM)*	3,267,325	membrane protein (ComM)	deletion	(T)7→6	Frame shift
ACN2438	ACIAD_RS01740 *(phaF)*	364,659	pH adaptation potassium efflux system protein F	insertion	(AAACAT)1 →2	2 AA insertion
	ACIAD_RS02090 (*rppH*) ← / → ACIAD_RS02095	454,209	dinucleoside polyphosphate hydrolase/haloacid dehalogenase	substitution	T→C	N/A
	ACIADtRNASer_34 → / ← ACIAD_RS05325	1,148,146	tRNA□Ser/hypothetical protein	substitution	C→T	N/A
	ACIAD_RS10375	2,236,529	lipoprotein	insertion	+A	Frame shift

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Table 2.3 Junctions and IS*1236* insertions in EASy evolved strains and amplicon replacement strain ACN2714

Evolved strain	Predicted description	Frequency*	Junction/ location of IS*1236* insertion**
ACN2041	New amplification junction including *lig* genes: ΩK^R and ACIAD_RS07905 (*quiX)*	98%	GCGAAACGATCCTCAGATAGGTATAATTCA
	90 kbp deletion: ACIAD_RS04420 and ACIAD_RS04820 (*metH*)	99%	ATTTTGGCTAGGATCAGTCCTTACGAATCA
ACN2429	SBF: ACIAD_RS07910 (*quiA*) and ACIAD_RS07915 (*pobS*)	93%	TTTACCTGTCACACGGGAATCGGAGTGGAC
	IS*1236* insertion: IS*1236* and intergenic region between ACIAD_RS10465 (*cca*) and ACIAD_RS10470 (*cydA*)	95/94%	2,261,252
ACN2430	SBF: ACIAD_RS07910 (*quiA*) and ACIAD_RS07915 (*pobS*)	95%	TTTACCTGTCACACGGGAATCGGAGTGGAC
ACN2431	SBF: ACIAD_RS07910 (*quiA*) and ACIAD_RS07915 (*pobS*)	94%	TTTACCTGTCACACGGGAATCGGAGTGGAC
ACN2432	IS*1236* insertion: IS*1236* and intergenic region between ACIAD_RS12710 and ACIAD_RS12715 (*citN*)	94/94%	2,755,247
ACN2433	SBF: ACIAD_RS07700 and ACIAD_RS07705	94%	GAATTAATTTGGCTCTCAAATGTGCTTGTT
ACN2434	SBF: ACIAD_RS07700 and ACIAD_RS07705	95%	GAATTAATTTGGCTCTCAAATGTGCTTGTT
ACN2435	SBF: ACIAD_RS07700 and ACIAD_RS07705	90%	GAATTAATTTGGCTCTCAAATGTGCTTGTT
	IS*1236* insertion: IS*1236* and ACIAD_RS06490	97/99% ***	1,400,282
	IS*1236* insertion: IS*1236* and intergenic region between ACIAD_RS08590 and ACIAD_RS08595	100/100%	1,860,413
	IS*1236* insertion: IS*1236* and intergenic region between ACIAD_RS09195 (*dadA*) and ACIAD_RS09200 (*ispD*)	97/99%	1,989,989
ACN2436	SBF: ACIAD_RS07910 (*quiA*) and ACIAD_RS07915 (*pobS*)	97%	TTTACCTGTCACACGGGAATCGGAGTGGAC
ACN2437	SBF: ACIAD_RS07910 (*quiA*) and ACIAD_RS07915 (*pobS*)	92%	TTTACCTGTCACACGGGAATCGGAGTGGAC
	Spontaneous amplification junction: ACIAD_RS12795 (*groEL*) and intergenic region between ACIAD_RS12800 (*groES*) and ACIAD_RS12805	48%	CTGTTGGTATCAACAAATATAGACTTAGTC
ACN2438	SBF: ACIAD_RS07910 (*quiA*) and ACIAD_RS07915 (*pobS*)	92%	TTTACCTGTCACACGGGAATCGGAGTGGAC
	Spontaneous amplification junction:	85%	CTGGCGTGCCACGCTTTATCGTTTATACAA

	ACIAD_RS12785 and ACIAD_RS12805 (including *groEL* and *groES*)		
ANC2714	IS*1236* insertion: IS*1236* and with deletion from ACIAD_RS06490 to ACIAD_RS06510 (*alkR*) (includes deletion of ACIAD_RS06505 (*ppiD*)	94/88%	1,400,282 – 1,403,648
	IS*1236* insertion: IS*1236* and intergenic region between ACIAD_RS08590 and ACIAD_RS08595	88/89%	1,860,413
	IS*1236* insertion: IS*1236* and intergenic region between ACIAD_RS09195 (*dadA*) and ACIAD_RS09200 (*ispD*)	89/92%	1,989,989

* Frequency refers to the percent of reads in which the junction is present. For example, if there is a chromosomal duplication, the frequency would be 50%. Insertions have two frequencies reported, one for the upstream region on one for the downstream region of the insertion.

** Numbers correspond to positions on the ADP1 chromosome in NCBI entry NC_005966.

Table: 2.4 Mutations in amplicon replacements strains

Amplicon Replacement strain	Gene locus ID (name)	Position*	Predicted description	Mutation type	DNA change	Protein effect
ACN2711	ACIAD_RS00015 (*recF*) → / → ACIAD_RS00020 (*gyrB*)	4087	recombination protein F/DNA gyrase, subunit B (type II topoisomerase)	substitution	A→C	N/A
	ACIAD_RS01435 (*nusG*)	299,455	transcription antitermination protein	substitution	G→A	G89D
	ACIAD_RS01760 *(phaAB)* – ACIAD_RS01765	370,006 – 370,826	monovalent cation/H+ antiporter subunit A,	deletion	Δ821 bp	deletion
	ACIAD_RS02085 *(ptsP)*	452,510	phosphoenolpyruvate □protein phosphotransferase	substitution	C→T	E340K
	ACIAD_RS03080 → / → ACIAD_RS03085	661,260	hypothetical protein/peptidase	insertion	(A)7→8	N/A
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP□type peptidylprolyl isomerase	deletion	(GATTTCAAC)2→1	Small deletion
	ACIAD_RS06135	1,332,722	hypothetical protein	substitution	C→A	P183Q
	ACIAD_RS14555 (*rpmC*)	3,130,381	50S ribosomal protein L29	deletion	Δ1 bp	Frame shift
ACN2712	ACIAD_RS02085 *(ptsP)*	451,463	phosphoenolpyruvate □protein phosphotransferase	insertion	+A	Frame shift

	ACIAD_RS02190 *(gltP)*	474,662	glutamate/aspartate:proton symporter	substitution	C→T	Q124*
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP□type peptidylprolyl isomerase	deletion	(GATTTCAAC)2→1	3 AA deletion
	ACIAD_RS07750 *(muck)*	1,686,451	major facilitator superfamily cis,cis□muconate transporter	substitution	T→C	Y243C
	ACIAD_RS11075 *(ppsA)*	2,389,497	phosphoenolpyruvate synthase	substitution	G→A	R515H
	ACIAD_RS12750	2,765,010	periplasmic binding protein of transport/transglycosylase	substitution	A→C	S353R
	ACIAD_RS12800 *(groES)*	2,776,999	GroES	substitution	A→G	*97Q
	ACIAD_RS14725 *(suhB)*	3,161,940 – 3,162,011	inositol monophosphatase	deletion	Δ72 bp	24 AA deletion
	ACIAD_RS15195 *(comM)*	3,266,721 – 3,266,804	membrane protein (ComM)	deletion	Δ84 bp	28 AA deletion
ACN2713	ACIAD_RS02090 (*rppH*) ← / → ACIAD_RS02095	454,209	dinucleoside polyphosphate hydrolase/haloacid dehalogenase	substitution	T→C	N/A
	ACIAD_RS10375	2,236,529	lipoprotein	insertion	+A	Frame shift
ACN2714	ACIAD_RS06285 *(clpA)*	1,359,837	ATP□binding protease component	deletion	Δ1 bp	Frame shift
	ACIAD_RS10375	2,236,391 – 2,236,421	lipoprotein	insertion	(GTGCCTGCATCATCTTTAAATAGCGATAATG)1→2	Frame shift
ACN2718	ACIAD_RS02085 *(ptsP)*	452,122	phosphoenolpyruvate□protein phosphotransferase	deletion	Δ1 bp	Frame shift
	ACIAD_RS05440 *(slyD)*	1,180,592 – 1,180,600	FKBP□type peptidylprolyl isomerase	deletion	(GATTTCAAC)2→1	3 AA deletion
	ACIAD_RS07750 (*mucK)* ← / → ACIAD_RS07755 (*caiB)*	1,687,308	major facilitator superfamily cis,cis□muconate transporter/L□carnitine dehydrogenase	substitution	G→T	N/A
	ACIAD_RS11230 → / ← ACIAD_RS11235 *(glnA)*	2,420,278	hypothetical protein/glutamine synthetase	insertion	+C	N/A
	ACIAD_RS14555 *(rpmC)*	3,130,364 – 3,130,378	50S ribosomal protein L29	insertion	(GGTTTTACGAGCAAT)1→2	5 AA insertion

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Table 2.5: Summary of *ptsP* mutations

Evolved Strain	Change	Position*	Effect
ACN2140	Δ1 bp	coding (1405/2295 nt)	C deletion in codon 469 causing a frameshift and a premature truncation after AA498 Abolishes binding sites, metal binding sites, and C680 active site
ACN2429	Δ39 bp	coding (1069-1107/2295 nt)	In frame deletion of 13 AA after Ala 354 Abolishes active site H367
ACN2432/ ACN2711	C→T E340K (GAA→AAA)	452,510	AA change Change in conserved residue
ACN2433/ ACN2718	Δ1 bp	coding (1405/2295 nt)	C deletion in codon 469 causing a frameshift and a premature truncation after AA498 Abolishes binding sites, metal binding sites, and C680 active site
ACN2434/ ACN2712	+A	coding (2063/2295 nt)	A insertion in codon 689 causing a frameshift and a premature truncation after AA706 Frameshift after binding/active sites
ACN2437	Δ15 bp	coding (2032□2046/2295 nt)	5 AA in frame deletion after V677 including S678-I679-C680-G681-E682 Abolishes active site C680
ACN2438/ ACN2713	T→C	intergenic (□111/□74)	Mutation in promoter of rppH/ptsP shifts prediction, probably affecting expression levels

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Table 2.6: Summary of peptidyl-prolyl isomerase mutations

Evolved Strain	Gene	Change	Position*	Effect
ACN2140	ACIAD_RS02420 (*tig*)	Δ2 bp	coding (986□987/1335 nt)	2 bp deletion AA328-329 causing frameshift and premature truncation after AA333
ACN2140	ACIAD_RS04740 (*fbp*)	Deletion (in 90 kbp)	N/A	N/A
ACN2429	ACIAD_RS05440(*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	coding (400□408/483 nt)	Small deletion
ACN2430	ACIAD_RS05440(*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	coding (400□408/483 nt)	Small deletion
ACN2431	ACIAD_RS05440(*slyD*)	C→T	E46K	AA replacement
ACN2432/ ACN2711	ACIAD_RS05440(*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	coding (400□408/483 nt)	Small deletion
ACN2433/ ACN2718	ACIAD_RS05440(*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	coding (400□408/483 nt)	Small deletion
ACN2434/ ACN2712	ACIAD_RS05440(*slyD*)	$(GATTTCAAC)_{2\rightarrow1}$	coding (400□408/483 nt)	Small deletion
ACN2714	ACIAD_RS06505 (*ppiD*)	Deletion	N/A	N/A

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Mutations conferring thermotolerant growth through the PCA-4,5 pathway

To test if mutations accrued during EASy evolution contribute to *A. baylyi* strains' ability to grow through the PCA-4,5 pathway at 30 °C, mutated alleles *ptsP52041* and *tig52041*, that arose from ACN2041, were used to replace the corresponding wild-type gene in one of the parent strains with a single copy of the *lig* genes, ACN2120. This strain was chosen since it matched the location and orientation of the *lig* genes in the strain these alleles were discovered, ACN2041. This process of allelic replacement will be termed "reconstruction" to indicate that I am engineering a specific genetic change that was discovered in my experiments into a different genetic background. These two alleles (*ptsP52041* and *tig52041)* were chosen because they represent the two most common type of changes found in strains selected for high-temperature growth. Since the alleles, affecting PtsP and a PPIase, cause premature truncations that should eliminate protein function, their effects may be more easily interpreted than other types of changes such as amino acid replacements. I also made deletions of *ptsP* and *tig* in ACN2120. Such deletions test if the loss of PtsP and trigger factor (encoded by *tig*) allow *A. baylyi* strains to grow on aromatics through the PCA-4,5 pathway at 30 °C.

I generated strains with different *ptsP* and *tig* alleles by allelic replacement of the corresponding gene of ACN2120. Next, I challenged each strain to grow on POB at 30 °C and measured the optical density at 595 nm (OD_{595}) at 72 hours post inoculation. As shown in **Figure 2.8,** *ptsP52041* or *tig52041* alone did not allow *A. baylyi* to grow through the PCA-4,5 pathway on POB at 30 °C. Additionally, deleting either coding sequence did not confer the growth

phenotype either. However, because *ptsP52041* and *tig52041* arose in the same strain, I decided to try reconstructing these alleles in combination. Strains containing both *ptsP52041* and *tig52041* or Δ*ptsP* and Δ*tig* were able to grow through the *lig* genes on POB at 30 °C (**Figure 2.8**). This result indicates that removing both PtsP and trigger factor allows *A. baylyi* to grow through the PCA-4,5 pathway at increased temperatures up to 30 °C.

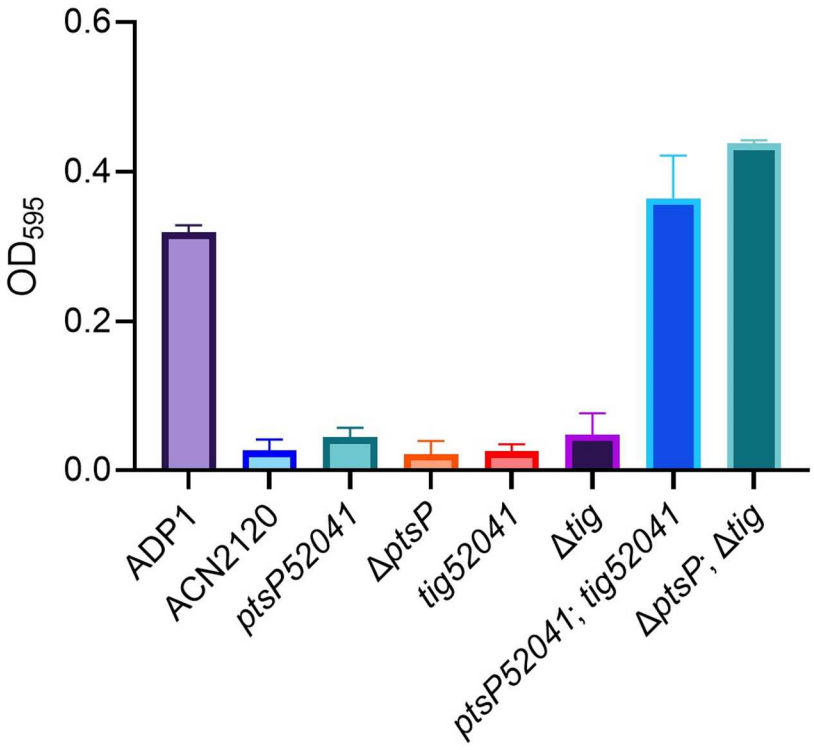


Figure 2.8: Growth on POB at 30 °C of strains with Tig and/or PtsP truncations or deletions using the PCA-4,5 pathway

Growth as measured by OD_{595} taken at 72 hours of growth on POB at 30 °C. ADP1 is the wild-type strain, and ACN2120 is the parent strain containing the *lig* genes in which all subsequent *ptsP* and *tig* alleles or deletions were reconstructed. Reconstructed strains are ACN2770 (*ptsP52041*), ACN2816 (Δ*ptsP*), ACN2772 (*tig52041*), ACN2828 (Δ*tig*), ACN2836 (*ptsP52041*;

tig52041), and ACN2838 ((Δ*ptsP;* Δ*tig*). All strains were done in biological duplicate. Error bars represent standard deviation.

Contribution of 90 kbp deletion to thermotolerant growth

Both *ptsP52041* and *tig52041* arose in the *A. baylyi* strain with the large 90 kbp deletion, ACN2041. This 90 kbp deletion contains 81 predicted coding sequences including the genes for the degradation of the aromatic compound vanillate, consumed through PCA. The vanillate transporter VanK, referenced earlier as an additional PCA and POB transporter, is also encoded in the large deletion and therefore not present in ACN2041. All genes in the 90 kbp deletion are listed in **Table 2.11**. Because there are so many genes missing in this strain, I sought to test the *ptsP* and *tig* alleles and deletions in this genetic background to see if there are differences in the phenotype compared to when they are reconstructed in a strain without this large deletion (such as ACN2120, above).

Figure 2.9 shows the results of introducing *ptsP52041*, *tig52041,* Δ*ptsP,* and Δ*tig* into ACN1910 (the parent strain of ACN2041 that contains the *lig* genes and the 90 kbp deletion). Once again, *tig52041* or Δ*tig* alone did not allow growth through the PCA-4,5 pathway at 30 °C. However, unlike in ACN2120 (which does not carry the 90 kbp deletion), *ptsP52041* and Δ*ptsP* alone allowed strains to grow on POB at 30 °C. This result suggested that the loss of some or all of the DNA in this 90 kbp deletion, contributes to the ΔPtsP-dependent effect on increased thermotolerant growth. Combining the *tig52041* or Δ*tig* with the *ptsP* allele or deletion did not confer any additional benefit (**Figure 2.9**).

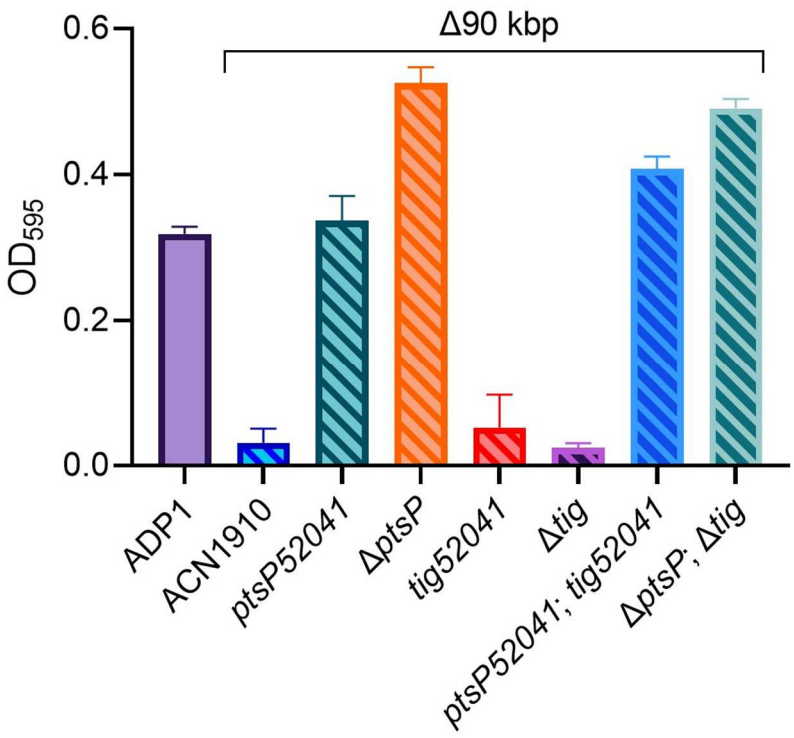


Figure 2.9: Growth on POB at 30 °C of strains with 90 kbp deletion and Tig/PtsP truncations or deletions using the PCA-4,5 pathway

Growth as measured by OD_{595} taken at 72 hours of growth on POB at 30 °C. Diagonal line pattern bars indicate that strains carry the 90 kbp deletion. ADP1 is the wild-type strain, and ACN1910 is the parent strain containing the *lig* genes and 90 kbp deletion in which all subsequent *ptsP* and *tig* alleles or deletions were reconstructed. Reconstructed strains are ACN2769 (*ptsP52041*), ACN2815 (Δ*ptsP*), ACN2771 (*tig52041*), ACN2827 (Δ*tig*), ACN2835 (*ptsP52041*; *tig52041*), and ACN2837 (Δ*ptsP*; Δ*tig*). All strains were done in biological duplicate. Error bars represent standard deviation.

After identifying all genes in the 90 kbp deletion, one gene stood out because it encodes a predicted FKBP-type PPIase. This gene is ACIAD_RS04740, designated *fbp*. In previous strains, the deletion of *ptsP* together with the deletion of *tig*, encoding a different FKBP-type PPIase,

was sufficient for growth. Based on this observation, I postulated that the deletion of *fbp* was conferring a similar benefit in strains with the 90 kbp deletion. I therefore reconstructed a deletion of *fbp* with the *ptsP* and *tig* alleles and deletions, using ACN2120 as the parent strain. Indeed, when *fbp* was deleted in combination with *ptsP* in strains without the 90 kbp deletion, they gained the ability to grow through the *lig* genes at temperatures up to 30 °C. No additional benefit was found by deleting or truncating both FKBP type PPIases in combination with *ptsP* (**Figure 2.10**).

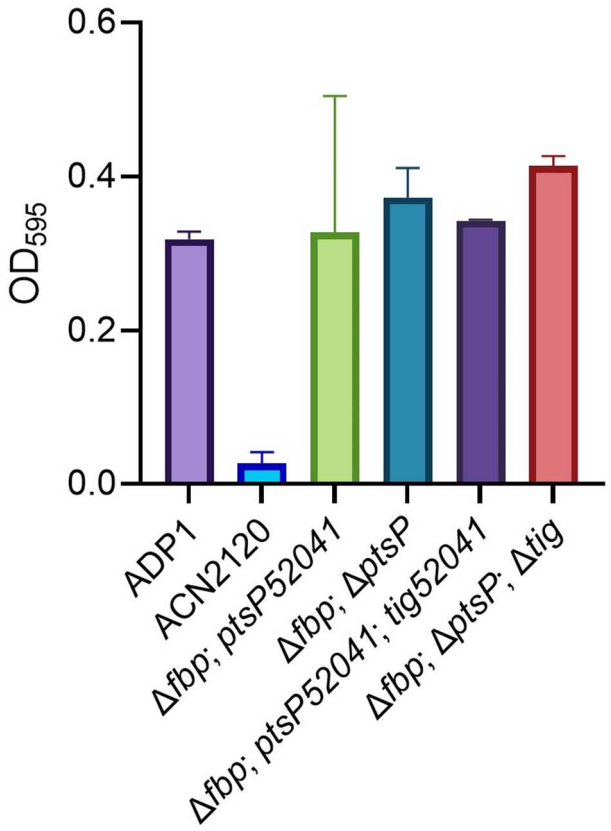


Figure 2.10: Growth of strains with *fbp* deletion in combination with *ptsP* on POB at 30 °C

Growth as measured by OD_{595} taken at 72 hours of growth on POB at 30 °C. ADP1 is the wild-type strain, and ACN2120 is the parent strain containing the *lig* genes in which all subsequent *ptsP* and *tig* alleles or deletions were reconstructed. Reconstructed strains are ACN2904 (Δ*fbp*; *ptsP52041*), ACN2905 (Δ*fbp*; Δ*ptsP*), ACN2906 (Δ*fbp*; *ptsP52041*; *tig52041*), and ACN2907

(Δ*fbp*; Δ*ptsP*; Δ*tig*). All strains were done in biological duplicate. Error bars represent standard deviation.

High expression of *groEL* and *groES* is sufficient for growth through the foreign pathway at 30 °C

The next set of strain constructions involved studies of the role played by the GroE chaperone. In strains selected by the EASy method, independent spontaneous chromosomal amplifications were discovered in which the gene dosage of *groEL* and *groES* increased. To test the significance of this observation, I sought to increase the expression of these genes by altering the relevant promoter. I tried to determine if higher expression of these genes contributes to the strain's ability to grow on aromatic compounds via the PCA-4,5 pathway at 30 °C. The constitutive *trc* promoter was put upstream *groES* because previous experiments determined that the *trc* promoter gives high expression in *A. baylyi* (12). I have yet to compare expression from the native *groES* promoter to the *trc* promoter, so assumptions about expression from *trc* being higher is still a hypothesis. Because of their related protein function and the relatively small intergenic region between *groES* and *groEL* (60 base pairs0, I expect these genes to be co-transcribed, so putting the *trc* promoter upstream *groES* should affect the expression of both genes.

Strains with the *trc* promoter upstream *groES* were, in fact, able to grow through the foreign PCA-4,5 pathway at 30 °C (green bars; **Figure 2.11**). Combining *ptsP*, *tig*, or *fbp* evolved alleles or deletions with increased expression of *groES* and *groEL* did not further improve the thermotolerance (orange bars; **Figure 2.11**). Lastly, the *groES52434* allele, which

extends the *groES* coding sequence by an additional 13 amino acids, did not appear to confer the thermotolerance phenotype (pink bars; **Figure 2.11**).

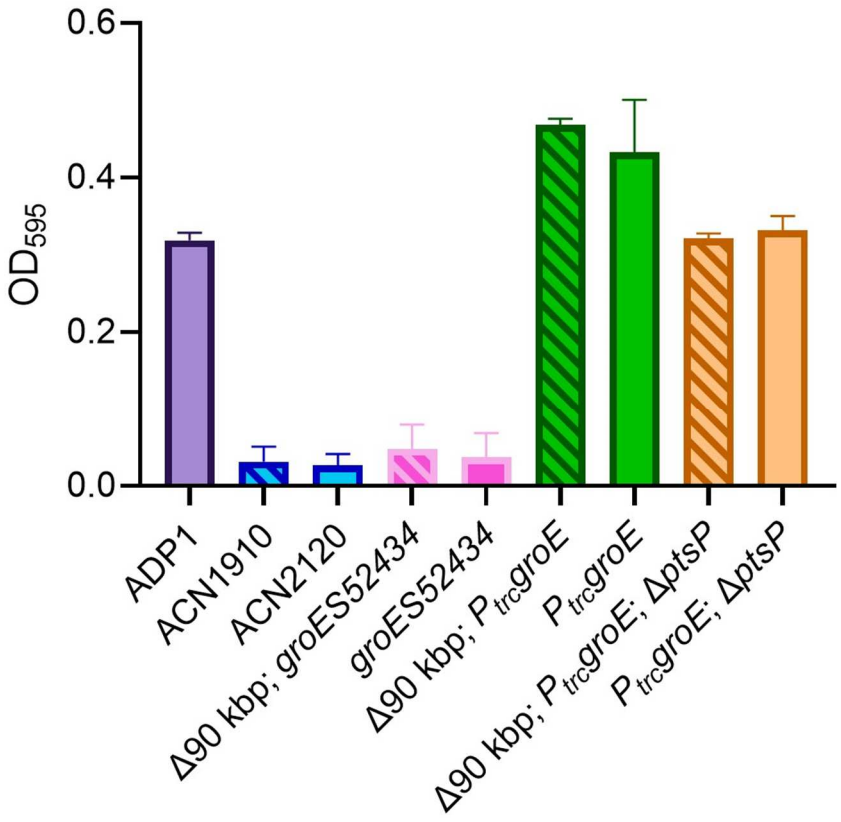


Figure 2.11: OD_{595} of strains overexpression *groES* and *groEL* on POB at 30 °C

Growth as measured by OD_{595} taken at 72 hours of growth on POB at 30 °C. Diagonal line pattern bars indicate that strains carry the 90 kbp deletion. Reconstructed strains are ACN2864 (Δ90 kbp; *groE52434*), ACN2866 (*groE52434*), ACN2865 (Δ90 kbp; P_{trc}*groE*), ACN2867 (P_{trc}*groE*), ACN2875 (Δ90 kbp; P_{trc}*groE*; Δ*ptsP*), and ACN2876 (P_{trc}*groE*; Δ*ptsP*). All strains done in biological duplicate. Error bars represent standard deviation.

Discussion

PCA-4,5 pathway from *Pseudoalteromonas atlantica* T6c is functionally expressed from the *Acinetobacter baylyi* chromosome

In *A. baylyi* strains missing the native PCA 3,4-cleavage pathway, chromosomal integration of a predicted PCA 4,5-cleavage pathway from *Pseudoalteromonas atlantica* T6c allowed strains to grow on PCA and aromatics funneled through PCA, such as POB, as sole carbon sources. *A. baylyi*-derived strains were able to express and utilize the large 7.3 kbp segment of foreign DNA containing seven catabolic genes and a transcriptional regulator, referred to as the *lig* genes, in various regions and orientations in the chromosome (**Figure 2.2** and **Table 2.1**). However, strains could only grow through the foreign pathway at temperatures up to 25 °C, and *A. baylyi* is typically grown at 30 or 37 °C (12, 13). Such temperature dependence could be resulting from a variety of problems such as temperature-dependent protein stability, enzyme activity, or transcriptional regulation. *P. atlantica* T6c itself is typically grown at 25 or 30 °C (54-58). The lower growth temperature of the marine bacterium, however, does not imply that there is anything inherent in the sequences of the Lig proteins that would affect thermal tolerance when heterologously expressed in *A. baylyi*.

No part of the *P. atlantica* PCA-4,5 pathway sequence was altered for integration into the *A. baylyi* chromosome. Genes were not codon optimized to match *A. baylyi* codon usage, nor were the native promoters or ribosome binding sites altered. Therefore, *A. baylyi* was able to recognize and express genes from the *P. atlantica* promoter(s) and coding sequences. Because the *P. atlantica* PCA-4,5 pathway has only been predicted and not characterized, the catabolic *lig* genes may be transcribed as a single operon, or from various promoters throughout the sequence. It is not known whether the transcriptional regulator from *P. atlantica* is required for growth.

Promoters and other regulatory elements are not always compatible between bacteria. For example, *A. baylyi* does not transcribe genes from the *E. coli lac* promoter used in many synthetic constructs (12). Therefore, chromosomal expression sufficient for growth through the foreign *lig* genes in *A. baylyi* could not have been predicted. Yet success of this approach bodes well for future studies and pathway engineering. For biotechnological purposes, the localization of foreign genes in the chromosome could be advantageous compared to placing them on plasmids, which can be lost. Plasmid-based expression systems can be problematic because they need to be maintained by selective pressure (often antibiotic resistance) and there can be concern about plasmid mobility with such recombinant constructs.

Manipulating the expression of complex catabolic pathways for aromatic compound degradation can be especially difficult because the intermediary metabolites are often toxic and all steps in a pathway must be properly balanced. Genes for aromatic compound catabolism are often very tightly regulated (3). It is possible that the initial success of achieving growth through the foreign pathway, at any temperature, was because of, and not in spite of, making no attempt to express the pathway at high levels where imbalances in regulation could have been toxic. Nevertheless, in trying to optimize expression for the intended long-term application, some manipulation was needed to increase the temperature at which *A. baylyi* could grow while using the foreign *lig* genes.

Without knowing the exact issue preventing *A. baylyi* cells from growing through the *lig* genes at temperatures above 25 °C, it was difficult to predict what targeted genetic changes could be made to allow growth. Therefore, I sought to overcome complex growth issues by employing the EASy method of laboratory evolution. This method allows dynamic changes in chromosomal gene dosage, thereby enabling natural selection rather than rational design to determine the

optimal expression level necessary for growth under specific conditions. The EASy method was developed, in part, to alter enzyme substrate specificity by increasing the copies of an enzyme with weak activity to improve catalysis. However, this method may be even more important when the potential problems are complex and unknown. Issues associated with thermotolerance likely involve protein instability, regulation, and problems with integrating the foreign pathway into the host metabolic networks. Therefore, increasing gene dosage could help by providing more protein to compensate for instability, protein degradation, improper folding, and/or insufficient regulation by the host's transcriptional and post-transcriptional mechanisms. As described in this thesis, I successfully generated strains that grow at 37 °C using the *lig* genes.

Increased chromosomal gene dosage of foreign PCA-4,5 pathway allowed growth at higher temperatures

Eleven *A. baylyi* populations that successfully grew on PCA or POB after EASy chromosomal amplification of the *lig* genes underwent long-term culturing for 16 weeks. This time period corresponds to approximately 1800 generations. During the sixteen weeks of culturing under selective pressure, growth on PCA or POB through the PCA-4,5 pathway at 37 °C, all populations greatly reduced the number of copies of the *lig* genes in their chromosomes. Two populations that started out with more than 150 copies of F87 only had 4 copies after 16 weeks. Such a great change in required copies of the *lig* genes led me to predict that mutations accrued that contribute to my strains' selected phenotype. Because I was selecting for thermotolerance, I suspected that strains would need to evolve mutations allowing the cells to incorporate the foreign pathway into *A. baylyi's* larger metabolic or regulatory networks.

Only one mutation was discovered in the foreign DNA. This *lig*-gene mutation encoded a D311G replacement in Patl_3889 (LigJ), the 2-keto-4-carboxy-3-hexenedioate hydratase. Initial investigation into this mutation has not revealed any benefit conferred by this change (data not shown). However, in other strains, mutations could have arisen in some amplicons of the *lig* genes that were missed by Illumina WGS. Illumina sequencing uses short read lengths, typically 50-150 bp (59). This method would only allow us to see mutations in the *lig* genes if they occurred in all or most of the amplicons since all reads from the amplicons would map to the same sequence in the reference. For example, if a mutation occurred in only 1 of the 10 copies of one of the *lig* genes, it would only appear in 10% of the reads. This low-level detection would not be interpreted as a mutation and would most likely be dismissed as a sequencing error. PacBio or nanopore sequencing can read many thousand bp in a single read (60, 61). Either of these methods of whole genome sequencing might help detect beneficial mutations that arise in a subset of amplicons.

Mutations that were selected during long-term culturing outside of the foreign *lig* genes commonly were in genes encoding proteins involved in post-translational modifications or regulation. These types of mutations suggest that the ability of *A. baylyi* strains to grow at higher temperatures is dependent on processes typically associated with thermotolerance, such as protein stability and maturation. For instance, a couple of strains evolved spontaneous amplifications including *groES* and *groEL*, genes encoding the subunits of the GroE chaperone. Increasing gene dosage of genes encoding the GroE chaperone has been shown to aid in stress tolerance in many other studies and in different bacteria (62). In my studies, when these genes were put under the control of a presumably high-expression synthetic promoter in *A. baylyi*, they allowed strains to grow through the foreign pathway at increased temperatures, up to 30 °C

(**Figure 2.11**). These mutations to post-translational modification systems are consistent with my finding that *A. baylyi* populations maintained multiple copies of the *lig* genes even after 16 weeks of culturing. The multiple copies of the *lig* genes may be aiding in issues of protein stability and/or folding by producing more protein than they would in single chromosomal copy.

Another common gene in which mutations were discovered was *ptsP*. This is the first discovery in an *Acinetobacter* strain of any role played by this gene. As discussed below, a comparison of PtsP studies in other bacteria suggest that an important player in complex metabolic regulation has been uncovered and warrants further investigation. Mutations discovered in *A. baylyi* strains in which foreign genes were evolved by the EASy method appear to be shedding light on how foreign genetic pathways are integrated into new hosts.

A novel regulatory circuit in *A. baylyi* revealed by *ptsP* mutations

Among the most intriguing mutations uncovered in my studies were several that affect ACIAD_RS02085, annotated as *ptsP*. As shown in (**Table 2.5**), six separate mutations were discovered in seven independently evolved isolates that appear to eliminate the function of an encoded protein (uniprot ID code Q6FEW8). This ADP1 protein is a homolog of the first enzyme in a phosphorelay system studied in other organisms (63). The significance of *ptsP* for the selected phenotype in my research was confirmed experimentally by demonstrating that deletion of this gene in combination with deletion of *tig* or *fbp*, which each encode a PPIase, was sufficient for a 5 °C increase in the maximal temperature for growth of *A. baylyi* strains utilizing a foreign catabolic pathway for aromatic compound consumption (**Figures 2.8-2.10**).

While PtsP has not been studied in ADP1 or any other *Acinetobacter* species (to the best of our knowledge), there are many recent reports that the PTS system in which it is predicted to

participate, PTS^{Ntr}, is critical for the global control, and integration, of cellular signals that balance carbon and nitrogen availability (64). As detailed in **Chapter 1** of this thesis, there are two general types of PTS systems, a carbohydrate-PTS that regulates sugar transport (63, 65, 66), and a three-component signal transduction pathway with diverse regulatory roles in different bacteria (63). This latter system, PTS^{Ntr}, is named in this fashion because it is often involved in nitrogen metabolism (67). As part of the connection to nitrogen metabolism, one component of the typical PTS^{Ntr}, PtsN (EIIANtr), is usually encoded by a gene localized together with *rpoN*. RpoN is a nitrogen-responsive RNA polymerase sigma factor (σ^{54}). However, in ADP1 there is no homolog of PtsN, raising questions of how a signal transduction system that involves PtsP in this bacterium compares with those that use PtsN for coordinating nitrogen metabolism.

One characteristic of PtsP that implicates it as a PTS^{Ntr} component rather than a component of a classical carbohydrate PTS, is its N-terminal GAF domain (discussed in **Chapter 1**). This domain is also found in NifA, which interacts with RpoN to serve as the key transcriptional activator of nitrogen fixation genes in diazotrophic bacteria (68). This domain creates another potential link to nitrogen sensing and takes its name from proteins where it was initially found (c**G**MP-specific phosphodiesterases, **a**denylyl cyclases and **F**hlA, a formate hydrogen lyase transcriptional activator). The GAF domain serves as a sensor of small molecules, and it plays a role in PtsP to modulate the activity of the adjacent EI domain, which catalyzes the transfer of a phosphoryl group from PEP through a phospho-histidine intermediate. In *E. coli*, the GAF domain of PtsP senses, and binds to, glutamine and α-ketoluglutarate as part of the global nitrogen-metabolism regulatory circuitry. GAF domains can also transduce signals from cyclic nucleotides (cAMP or cGMP), which are common second messengers in the global

regulation of carbon metabolism (50, 69). In some bacteria cross-talk can occur between sugar PTS and PTS^{Ntr} to effect an intricate balance of carbon and nitrogen pools (64).

For canonical PTS, the phosphoryl group from PEP is ultimately used for the coupled phosphorylation and translocation of a sugar across the cytoplasmic membrane. In ADP1, there is one only sugar PTS, and it is functional for fructose uptake despite the inability of ADP1 to use fructose as a carbon source (described in **Chapter 1**). The fate of a phosphoryl group that might be transferred from PEP by PtsP in ADP1 is completely unknown. However, studies in *Legionella pneumophila* provide a framework for speculation (70). In *L. pneumophila*, as in ADP1, there is no PtsN homolog, but a two-component PtsP-PtsO system was shown to be functional and to participate in complex global regulation.

As in *L. pneumophila*, we hypothesize that PtsP, and a putative PtsO, connect with another common type of bacterial phosphotransfer cascade, two-component transcriptional regulatory systems. It is possible that PtsP is connected directly or indirectly to a GacA-GacS two component system in ADP1. There are several reasons for this supposition. The presence of *ptsP* is highly conserved in the *Acinetobacter* genus, and in acineetobacters as in many other bacteria, *ptsP* is chromosomally localized with an adjacent gene that is highly conserved. The gene product of this conserved neighbor, ACIAD_RS02090, is homologous to the mRNA processing enzyme RppH. This enzyme de-phosphorylates the 5' end of mRNA, targeting transcripts for degradation (71).

Studies in diverse bacteria support the possibility of interplay between RppH (and maybe PtsP) and CsrA, a carbon storage regulator that another member of the Neidle lab, Emily McIntyre, has been studying in the context of aromatic compound catabolism in *A. baylyi* (unpublished data). In many EASy studies of aromatic compound degradation, both published

and unpublished research, mutations in *gacA* and *csrA* have been discovered frequently. Our research supports a model in which the *gacA* and *gacS* transcriptional regulatory system controls transcription of multiple small noncoding RNAs. In turn these, RNA molecules sequester the CsrA protein. CsrA acts as a global regulator for the post-transcriptional control of aromatic compound catabolism in ADP1. The possible link to PtsP is that RppH has been shown to affect CsrA-mediated regulation through the degradation of the non-coding RNAs that control CsrA availability in a related system in *Azotobacter vinlandii* (72). Moreover, the studies of PtsP in *L. pneumophila* also connect indirectly with a CsrA homolog (70).

The complex scenario outlined above is the starting point for a regulatory model that generates testable hypotheses. Experimental exploration of this model will be done in the future. Regardless of its veracity, my experiments demonstrate that PtsP is important in the acquisition of higher growth temperatures after chromosomal integration of a foreign pathway for aromatic compound catabolism. It is likely that an important aspect of functionality of this pathway is the integration of its regulation with that of overall carbon and nitrogen metabolism in ADP1. All studies of PtsP in other organisms indicate it plays important metabolic roles that differ according to the details of the bacteria in which it is expressed. It appears that my *lig* studies have also identified PtsP in ADP1 as an important global regulator, even if the details of its physiological role remain to be identified.

PPIases of *A. baylyi* ADP1

A. baylyi encodes eleven predicted peptidyl-prolyl isomerases (PPIases) in its chromosome (**Table 1.5**). In my long-term culturing experiments, strains evolved mutations in or deletions of four different PPIases (**Table 2.6**). After reconstructing mutations or deletions of the

genes *tig* and *fbp*, encoding two separate PPIases, it became clear that deleting one PPIase in combination with *ptsP* allows *A. baylyi* strains to grow through the foreign PCA-4,5 pathway at 30 °C (**Figures 2.8, 2.9, and 2.10**). PPIases are likely affecting the post-translational maturation or stability of certain proteins. Removing a PPIase could be preventing or slowing the maturation of *A. baylyi* proteins that hinder the use of the *lig* genes at increased temperatures, but further studies need to be done to determine what exactly is happening in these strains. Deleting two PPIases encoding genes together however did not have a beneficial or a detrimental effect the strain's ability to grow through the foreign pathway (**Figures 2.9 and 2.10**). However, even though I have not elucidated the mechanism in my *A. baylyi* strains, these results indicate that this class of enzymes, PPIases, is important to the adaptation of *A. baylyi* when growing through the *lig* pathway at increased temperatures.

Unprecedented expansion of the *A. baylyi* chromosome

Previous studies in *A. baylyi* had introduce and evolved several catabolic genes, yet this current study of the *lig*-genes involves the largest set of related genes (to date) to be put into the *A. baylyi* chromosome for EASy evolution (14, 29). My initial success in generating *A. baylyi* strains that can grow through the foreign pathway at 25 °C demonstrate *A. baylyi's* ability to incorporate and express large foreign catabolic pathways in its chromosome. In these studies, I also expanded the chromosome by an unprecedented amount. Previous studies of spontaneous chromosomal amplifications demonstrated that *A. baylyi* could expand its ~3.6 Mbp chromosome by ~1 Mbp, increasing its chromosome size by around 25% (48). In one of my strains, the 11,512 bp amplicon (location 1) had 172 copies in the chromosome at the start of EASy evolution, as measured by the F87 fragment. Assuming that this dosage is uniform for the

entire amplicon, this selected increase in copy number results in 1,980,064 bp, or almost 2 Mbp of extra DNA in the *A. baylyi* chromosome. This quantification equates to increasing the *A. baylyi* chromosome by approximately 50%. This result speaks to the incredible genetic malleability of *A. baylyi* and its potential for bioengineering.

The genetic method to remove multiple amplicons from *A. baylyi* strains developed in this study also points to a useful variation of common types of chromosomal manipulation. I was able to remove up to 10 copies of a 12,853-bp amplicon (location 2) from the chromosome of strains by using a selectable/counter-selectable marker via natural transformation and homologous recombination. With this approach, I could transform the recipient strains with linear DNA carrying a segment on either side of the cassette with 2-4 kbp of sequence identity to the chromosome. This process equates to about 130 kbp deleted from the chromosome in a single selective step. Such a large deletion being easily engineered in *A. baylyi* implies that techniques for chromosomal rearrangements, testing gene essentiality, and developing a minimal *A. baylyi* genome could be relatively facile (13). Genetic manipulations to reduce the genome size could prove impossible by this method after genes are removed that affect natural transformation and recombination. Nevertheless, the creation of specific chromosomal deletions could prove useful in testing a wide variety of hypotheses about genetic function and localization.

This method of amplicon removal and replacement with a single copy of the *lig* genes allowed me to test the significance of the full complement of all chromosomal mutations that arose outside the region of targeted amplification during the course of the 16-week experiments. It was possible to determine if they were sufficient for thermotolerant growth with a single copy of the foreign PCA-4,5 pathway. This is a useful technique because it is typical for strains to acquire multiple chromosomal mutations during laboratory evolution (14, 15, 29, 35). It can be

difficult to test the role of individual mutations or a subset of different combinations. It may not be clear which mutations are providing a benefit and which are neutral, or even detrimental. . I contributed to one novel approach to this problem in a separate study of analyzing mutations after EASy amplification. In a newly accepted publication (Luo, et al), we describe that DNA fragments with specific mutations can be added back to growing cultures, and those that become enriched in the chromosomes of the evolving populations are implicated for playing a role in the selected phenotype (15).

In my study of the *lig* genes, ACN2429 and ACN2430 each acquired six chromosomal mutations. Reconstructing all six mutations in an ADP1 strain would take a long time, even with ADP1's efficient genetic system. However, by removing and replacing the *lig* genes, I could rapidly test the relative contributions of multiple chromosomal mutations outside the target region compared to those within the target region. Target-region effects could arise from mutations and/or from the impact of having multiple amplicon copies. In these specific studies, the combination of chromosomal mutations outside the amplicons were, in most cases, sufficient for growth at 30 °C but not 37 °C. There are additional experiments that become possible for the reconstructed strains. For example, for strains with a single *lig*-gene pathway and a known set of chromosomal mutations, it would be possible to have them undergo a second round of amplification and evolution. Perhaps additional mutations would arise in these strains and higher temperature growth could be achieved without requiring multiple amplicons. Various types of iterative experimental approaches could prove useful.

Not all strains derived from the evolving populations were able to be analyzed by this genetic method of removing amplicons and replacing them with a single pathway copy. When *A. baylyi* ADP1 undergoes laboratory evolution, it has been reported to lose competency for natural

transformation (35, 73, 74). I attempted to transform ten of the *A. baylyi* strains post long-term culturing, and six of these were successful (**Figure 2.7 Panel B**). Only one of these strains had a mutation in a gene known to affect *A. baylyi* competency, *comM* in ACN2437 (75). No genetic changes found explain the potential loss of competency of the other three strains. It is possible that these other strains are still competent. Previous studies have developed quantitative competence assays which could be used with my strains to determine if the failure to obtain the desired transformants resulted from loss of competency (74).

These experiments demonstrated that the EASy method is robust enough to be utilized for whole pathway evolution. Integration of the PCA-4,5 into *A. baylyi* opens up the possibility of further metabolic expansion through the incorporation of other catabolic pathways that funnel into intermediates of the PCA-4,5 pathway (**Figure 2.1**). Using EASy to evolve such pathways will simplify future experiments seeking to expand ADP1's metabolic repertoire further.

Experimental Procedures

Media and growth conditions

Defined ADP1 medium was used for culturing all *A. baylyi* strains as previously described (33). This medium was supplemented with 5 mM PCA, POB, or pyruvate and incubated at 37 °C with liquid cultures shaking at 250 rpm, unless otherwise indicated. For plasmid cloning, *E. coli* XLI Blue or DH5α (Stratagene) was grown in Luria-Bertani medium (LB) (76). All cultures were incubated at 37 °C with liquid cultures shaking at 250 rpm, unless otherwise indicated. Genomic DNA from *P. atlantica* T6c was harvested from cells grown on Marine Broth (BD Difco™) shaking at 25 °C and 250 rpm. Antibiotics were used at final concentrations of 25 μg/ml for kanamycin, 12.5 μg/ml each for streptomycin and spectinomycin,

and 150 μg/ml for ampicillin. Increased concentration of kanamycin used to select for EASy amplification mutants was 1 mg/ml.

Strain and plasmid construction

DNA used for plasmid constructed was PCR-amplified using the high-fidelity polymerases PrimeSTAR Max (Takara Biosciences) or Phusion (New England Biosciences). All primers used can be found in the primer table (**Table 2.9**). Plasmids were constructed by overlapping sequence assembly by NEBuilder (New England Biosciences) or in *E. coli* XL1-Blue or DH5α competent cells (Stratagene) as previously described (cloning should be easy). Standard restriction digest followed by ligation (Quick Ligation Kit, New England Biosciences) was also used for plasmid construction. Plasmids were confirmed by DNA restriction mapping and/or regional DNA sequencing (Eton Biosciences or Eurofins Genomics). Plasmids are described further in **Table 2.8**.

All chromosomal changes in ADP1 were constructed by allelic replacement as previously described (77). Strains were confirmed by PCR and/or regional DNA sequencing. All strains are further described in **Table 2.7**.

Laboratory Evolution by EASy

A. baylyi derived strains carrying the *P. atlantica* T6c PCA-4,5 pathway were evolved for thermal tolerance by the EASy method. Strains were transformed with donor DNA called the **s**ynthetic **b**ridging **f**ragment (SBF), linearized pBAC1414 or pBAC1415 in this study. This fragment induces precise, defined duplications in the chromosome that can undergo subsequent amplification (48, 78). Transformants were then exposed to 1 mg/ml of kanamycin. This

concentration of antibiotic kills cells carrying a single copy of the ΩK^R cassette but not cells that underwent chromosomal amplification (28). Therefore, this step selects for strains with the desired chromosomal amplifications including ΩK^R and the PCA-4,5 pathway. Strains that grew on the high concentration of kanamycin were then moved to liquid ADP1 minimal media with POB or PCA as the sole carbon source at 37 °C. Strains were then cultured with constant selective pressure, growth on POB or PCA at 37 °C. For the first week 10% inoculations were used for daily subcultures. This daily inoculation was then tapered down over multiple days to 0.1% for the rest of the evolution. The number of copies of the amplicon in evolving strains was analyzed by quantitative (q) PCR as described below.

qPCR Copy Number Analysis

The qPCR method used to assess copy number during EASy experiments was previous described (79). Primers for qPCR are listed in the primer table (**Table 2.9**). Standard curves used genomic DNA from strains with a single copy of the amplicon (ACN1910, ACN2118, or ACN2119) for both *rpoA* and ΩK^R. The ratio of calculated concentration for ΩK^R relative to *rpoA* was presumed to equal the copy number of that fragment.

Next Generation Sequencing and Analysis

Genomic DNA was isolated from ADP1 derived strains and sonicated to approximately 500 bp size fragments in a total of 500 ng. Genomic libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) including end-repair, adaptor ligation, and addition of an individual i5/i7 index primer. Sequencing was performed on the

Illumina NextSeq500 instrument at the Georgia Genomics and Bioinformatic Core at the University of Georgia.

Breseq version 0.35.4 was used to analyze WGS data on the UGA GACRC High-Performance Computing Cluster (Sapelo2), Linux based. Version 2.4.1 of bowtie2 and version 4.0.0 of R were used in the pipeline. The consensus mode was used with the default consensus frequency cutoff of 0.8 and polymorphism frequency cutoff of 0.2. The reference genome used for ADP1 was from NCBI entry GenBank CR543861. Differences between our laboratory ADP1 sequence and the NCBI entry can be found in **Table 2.10**.

Growth Analysis

For growth experiments testing final optical density, strains were initially grown overnight at 37 °C with the non-selective carbon source pyruvate. Cultures of ADP1 minimal media with 5 mM POB as the sole carbon source were subsequently inoculated 5 % and placed at either 25 °C, 30 °C, or 37 °C. The optical density was measured at 595 nm on Eppendorf Biophotometer Plus during growth or after 72 hours. All experiments were done in three biological replicates.

For growth experiment to calculate doubling times, strains were grown overnight at 25 or 30 °C on ADP1 minimal medium with 5 mM POB as the sole carbon source. 50 ml of fresh ADP1 minimal medium with 5 mM POB was subsequently inoculated 2 % and placed at either 25 or 30 °C. The optical density was measured at 595 nm hourly during growth on Eppendorf Biophotometer Plus. All experiments were done in three biological replicates.

Table 2.7: Strains used in Chapter 2 study

Strain	Relevant Characteristics	Source
ADP1	Wild type (BD413)	(80)
ACN462	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820	(32)
ACN1910	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910* Insertion of PCA-4,5 pathway genes (4,681,539 – 4,688,831) from *P. atlantica* T6c NCBI entry NC_008228 along with ΩK^R cassette after **1,723,583^b** into strain without native PCA-3,4 dioxygenase and with 90 kbp deletion. Selected by ΩK^R pBAC1488/NsiI X ACN462	This study
ACN1965	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910*; *SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699^b** (in *quiA* ACIAD_RS07910) and **1,724,607^b** (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN1910	This study
ACN2041	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910*; *SBF52041* EASy-derived POB37$^+$ isolate from evolving population of ACN1965 after **14** weeks with multiple copies of a spontaneous new amplicon encompassing PCA-4,5 pathway and part of the ΩK^R cassette found by WGS and other mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2106	Δ*pcaUIJFBDCHG*::lac-*pcaK52106* Deletion of DNA containing native PCA degradation genes from **1,707,705 - 1,717,170^b** replaced with *pcaK* controlled by the lac promoter. Screened for loss of the ability to grow on POB pBAC1561 X ADP1	This study
ACN2118	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^R*52118* Insertion of PCA-4,5 pathway genes (4,681,539 – 4,688,831) from *P. atlantica* T6c NCBI entry NC_008228 along with ΩK^R cassette after **1,723,583^b** into strain without native PCA degradation genes. Selected by K^R pBAC1461/NsiI X ACN2106	This study
ACN2119	Δ*pcaUIJFBDCHG*::lac-*pcaK52106;* ACIAD_RS07700*:ligRKUJIABC52119;* ΩK^R *52119* Insertion of PCA-4,5 pathway genes (4,681,539 – 4,688,831) from *P. atlantica* T6c NCBI entry NC_008228 along with ΩK^R cassette after **1,677,020^b** into strain without native PCA degradation genes. Selected by K^R pBAC1473/NsiI X ACN2106	This study
ACN2120	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910* Insertion of PCA-4,5 pathway genes (4,681,539 – 4,688,831) from *P. atlantica* T6c NCBI entry NC_008228 along with ΩK^R cassette after **1,723,583^b** into strain without native PCA degradation genes. Selected by K^R pBAC1488/NsiI X ACN2106	This study
ACN2121	Δ*pcaUIJFBDCHG*::lac-*pcaK52106;* ACIAD_RS07700*:ligRKUJIABC52121;* ΩK^R*52121* Insertion of PCA-4,5 pathway genes (4,681,539 – 4,688,831) from *P. atlantica* T6c NCBI entry NC_008228 along with ΩK^R cassette after **1,677,020^b** into strain without native PCA degradation genes. Selected by K^R pBAC1489/NsiI X ACN2106	This study
ACN2126	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^R*52118 ; SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699^b** (in *quiA* ACIAD_RS07910) and **1,724,607^b** (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN2118	This study
ACN2127	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 quiA:ligRKUJIABC52118;* ΩK^R*52118; SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699^b** (in *quiA* ACIAD_RS07910) and **1,724,607^b** (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN2118	This study
ACN2128	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ; quiA:ligRKUJIABC52118;* ΩK^R*52118; SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699^b** (in *quiA* ACIAD_RS07910) and **1,724,607^b** (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN2118	This study
ACN2129	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700*:ligRKUJIABC52119;* ΩK^R *52119; SBF52129*	This study

	multiple chromosomal copies of a 12.8 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,676,150** [b] (in ACIAD_RS07700) and **1,679,399**[b] (in ACIAD_RS07705). Selected by high-level Kanamycin, POB37$^+$ pBAC1414/AatII X ACN2119	
ACN2130	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^R *52119; SBF52129* multiple chromosomal copies of a 12.8 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,676,150** [b] (in ACIAD_RS07700) and **1,679,399**[b] (in ACIAD_RS07705). Selected by high-level Kanamycin, POB37$^+$ pBAC1414/AatII X ACN2119	This study
ACN2131	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52121;* ΩK^R *52121; SBF52129* multiple chromosomal copies of a 12.8 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,676,150** [b] (in ACIAD_RS07700) and **1,679,399**[b] (in ACIAD_RS07705). Selected by high-level Kanamycin, POB37$^+$ pBAC1414/AatII X ACN2121	This study
ACN2132	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^R *52119; SBF52129* multiple chromosomal copies of a 12.8 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,676,150** [b] (in ACIAD_RS07700) and **1,679,399**[b] (in ACIAD_RS07705). Selected by high-level Kanamycin, POB37$^+$ pBAC1414/AatII X ACN2119	This study
ACN2133	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R*51910; SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699**[b] (in *quiA* ACIAD_RS07910) and **1,724,607**[b] (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN2120	This study
ACN2134	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R*51910; SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699**[b] (in *quiA* ACIAD_RS07910) and **1,724,607**[b] (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN2120	This study
ACN2135	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R*51910; SBF51965* multiple chromosomal copies of a 11.5 kbp amplicon including *ligRKUJIABC* and ΩK^R; Precise chromosomal duplication connects **1,722,699**[b] (in *quiA* ACIAD_RS07910) and **1,724,607**[b] (in ACIAD_RS07915). Selected by high-level Kanamycin, POB37$^+$ pBAC1415/AatII X ACN2120	This study
ACN2401	D5H6-*pcaK52401*; Δ*pcaUIFJBDKCHG52401* Deletion of *pca* genes **(1,707,705 - 1,717,170)**[b] replaced with the D5H6 promoter controlling *pcaK* **(1,713,983 - 1,715,356)**[b]. Screened for loss of ability to grow on POB pBAC1717/AatII X ADP1[c]	This study
ACN2429	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^R*52118 ; SBF52126* EASy-derived POB37$^+$ isolate from evolving population of ACN2126 after 14 weeks with 12 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2430	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 quiA:ligRKUJIABC52118;* ΩK^R*52118; SBF52127* EASy-derived POB37$^+$ isolate from evolving population of ACN2127 after 14 weeks with 14 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2431	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ; quiA:ligRKUJIABC52118;* ΩK^R*52118; SBF52128* EASy-derived POB37$^+$ isolate from evolving population of ACN2128 after 14 weeks with 10 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2432	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^R *52119; SBF52129* EASy-derived POB37$^+$ isolate from evolving population of ACN2129 after 14 weeks with 13 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2433	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^R *52119; SBF52130* EASy-derived POB37$^+$ isolate from evolving population of ACN2130 after 14 weeks with 15 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2434	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52121;* ΩK^R *52121; SBF52131*	This study

	EASy-derived POB37^{+} isolate from evolving population of ACN2131 after 14 weeks with 14 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	
ACN2435	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^{R} *52119; SBF52132* EASy-derived POB37^{+} isolate from evolving population of ACN2132 after 14 weeks with 6 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2436	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R}*51910; SBF52133* EASy-derived POB37^{+} isolate from evolving population of ACN2133 after 14 weeks with 21 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2437	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R}*51910; SBF52134* EASy-derived POB37^{+} isolate from evolving population of ACN2134 after 14 weeks with 11 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2438	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R}*51910; SBF52135* EASy-derived POB37^{+} isolate from evolving population of ACN2135 after 14 weeks with 7 copies of 11.5 kbp amplicon. Mutations identified in genome (see **Tables 2.2/2.3**)	This study
ACN2456	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ; quiA*:*sacB-S*^{R}*52456* PCA-4,5 amplicons including **1,721,686 – 1,724,335**[b] replaced with *sacB-S*^{R} pBAC1777 AatII X ACN2429	This study
ACN2457	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*sacB-S*^{R}*52457* PCA-4,5 amplicons including **1,675,335 – 1,678,911**[b] replaced with *sacB-S*^{R} pBAC1778 AatII X ACN2432	This study
ACN2458	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*sacB-S*^{R}*52457* PCA-4,5 amplicons including **1,675,335 – 1,678,911**[b] replaced with *sacB-S*^{R} pBAC1778 AatII X ACN2434	This study
ACN2459	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ; quiA*:*sacB-S*^{R}*52456* PCA-4,5 amplicons including **1,721,686 – 1,724,335**[b] replaced with *sacB-S*^{R} pBAC1777 AatII X ACN2438	This study
ACN2462	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ; quiA*:*sacB-S*^{R}*52456* PCA-4,5 amplicons including **1,721,686 – 1,724,335**[b] replaced with *sacB-S*^{R} pBAC1777 AatII X ACN2431	This study
ACN2463	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*sacB-S*^{R}*52457* PCA-4,5 amplicons including **1,675,335 – 1,678,911**[b] replaced with *sacB-S*^{R} pBAC1778 AatII X ACN2435	This study
ACN2464	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*sacB-S*^{R}*52457* PCA-4,5 amplicons including **1,675,335 – 1,678,911**[b] replaced with *sacB-S*^{R} pBAC1778 AatII X ACN2433	This study
ACN2535	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;*K^{R} *52119* Single copy PCA-4,5 pathway replacing *sacB-S*^{R} reintroducing **1,675,335 – 1,678,911**[b] pBAC1807 SbfI X ACN2457	This study
ACN2536	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;*K^{R} *52119* Single copy PCA-4,5 pathway replacing *sacB-S*^{R} reintroducing **1,675,335 – 1,678,911**[b] pBAC1807 SbfI X ACN2458	This study
ACN2537	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^{R}*52118* Single copy PCA-4,5 pathway replacing *sacB-S*^{R} reintroducing **1,721,686 – 1,724,335**[b] pBAC1806 SbfI X ACN2459	This study
ACN2538	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^{R}*52118* Single copy PCA-4,5 pathway replacing *sacB-S*^{R} reintroducing **1,721,686 – 1,724,335**[b] pBAC1806 SbfI X ACN2462	This study
ACN2539	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^{R} *52119* Single copy PCA-4,5 pathway replacing *sacB-S*^{R} reintroducing **1,675,335 – 1,678,911**[b] pBAC1807 SbfI X ACN2463	This study
ACN2540	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^{R} *52119* Single copy PCA-4,5 pathway replacing *sacB-S*^{R} reintroducing **1,675,335 – 1,678,911**[b] pBAC1807 SbfI X ACN2464	This study
ACN2711	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^{R} *52119* ACN2535 growing on POB overnight at 30 °C as a full patch; copy number 1	This study
ACN2712	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^{R} *52119* ACN2536 growing on POB overnight at 30 °C as a full patch; copy number 1	This study
ACN2713	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^{R}*52118* ACN2537 growing on POB overnight at 30 °C as a full patch; copy number 1	This study
ACN2714	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^{R} *52119*	This study

	ACN2539 growing on POB overnight at 30 °C as a full patch; copy number 1	
ACN2718	Δ*pcaUIJFBDCHG*::lac-*pcaK52106 ;* ACIAD_RS07700:*ligRKUJIABC52119;* ΩK^R *52119* ACN2540 growing on POB overnight at 30 °C as a full patch; copy number 1	This study
ACN2719	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^R*52118* ACN2538 spontaneous mutant growing on patch from POB plate at 30 °C #1; copy number 1.2	This study
ACN2720	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^R*52118* ACN2538 spontaneous mutant growing on patch from POB plate at 30 °C #2; copy number 3.5	This study
ACN2721	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC52118;* ΩK^R*52118* ACN2538 spontaneous mutant growing on patch from POB plate at 30 °C #3; copy number 1.5	This study
ACN2756	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* Ω*K*^{R} *51910; tig*::*sacB-S*^{R}*52756* *sacB-S*^{R} interrupting *tig* after **520,256**[b] in ACN1910 background pBAC1881 AatII X ACN1910	This study
ACN2757	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* Ω*K*^{R} *51910; tig*::*sacB-S*^{R}*52756* *sacB-S*^{R} interrupting *tig* after **520,256**[b] in ACN2120 background pBAC1881 AatII X ACN2120	This study
ACN2758	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* Ω*K*^{R} *51910; ptsP*::*sacB-S*^{R}*52758* *sacB-S*^{R} interrupting *ptsP* after **452,882**[b] in ACN1910 background pBAC1882 AatII X ACN1910	This study
ACN2759	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* Ω*K*^{R} *51910; ptsP*::*sacB-S*^{R}*52758* *sacB-S*^{R} interrupting *ptsP* after **452,882**[b] in ACN2120 background pBAC1882 AatII X ACN2120	This study
ACN2769	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910; ptsP52041* *ptsP52041* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN1910 background pBAC1873 AatII/AhdI X ACN2758	This study
ACN2770	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; ptsP52041* *ptsP52041* allele (**Table X**) replacing *sacB-S*^{R} in ACN2120 background pBAC1873 AatII/AhdI X ACN2759	This study
ACN2771	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910; tig*52041 *tig52041* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN1910 background pBAC1872 AatII/AhdI X ACN2756	This study
ACN2772	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; tig*52041 *tig52041* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN2120 background pBAC1872 AatII/AhdI X ACN2757	This study
ACN2815	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815* Δ*ptsP (***451,236 – 453,524**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN1910 background pBAC1896 AatII X ACN2758	This study
ACN2816	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815* Δ*ptsP (***451,236 – 453,524**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background pBAC1896 AatII X ACN2759	This study
ACN2827	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*tig52827* Δ*tig (***520,059 – 521,387**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN1910 background pBAC1897 AatII X ACN2756	This study
ACN2828	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*tig52827* Δ*tig (***520,059 – 521,387**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background pBAC1897 AatII X ACN2757	This study
ACN2829	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910; tig52041; ptsP*::*sacB-S*^{R}*52758* *sacB-S*^{R} interrupting *ptsP* after **452,882**[b] in ACN1910 background with *tig52041* allele pBAC1882 AatII X ACN2771	This study
ACN2830	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; tig52041; ptsP*::*sacB-S*^{R}*52758*	This study

	sacB-S^{R} interrupting *ptsP* after **452,882**[b] in ACN2120 background with *tig52041* allele pBAC1882 AatII X ACN2772	
ACN2831	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^{R} *51910;* Δ*ptsP52815; tig*::*sacB-S*^{R}*52756* *sacB-S*^{R} interrupting *tig* after **520,256**[b] in ACN2120 background with Δ*ptsP* pBAC1881 AatII X ACN2815	This study
ACN2832	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *51910;* Δ*ptsP52815; tig*::*sacB-S*^{R}*52756* *sacB-S*^{R} interrupting *tig* after **520,256**[b] in ACN2120 background with Δ*ptsP* pBAC1881 AatII X ACN2816	This study
ACN2835	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^{R} *51910; tig52041; ptsP52041* *tig52041* allele (**Table 2.2**) in ACN1910 background with *ptsP52041* allele (**Table 2.2**) pBAC1873 AatII X ACN2829	This study
ACN2836	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *51910; tig52041; ptsP52041* *ptsP52041* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN2120 background with *tig52041* allele (**Table 2.2**) pBAC1873 AatII X ACN2830	This study
ACN2837	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^{R} *51910;* Δ*ptsP52815;* Δ*tig52827* Δ*tig (***520,059 – 521,387**[b]) in between start and stop codons in ACN1910 background with Δ*ptsP* pBAC1897 AatII X ACN2831	This study
ACN2838	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *51910;* Δ*ptsP52815;* Δ*tig52827* Δ*tig (***520,059 – 521,387**[b]) in between start and stop codons in ACN2120 background with Δ*ptsP* pBAC1897 AatII X ACN2832	This study
ACN2839	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *52118; tig*::*sacB-S*^{R}*52756* *sacB-S*^{R} interrupting *tig* after **520,256**[b] in ACN2118 pBAC1881 AatII X ACN2118	This study
ACN2840	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *52118; ptsP*::*sacB-S*^{R}*52758* *sacB-S*^{R} interrupting *ptsP* after **452,882**[b] in ACN2118 background pBAC1882 AatII X ACN2118	This study
ACN2841	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^{R} *51910; rppH*::*sacB-S*^{R}*52841* *sacB-S*^{R} upstream *rppH* (after **454,098**[b]) in ACN1910 background pBAC1913 AatII X ACN1910	This study
ACN2842	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *52118; rppH*::*sacB-S*^{R}*52841* *sacB-S*^{R} upstream *rppH* (after **454,098**[b]) in ACN2118 background pBAC1913 AatII X ACN2118	This study
ACN2843	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *51910; rppH*::*sacB-S*^{R}*52841* *sacB-S*^{R} upstream *rppH* (after **454,098**[b]) in ACN2120 background pBAC1913 AatII X ACN2120	This study
ACN2844	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *52118;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2118 background pBAC1920 AatII X ACN2118	This study
ACN2845	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^{R} *51910;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN1910 background pBAC1920 AatII X ACN1910	This study
ACN2846	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *51910;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2120 background pBAC1920 AatII X ACN2120	This study
ACN2856	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *52118; tig52041* *tig52041* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN2118 background pBAC1872 BsaI X ACN2839	This study
ACN2857	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^{R} *52118;* Δ*tig52827* Δ*tig (***520,059 – 521,387**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2118 background pBAC1897 BsaI X ACN2839	This study

ACN2858	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910; rppH52438* *rppH* allele (**Table 2.2**) from ACN2438 replacing *sacB-S*^{R} in ACN1910 background pBAC1899 BsaI X ACN2841	This study
ACN2859	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910; ΔrppHptsP2859* Δ*rppHptsP* (**451,233 – 454,098**[b]) in ACN1910 background pBAC1916 BsaI X ACN2841	This study
ACN2860	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; rppH52438* *rppH* allele (**Table X**) from ACN2438 replacing *sacB-S*^{R} in ACN2118 background pBAC1899 BsaI X ACN2842	This study
ACN2861	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118;* Δ*rppHptsP2859* Δ*rppHptsP* (**451,233 – 454,098**[b]) in ACN2118 background pBAC1916 BsaI X ACN2842	This study
ACN2862	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*rppHptsP2859* Δ*rppHptsP* (**451,233 – 454,098**[b]) in ACN2120 background pBAC1916 BsaI X ACN2843	This study
ACN2863	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; groES52863* *trc* promoter controlling *groES* (inserted after **2,777,287**[b]) in ACN2118 background pBAC1919 BsaI X ACN2844	This study
ACN2864	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910; groES52434* *groES52434* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN1910 background pBAC1904 BsaI X ACN2845	This study
ACN2865	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910; groES52863* *trc* promoter controlling *groES* (inserted after **2,777,287**[b])in ACN1910 background pBAC1919 BsaI X ACN2845	This study
ACN2866	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; groES52434* *groES52434* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN2120 background pBAC1904 BsaI X ACN2846	This study
ACN2867	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; groES52863* *trc* promoter controlling *groES* (inserted after **2,777,287**[b]) in ACN2120 background pBAC1919 BsaI X ACN2846	This study
ACN2868	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; ptsP52041* *ptsP52041* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN2118 background pBAC1873 AatII X ACN2840	This study
ACN2869	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118;* Δ*ptsP52815* Δ*ptsP (***451,236 – 453,524**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2118 background pBAC1896 BsaI X ACN2840	This study
ACN2870	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; rppH52438* *rppH* allele (**Table 2.2**) from ACN2438 replacing *sacB-S*^{R} in ACN2120 background pBAC1899 BsaI X ACN2843	This study
ACN2871	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; groES52434* *groES52434* allele (**Table 2.2**) replacing *sacB-S*^{R} in ACN2118 background pBAC1904 BsaI X ACN2844	This study
ACN2872	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN1910 background with Δ*ptsP* pBAC1920 AatII X ACN2815	This study
ACN2873	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2120 background with Δ*ptsP* pBAC1920 AatII X ACN2816	This study
ACN2875	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815; groES52863* *trc* promoter controlling *groES* (inserted after **2,777,287**[b]) in ACN1910 background with Δ*ptsP* pBAC1919 BsaI X ACN2872	This study
ACN2876	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815; groES52863*	This study

	trc promoter controlling *groES* (inserted after **2,777,287**[b]) in ACN2120 background with Δ*ptsP* pBAC1919 BsaI X ACN2873	
ACN2877	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118;* Δ*tig52827 ptsP*::*sacB-S*^{R}*52758* *sacB-S*^{R} interrupting *ptsP* after **452,882**[b] in ACN2118 background with Δ*tig* pBAC1882 AatII X ACN2857	This study
ACN2878	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910; rppH52438;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN1910 background with *rppH52438* allele pBAC1920 AatII/AhdI X ACN2858	This study
ACN2879	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R*51910;* Δ*rppHptsP2859;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN1910 background with Δ*rppHptsP* pBAC1920 AatII/AhdI X ACN2859	This study
ACN2880	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; rppH52438;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2118 background with *rppH52438* pBAC1920 AatII/AhdI X ACN2860	This study
ACN2881	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*rppHptsP2859;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2120 background with Δ*rppHptsP* pBAC1920 AatII/AhdI X ACN2862	This study
ACN2882	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; rppH52438;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2120 background with *rppH52438* pBAC1920 AatII/AhdI X ACN2870	This study
ACN2883	Δ*pcaHG;* ΔACIAD_RS04410-ACIAD_RS04820*; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* Δ*tig52827;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN1910 background with Δ*ptsP* and Δ*tig* pBAC1920 AatII/AhdI X ACN2837	This study
ACN2884	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118;* Δ*rppHptsP2859;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2118 background with Δ*rppHptsP* pBAC1920 AatII/AhdI X ACN2861	This study
ACN2885	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* Δ*tig52827;* ACIAD_RS12805::*sacB-S*^{R}*52844* *sacB-S*^{R} interrupting ACIAD_RS12805 after **2,778,224**[b] in ACN2120 background with Δ*ptsP* and Δ*tig* pBAC1920 AatII/AhdI X ACN2838	This study
ACN2886	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; groES52863; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2118 background with *trc* promoter controlling *groES* pBAC1933 AatII/AhdI X ACN2863	This study
ACN2887	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; ptsP52041; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2120 background with *ptsP52041* pBAC1933 AatII/AhdI X ACN2770	This study
ACN2888	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2120 background with Δ*ptsP* pBAC1933 AatII/AhdI X ACN2816	This study
ACN2889	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; tig52041; ptsP52041; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2120 background with *ptsP52041* and *tig52041* pBAC1933 AatII/AhdI X ACN2836	This study

ACN2890	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* Δ*tig52827; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2120 background with Δ*ptsP* and Δ*tig* pBAC1933 AatII/AhdI X ACN2838	This study
ACN2891	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815; groES52863; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2120 background with Δ*ptsP* and *trc* promoter controlling *groES* pBAC1933 AatII/AhdI X ACN2876	This study
ACN2892	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; groES52863; fbp::sacB-S*^{R}*52886* *sacB-S*^{R} replacing *fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons in ACN2120 background with *trc* promoter controlling *groES* pBAC1933 AatII/AhdI X ACN2867	This study
ACN2903	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *52118; groES52863;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2118 background with *trc* promoter controlling *groES* pBAC1932 AatII/AhdI X ACN2886	This study
ACN2904	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; ptsP52041;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background with *ptsP52041* pBAC1932 AatII/AhdI X ACN2887	This study
ACN2905	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background with Δ*ptsP* pBAC1932 AatII/AhdI X ACN2888	This study
ACN2906	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; tig52041; ptsP52041;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background with *ptsP52041* and *tig52041* pBAC1932 AatII/AhdI X ACN2889	This study
ACN2907	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815;* Δ*tig52827;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background with Δ*ptsP* and Δ*tig* pBAC1932 AatII/AhdI X ACN2890	This study
ACN2908	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910;* Δ*ptsP52815; groES52863;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background with Δ*ptsP* and *trc* promoter controlling *groES* pBAC1932 AatII/AhdI X ACN2891	This study
ACN2909	Δ*pcaUIJFBDCHG*::lac-*pcaK52106; quiA:ligRKUJIABC51910;* ΩK^R *51910; groES52863;* Δ*fbp52903* Δ*fbp* (**1,019,799 – 1,020,128**[b]) in between start and stop codons replacing *sacB-S*^{R} in ACN2120 background with *trc* promoter controlling *groES* pBAC1932 AatII/AhdI X ACN2892	This study

a. *A. baylyi* strains were derived from ADP1, previously known as *Acinetobacter calcoaceticus* or *Acinetobacter sp. (81)*

b. Bold numbers correspond to positions on the ADP1 chromosome in NCBI entry NC_005966.

c. Underlined text indicates the donor DNA and, where relevant, the restriction enzyme used to linearize a plasmid (pBAC number/Enzyme). The donor DNA was used to transform (X) the indicated recipient strain.

Table 2.8: Plasmids used in Chapter 2 study

Plasmid	Relevant Characteristics	Source
pUC18	ApR; cloning vector	(82)
pUC19	ApR; cloning vector	(82)
pUI1637	Ap^R, K^R; source of ΩK^R cassette	(46)
pBAC1393	Plasmid for inserting lig genes into location 1. Fragment upstream location 1 (MTV580 and MTV581) and downstream (MTV582 and MTV583) and pUC18 cut with SacI and PstI assembled by NEBuilder	This Study
pBAC1397	Plasmid for inserting lig genes into location 2. Fragment upstream location 2 (MTV568 and MTV609) and downstream (MTV569 and MTV608) and pUC18 (MTV570 and MTV571) assembled by NEBuilder	This Study
pBAC1396	Plasmid for inserting *lig* genes into location 1 with Km resistance marker. Constructed by enzyme restriction and ligation of ΩK^R cassette cut from pUI1637 with SacI and XmnI inserted into pBAC1393 cut with SacI	This Study
pBAC1399	Plasmid for inserting *lig* genes into location 2 with Km resistance marker. Constructed by enzyme restriction and ligation of ΩK^R cassette cut from pUI1637 with SacI inserted into pBAC1397 cut with SacI	This Study
pBAC1414	Synthetic bridging fragment for ACIAD_RS07700 and ACIAD_RS07705	This Study
pBAC1415	Synthetic bridging fragment for *quiA* (ACIAD_RS07910) and ACIAD_RS07915	This Study
pBAC1461	Predicted PCA-4,5 pathway PCR amplified with SRB11 and SRB12 from *P. atlantica* digested with AvrII and PspOMI inserted into pBAC1396 digested with SpeI and PspOMI to insert pathway downstream *quiA* on leading strand of *A. baylyi* chromosome	This Study
pBAC1473	Predicted PCA-4,5 pathway PCR amplified with SRB11 and SRB12 from *P. atlantica* digested with AvrII and PspOMI inserted into pBAC1399 digested with SpeI and PspOMI to insert pathway downstream ACIAD_RS07700 on leading strand of *A. baylyi* chromosome	This Study
pBAC1478	pBAC1396 digested with SacI and relegated to put ΩK^R in reverse direction (on lagging strand of *A. baylyi* chromosome)	This Study
pBAC1484	pBAC1399 digested with SacI and relegated to put ΩK^R in reverse direction (on lagging strand of *A. baylyi* chromosome)	
pBAC1488	Predicted PCA-4,5 pathway from *P. atlantica* on pBAC1461 digested with NotI and PspOMI inserted into pBAC1478 digested with NotI and PspOMI to insert pathway downstream *quiA* on lagging strand of *A. baylyi* chromosome	This Study
pBAC1489	Predicted PCA-4,5 pathway from *P. atlantica* on pBAC1473 digested with NotI and PspOMI inserted into pBAC1484 digested with NotI and PspOMI to insert pathway downstream ACIAD_RS07700 on lagging strand of *A. baylyi* chromosome	This Study
pBAC1548	Ap^R, S^R; source of *sacB*-S^R cassette	(83)
pBAC1561	*lac* promoter controlling *pcaK* replacing *pcaUIJFBDCHG* made by NEBuilder of fragments for DNA upstream pca operon (SRB47 and SRB48), *lac* promoter (SRB49 and SRB50), *pcaK* (SRB51 and SRB52), DNA downstream pca operon (SRB53 and SRB54), and pUC19 (digested with BamHI and SalI).	This Study
pBAC1702	*pcaK* and mCherry controlled by D5H6 promoter after ACIAD_RS07840	(12)
pBAC1717	*pcaK* controlled by D5H6 promoter after ACIAD_RS07840 made by self-closure of PCR fragment pBAC1702 amplified with SRB233 and SRB234	This Study
pBAC1773	Deletion of amplicons from *quiA* site plasmid removing all chromosomal DNA included in SBF to avoid knocking out only some of the amplicons. Made by E. coli assembly of PCR fragments pUC19 (SRB346 and SRB347), upstream DNA (SRB348 and SRB349), downstream fragments (SRB350 and SRB351) with BamHI site engineered in between upstream and downstream.	This Study

pBAC1774	Deletion of amplicons from ACIAD_RS07700 site plasmid removing all chromosomal DNA included in SBF to avoid knocking out only some of the amplicons. Generated by E. coli assembly of PCR fragments pUC19 (SRB352 and SRB353), upstream DNA (SRB354 and SRB355), downstream fragments (SRB356 and SRB357) with BamHI site engineered in between upstream and downstream.	This Study
pBAC1777	*sacB*-S^R interrupting and replacing *quiA* through ACIAD_RS07915 to knockout multiple amplicons of PCA-4,5 pathway. *sacB*-S^R from pBAC1548 ligated into BamHI site of pBAC1773	This Study
pBAC1778	*sacB*-S^R interrupting and replacing ACIAD_RS07695 through ACIAD_RS07705 to knockout multiple amplicons of PCA-4,5 pathway. *sacB*-S^R from pBAC1548 ligated into BamHI site of pBAC1773	This Study
pBAC1806	Same as pBAC1461 but with more homology to the ADP1 chromosome. Made by *E. coli* assembly with PCR fragments of PCA-4,5 pathway from pBAC1461 (SRB409 and SRB410) and pBAC1777 backbone (SRB411 and SRB412)	This Study
pBAC1807	Same as pBAC1473 but with more homology to the ADP1 chromosome. Made by *E. coli* assembly with PCR fragments of PCA-4,5 pathway from pBAC1461 (SRB413 and SRB414) and pBAC1777 backbone (SRB415 and SRB416)	This Study
pBAC1872	*tig52041* allele in pUC19. Made by *E. coli* assembly of PCR fragments *tig52041* (SRB560 and SRB561) and pUC19 (SRB562 and SRB563)	This Study
pBAC1873	*ptsP52041* allele in pUC19. Made by *E. coli* assembly of PCR fragments *ptsP52041* (SRB546 and SRB547) and pUC19 (SRB548 and SRB549)	This Study
pBAC1881	*sacB*-S^R interrupting *tig*. *sacB*-S^R from pBAC1548 cut with EcoRV and SfoI and ligated into HpaI site of pBAC1872	This Study
pBAC1882	*sacB*-S^R interrupting *ptsP*. *sacB*-S^R from pBAC1548 cut with EcoRV and SfoI and ligated into EcoRV site of pBAC1873	This Study
pBAC1896	Δ*ptsP* in pUC19. Made by *E. coli* assembly removing *ptsP* in between start and stop codons from pBAC1873 (SRB659 and SRB660)	This Study
pBAC1897	Δ*tig* in pUC19. Made by *E. coli* assembly removing *tig* in between start and stop codons from pBAC1872 (SRB661 and SRB662)	This Study
pBAC1898	*rppHptsP* in pUC19. Made by *E. coli* assembly of PCR fragments *rppH and ptsP* (SRB663 and SRB664) and pUC19 (SRB665 and SRB666)	This Study
pBAC1899	*rppH52438* in pUC19. Made by *E. coli* assembly of PCR fragments *rppH52438* (SRB663 and SRB664) and pUC19 (SRB665 and SRB666)	This Study
pBAC1903	*groES* in pUC19. Made by *E. coli* assembly of PCR fragments *groES* (SRB722 and SRB727) and pUC19 (SRB728 and SRB729)	This Study
pBAC1904	*groES52434* in pUC19. Made by *E. coli* assembly of PCR fragments *groES52434* (SRB722 and SRB727) and pUC19 (SRB728 and SRB729)	This Study
pBAC1913	*sacB*-S^R upstream *rppH*. Made by *E. coli* assembly of PCR fragments *sacB*-S^R from pBAC1548 (SRB688 and SRB689) and the backbone from pBAC1899 (SRB690 and SRB691)	This Study
pBAC1916	Δ*rppHptsP* in pUC 19. Made by *E. coli* assembly removing *rppH and ptsP* in between start codon of *rppH* and stop codon of *ptsP* from pBAC1898 (SRB667 and SRB668)	This Study
pBAC1919	*trc* promoter controlling *groES*. Made by *E. coli* assembly of PCR fragments *trc* promoter from pBAC1660 (SRB724 and SRB725) and pBAC1903 backbone (SRB723 and SRB726)	This Study
pBAC1920	*sacB*-S^R interrupting ACIAD_RS12805. *sacB*-S^R from pBAC1548 cut with XbaI and ligated into XbaI site of pBAC1903	This Study
pBAC1932	Δ*fbp* in between start and stop codons in pUC19. Made by *E. coli* assembly of PCR fragments pUC19 (SRB739 and SRB740), upstream *fbp* (SRB741 and SRB737), and downstream *fbp* (SRB738 and SRB742)	This Study
pBAC1933	*sacB*-S^R replacing *fbp*. Make by *E. coli* assembily of PCR fragments from *sacB*-S^R pBAC1548 (SRB743 and SRB744), pUC19 (SRB739 and SRB740), upstream *fbp* (SRB741 and SRB746), and downstream *fbp* (SRB745 and SRB742)	This Study

Table 2.9: Primers used in Chapter 2 study

Primer	Sequence	Use
MTV274	GCTCGACGCCTTCTATTTCAA	pPCR for *rpoA*
MTV275	TTTACGTCGCATTCTATTGTCTTCTT	pPCR for *rpoA*

MTV302	GCGTTGGCTACCCGTGATA	pPCR for ΩK^R
MTV303	GGAAGCGGTCAGCCCATT	pPCR for ΩK^R
MTV568	CCATGATTACGAATTCATGACACCAGCTTGCAAATTAC	Amplify upstream insertion location 2 for plasmid cloning with MTV609. Overlaps pUC18
MTV569	TTGCATGCCTGCAGTTAGCCACAGTTCAAATGTGCTT	Amplify downstream insertion location 2 for plasmid cloning with MTV608. Overlaps pUC18
MTV570	CATTTGAACTGTGGCTAACTGCAGGCATGCAAGCTT	Amplify pUC18 for location 2 plasmid cloning with MTV571
MTV571	TGCAAGCTGGTGTCATGAATTCGTAATCATGG	Amplify pUC18 for location 2 plasmid cloning with MTV570
MTV580	CCATGATTACGAATTCTTGAAAGTATGGTGTTTTATA	Amplify upstream insertion location 1 for plasmid cloning with MTV581. Overlaps pUC18
MTV581	TTAAGCGATTGAAGAGCTCTTATTTTTCAAGGGCA	Amplify upstream insertion location 1 for plasmid cloning with MTV580. Overlaps downstream fragment with engineered SacI site
MTV582	TTGAAAAATAAGAGCTCTTCAATCGCTTAAAAAAAAGGGAC	Amplify downstream insertion location 1 for plasmid cloning with MTV583. Overlaps upstream fragment with engineered SacI site
MTV583	CAAGCTTGCATGCATGGGTACCAATATATC	Amplify downstream insertion location 1 for plasmid cloning with MTV582. Overlaps pUC18
MTV595	TTGTTGTTTCAGCAGGTGGTG	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *quiA*
MTV608	CAGTGATTAGATAGTTGAGCTCTTAGTTAATGCCATAGCACTGCTT	Amplify downstream insertion location 2 for plasmid cloning with MTV569. Overlaps upstream fragment with engineered SacI site
MTV609	CTATGGCATTAACTAAGAGCTCAACTATCTAATCACTGAAATTTTTATTC	Amplify upstream insertion location 2 for plasmid cloning with MTV568. Overlaps downstream fragment with engineered SacI site
SRB11	GAGTCACCTAGGATCTCATCAAACCGTAGCG	Amplify predicted PCA-4,5 pathway from *P. atlantica* with SRB12. Sits in *ligR*
SRB12	GATGATAGGGCCCTTATACAAGTTGCTGTTCAAGGTC	Amplify predicted PCA-4,5 pathway from *P. atlantica* with SRB11. Sits in *ligC*
SRB13	TTTTCGCTGAACTATCTCAAGGTC	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligR*
SRB14	GCAACGAGGCACAATTTCAGT	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligR*
SRB15	CATTATAGAAAAAACTAATGGTGAGTGTCAT	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligR*
SRB16	CCAAAGTAGAGCAACGCATG	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligK*
SRB17	TGCAGCACCAGTCTTAATTGATTT	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligJ*
SRB18	AGACACTTTAGACCCGCAG	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligJ*
SRB19	GTTTCATTATTCCACATGGCGG	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligU*
SRB20	TGCTGGCTTCCAATGGTC	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligI*
SRB21	TGGCCTGAACAGAGCATTAAA	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligI*
SRB22	TAAGGCGACCACTTCAACATC	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligC*
SRB23	GGGTTCTGAAGGTGTTGAGTT	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in *ligB*
SRB47	TACGAATTCGAGCTCGGTACCCGGGGGAATTCATGGTAATGCAAC	Amplify DNA upstream pca operon. Sits in pUC19 and ACIAD_RS07835
SRB48	GAGGCTTTTGACTTTAATCGAAGTTTGGTTGC	Amplify DNA upstream pca operon. Sits in tonB terminator and ACIAD_RS07840
SRB49	AAACTTCGATTAAAGTCAAAAGCCTCCGGTC	Amplify *lac* promoter. Sits in ACIAD_RS07840 and tonB terminator

SRB50	TTGCCTCCTTAGGCATAGCTGTCCTCCTCTCTAGAG	Amplify lac operon. Sits in *pcaK* and *lac* promoter
SRB82	TAAGTCAAGTTGCGATTATGCCCGT	Sequence confirmation and PCR verification *P. atlantica* PCA-4,5 pathway. Sits in ACIAD_RS07700
SRB233	CAACTTGAATTTTTGACTATGGAACGGGCCGTA	Amplify pBAC1702 sits in *pcaK* and intergenic region between ACIAD_RS07840 and ACIAD_RS07845
SRB234	CAAAAATTCAAGTTGCATCTGTATGTGCAACAAG	Amplify pBAC1702 sits in *pcaK* and intergenic region between ACIAD_RS07840 and ACIAD_RS07845
SRB346	GCTACCGGGTCGCGTGCCGTACCCCCTGCAGGCATGCAAGCTTGG	Amplify pUC19. Sits in pUC19 and ACIAD_RS07930
SRB347	AGTTTCTGCCGCAGCAAAACGGTACCGAGCTCGAATTCAC	Amplify pUC19. Sits in pUC19 and ACIAD_RS07900
SRB348	GGCCAGTGAATTCGAGCTCGGTACCGTTTTGCTGCGGCAG	Amplify DNA upstream PCA-4,5 amplicon. Sits in pUC19 and ACIAD_RS07900
SRB349	ATATTCAATATCACAAGGATCCTTACCATAATGATCCCAATTTACCTGTT	Amplify DNA upstream PCA-4,5 amplicon. Sits in ACIAD_RS07915 and ACIAD_RS07910 separated by engineered BamHI site
SRB350	TGGGATCATTATGGTAAGGATCCTTGTGATATTGAATATCTGGATTGCTC	Amplify DNA downstream PCA-4,5 amplicon. Sits in ACIAD_RS07915 and ACIAD_RS07910 separated by engineered BamHI site
SRB351	GATTACGCCAAGCTTGCATGCCTGCAGGGGGTACGGCACGCG	Amplify DNA upstream PCA-4,5 amplicon. Sits in pUC19 and ACIAD_RS07930
SRB352	AAAAAGCGATGCGGCATCCAGTAACCCTGCAGGCATGCAAGCTTG	Amplify pUC19. Sits in pUC19 and ACIAD_RS07730
SRB353	CAGGCGCCTTCCCTGCATATTGTGGTACCGAGCTCGAATTCACTGG	Amplify pUC19. Sits in pUC19 and ACIAD_RS07680
SRB354	CCAGTGAATTCGAGCTCGGTACCACAATATGCAGGGAAGGCGCCTG	Amplify DNA upstream PCA-4,5 amplicon. Sits in pUC19 and ACIAD_RS07680
SRB355	CCCATAGCAACACTCTAGACTTAATGATTTATCTTCATCTTTCTGCAAAG	Amplify DNA upstream PCA-4,5 amplicon. Sits in ACIAD_RS07705 and ACIAD_RS07695 separated by engineered BamHI site
SRB356	AAGATGAAGATAAATCATTAAGTCTAGAGTGTTGCTATGGGGGATATTCG	Amplify DNA downstream PCA-4,5 amplicon. Sits in ACIAD_RS07705 and ACIAD_RS07695 separated by engineered BamHI site
SRB357	TACGCCAAGCTTGCATGCCTGCAGGGTTACTGGATGCCGCATCGC	Amplify DNA upstream PCA-4,5 amplicon. Sits in pUC19 and ACIAD_RS07730
SRB409	AAGTCTCACATCATATTGAAAGTATGGTGTTTTATACTTGGGCTG	Amplify *quiA* for cloning
SRB410	TTTTAAGAGAAATATATGGGTACCAATATATCTGGAATGATTTTTAAATG	Amplify ACIAD_RS07915 for clonging
SRB411	CCCATATATTTCTCTTAAAATTTCTATAAAATAATTTACTAATAAGTTGC	Amplify ACIAD_RS07915 for clonging
SRB412	AAAACACCATACTTTCAATATGATGTGAGACTTTTCTTGAGGGTC	Amplify *quiA* for cloning
SRB413	TATTTTGGCTTTAGATAGACAATGACACCAGCTTGCAAATTACTTAAAAC	Amplify ACIAD_RS07695 for cloning
SRB414	TTTATAGATAATGTTCAGTCATTAGCCACAGTTCAAATGTGCTTG	Amplify ACIAD_RS07705 for cloning
SRB415	AACTGTGGCTAATGACTGAACATTATCTATAAAACTGATTCATTTAATTG	Amplify ACIAD_RS07705 for cloning
SRB416	GGTGTCATTGTCTATCTAAAGCCAAAATAAATTTAATTATAAGTTCTTTC	Amplify ACIAD_RS07695 for cloning
SRB546	AAAACGACGGCCAGTGAATTCGAGATGATGACAAGGATACCCGATCAGTC	Amplify *ptsP* and clone into pUC19. Sits in pUC19 and ACIAD_RS02080
SRB547	AACAGCTATGACCATGATTACGCCAAGATGATCGACTCAGAAGGTTTCCG	Amplify *ptsP* and clone into pUC19. Sits in pUC19 and *nudH*
SRB548	GGGTCGGAAACCTTCTGAGTCGATCATCTTGGCGTAATCATGGTCATAGC	Amplify *ptsP* and clone into pUC19. Sits in pUC19 and *nudH*
SRB549	ATCAGACTGATCGGGTATCCTTGTCATCATCTCGAATTCACTGGCCGTCG	Amplify *ptsP* and clone into pUC19. Sits in pUC19 and ACIAD_RS02080

SRB560	AACGACGGCCAGTGAATTCGAGAAGACCCAA ATCAGTACTTTGATCACTG	Amplify *tig* and clone into pUC19. Sits in pUC19 and ACIAD_RS02495
SRB561	ATGATTACGCCAAGTTATGCAGATTTAGAATC AATGACCTTTAAATCAGC	Amplify *tig* and clone into pUC19. Sits in pUC19 and *clpX*
SRB562	AAGGTCATTGATTCTAAATCTGCATAACTTGGC GTAATCATGGTCATAGC	Amplify *tig* and clone into pUC19. Sits in pUC19 and *clpX*
SRB563	TACAGTGATCAAAGTACTGATTTGGGTCTTCTC GAATTCACTGGCCGTCG	Amplify *tig* and clone into pUC19. Sits in pUC19 and ACIAD_RS02495
SRB659	ATTTACATGGATGCATACTCCAAAATTT	Remove *ptsP* between start and stop codons
SRB660	CCATGTAAATTAGGCTTCATCAGAAAAGGGCA	Remove *ptsP* between start and stop codons
SRB661	CCATGTAATTCTGGATCGTAAAAAAGGCG	Remove *tig* between start and stop codons
SRB662	AATTACATGGCTTCCTCGATAAATTAGG	Remove *tig* between start and stop codons
SRB663	GACGGCCAGTGAATTCGAGATGATGACAAGGA TACCCGATCAGTC	Amplify *rppH52438* and clone into pUC19. Sits in pUC19 and ACIAD_RS02080
SRB664	GCTATGACCATGATTACGCCAAGATGAGTATC CCCTTTTCTCGCT	Amplify *rppH52438* and clone into pUC19. Sits in pUC19 and ACIAD_RS02100
SRB665	AGCGAGAAAAGGGGATACTCATCTTGGCGTAA TCATGGTCATAGC	Amplify *rppH52438* and clone into pUC19. Sits in pUC19 and ACIAD_RS02100
SRB666	ACTGATCGGGTATCCTTGTCATCATCTCGAATT CACTGGCCGTCG	Amplify *rppH52438* and clone into pUC19. Sits in pUC19 and ACIAD_RS02080
SRB688	GTCGGAAACCTTCTGAGTCGATCATCGATACC GTCGACCTCGA	Insert *sacB*-S^R upstream *rppA*. Sits in *sacB*-S^R and *rppA*
SRB689	ATATTGATAAATTTTAGGTAGCCAAGCCCATG CAACAGAAACTAT	Insert *sacB*-S^R upstream *rppA*. Sits in *sacB*-S^R and upstream *rppA*
SRB690	AGTTTCTGTTGCATGGGCTTGGCTACCTAAAAT TTATCAATATCGTTGAC	Insert *sacB*-S^R upstream *rppA*. Sits in *sacB*-S^R and upstream *rppA*
SRB691	TCGAGGTCGACGGTATCGATGATCGACTCAGA AGGTTTCCGAC	Insert *sacB*-S^R upstream *rppA*. Sits in *sacB*-S^R and *rppA*
SRB692	CTAATTTGGCTACCTAAAATTTATCAATATCGT TGAC	Remove *rppH* and *ptsP* between start and stop codons
SRB693	GCCAAATTAGGCTTCATCAGAAAAGGGCAAAT C	Remove *rppH* and *ptsP* between start and stop codons
SRB722	CGACGGCCAGTGAATTCGAGATGCTCAAACCA TTATCTATCATTTATAAC	Amplify *groES52434* and clone into pUC19. Sits in pUC19 and ACIAD_RS12810
SRB723	AGCCTCCGACCGGAGGCTTTTGACTAAATAAC TCCATCAGACTAAGTCTA	Amplify pBAC1903 backbone and insert *trc* promoter. Sits in *tonB* terminator and downstream ACIAD_RS12805
SRB724	AGCCTCCGGTCGGAGGCTTTTGACTGCGCAAC GCAATTAATGTGAG	Amplify pBAC1660 *trc* promoter. Sits in *tonB* terminator and *trc* promoter
SRB725	TAATGGACGAATGTTGCTCATGGTCTGTTTCCT GTGTGAAATTGTTATCC	Amplify pBAC1660 *trc* promoter. Sits in *trc* promoter and *groES*
SRB726	AACAATTTCACACAGGAAACAGACCATGAGCA ACATTCGTCCATTACATG	Amplify pBAC1903 backbone and insert *trc* promoter. Sits in *trc* promoter and *groES*
SRB727	TATGACCATGATTACGCCAAGTTGTTCTGCGA AATTGTAATGAACCAATC	Amplify *groES52434* and clone into pUC19. Sits in pUC19 and ACIAD_RS12790
SRB728	TTGGTTCATTACAATTTCGCAGAACAACTTGGC GTAATCATGGTCATAGC	Amplify *groES52434* and clone into pUC19. Sits in pUC19 and ACIAD_RS12790
SRB729	GTTATAAATGATAGATAATGGTTTGAGCATCT CGAATTCACTGGCCGTCG	Amplify *groES52434* and clone into pUC19. Sits in pUC19 and ACIAD_RS12810
SRB737	GTTTACATAGCATGTCTACCTTTAATATAAATT CTGTC	Amplify upstream *fbp*. Sits in downstream *fbp* and upstream *fbp*
SRB738	CTATGTAAACCAGGATTATAAAAAAGCCCATT CTG	Amplify downstream *fbp*. Sits in upstream *fbp* and downstream *fbp*
SRB739	GTTGGATCGTATCCGCTTGGCGTAATCATGGTC ATAGCTG	Amplify pUC19 and insert *fbp*. Sits in downstream *fbp* and pUC19
SRB740	ATTGTATTTGTGGTGCTCGAATTCACTGGCCGT CGTTTTAC	Amplify pUC19 and insert *fbp*. Sits in upstream *fbp* and pUC19
SRB741	GCCAGTGAATTCGAGCACCACAAATACAATAC AGGTGATGTAGAAAAG	Amplify upstream *fbp*. Sits in pUC19 and upstream ACIAD_RS04735
SRB742	CATGATTACGCCAAGCGGATACGATCCAACAA GTGTCCCA	Amplify downstream *fbp*. Sits in pUC19 and downstream *fbp*

SRB743	TTTATAATCCTGGTTTAAATACAGAGAATGAAAAGAAACAGATAGATTTT	Amplify *sacB*-S^R to replace *fbp*. Sits in *sacB*-S^R and downstream *fbp*
SRB744	CAGAATTTATATTAAAGGTAGACATGCTATGCGATACCGTCGACCTCGAG	Amplify *sacB*-S^R to replace *fbp*. Sits in *sacB*-S^R and upstream *fbp*
SRB745	TTCTTTTCATTCTCTGTATTTAAACCAGGATTATAAAAAAGCCCATTCTG	Amplify pBAC1932 and insert *sacB*-S^R. Sits in *sacB*-S^R and downstream *fbp*
SRB746	CGAGGTCGACGGTATCGCATAGCATGTCTACCTTTAATATAAATTCTGTC	Amplify pBAC1932 and insert *sacB*-S^R. Sits in *sacB*-S^R and upstream *fbp*

Table 2.10: Allele differences between ADP1 NCBI entry NC_005966 and our laboratory strain

DNA change	Annotation	Loci	Description
(G)10→9	intergenic (+100/-74)	ACIAD_RS00825 → / → ACIAD_RS008	high affinity Zn ABC transporter periplasmic substrate-binding protein/ATP synthase protein I
C→T	D388N (GAC→AAC)	ACIAD_RS05435 ←	deoxyribodipyrimidine photolyase (photoreactivation), FAD-binding
C→T	I104I (ATC→ATT)	ACIAD_RS08285 →	hypothetical protein
T→C	A105A (GCT→GCC)	ACIAD_RS08285 →	hypothetical protein
C→T	V118V (GTC→GTT)	ACIAD_RS08285 →	hypothetical protein
T→C	Y119Y (TAT→TAC)	ACIAD_RS08285 →	hypothetical protein
G→A	K123K (AAG→AAA)	ACIAD_RS08285 →	30hypothetical protein
A→G	A125A (GCA→GCG)	ACIAD_RS08285 →	hypothetical protein
T→C	D128D (GAT→GAC)	ACIAD_RS08285 →	hypothetical protein
A→G	L129L (TTA→TTG)	ACIAD_RS08285 →	hypothetical protein
C→T	G131G (GGC→GGT)	ACIAD_RS08285 →	hypothetical protein
+C	intergenic (+56/+34)	ACIAD_RS11230 → / ← ACIAD_RS112305	hypothetical protein/glutamine synthetase
G→A	I19I (ATC→ATT)	ACIAD_RS11750 ←	GTP-binding protein LepA
A→C	S353R (AGT→CGT)	ACIAD_RS12750 →	periplasmic binding protein of transport/transglycosylase
G→A	intergenic (-237/-520)	ACIAD_RS15545 ← / → ACIAD_RS15555	hypothetical protein/hypothetical protein
C→T	A769V (GCA→GTA)	ACIAD_RS15680 →	two-component sensor

Table 2.11: Loci in 90 kbp deletion

Locus	Gene	Predicted description
ACIAD_RS04410		MULTISPECIES: IS3-like element IS*1236* family transposase
ACIAD_RS04425		NAD-dependent succinate-semialdehyde dehydrogenase
ACIAD_RS04430		hypothetical protein
ACIAD_RS04435		MULTISPECIES: NUDIX hydrolase
ACIAD_RS04440		nicotinate phosphoribosyltransferase
ACIAD_RS04445		ribose-phosphate pyrophosphokinase
ACIAD_RS04450		phosphatase PAP2 family protein
ACIAD_RS04455		amino acid transporter
ACIAD_RS04460		ArgP/LysG family DNA-binding transcriptional regulator
ACIAD_RS04465		GntR family transcriptional regulator
ACIAD_RS04470		ABC transporter ATP-binding protein
ACIAD_RS04475		iron ABC transporter permease
ACIAD_RS04480		ABC transporter substrate-binding protein

ACIAD_RS04485		TonB-dependent receptor
ACIAD_RS04490		GNAT family N-acetyltransferase
ACIAD_RS04495		hypothetical protein
ACIAD_RS04500		isopenicillin N synthase family oxygenase
ACIAD_RS04505		GntR family transcriptional regulator
ACIAD_RS04510	*vanB*	oxidoreductase
ACIAD_RS04515	*vanA*	aromatic ring-hydroxylating dioxygenase subunit alpha
ACIAD_RS04520	*vanK*	MFS transporter
ACIAD_RS04525		outer membrane porin, OprD family
ACIAD_RS04530		FAD-dependent monooxygenase
ACIAD_RS04535		IclR family transcriptional regulator
ACIAD_RS04540		non-heme iron oxygenase ferredoxin subunit
ACIAD_RS04545		hypothetical protein
ACIAD_RS04550		aromatic ring-hydroxylating dioxygenase subunit alpha
ACIAD_RS04555		SDR family oxidoreductase
ACIAD_RS04560		VOC family protein
ACIAD_RS04565		FAD-dependent oxidoreductase
ACIAD_RS04570		MFS transporter
ACIAD_RS04575		cupin domain-containing protein
ACIAD_RS04580		alpha/beta hydrolase
ACIAD_RS04585		SDR family oxidoreductase
ACIAD_RS04590		aspartate dehydrogenase
ACIAD_RS04595		aldehyde dehydrogenase
ACIAD_RS04600		thiamine pyrophosphate-binding protein
ACIAD_RS04605		hypothetical protein
ACIAD_RS04610		PepSY domain-containing protein
ACIAD_RS04615		GNAT family N-acetyltransferase
ACIAD_RS04620		TonB-dependent siderophore receptor
ACIAD_RS04625		DUF1624 domain-containing protein
ACIAD_RS04630		hypothetical protein
ACIAD_RS04635		MULTISPECIES: MFS transporter
ACIAD_RS04640	*mqo*	malate dehydrogenase (quinone)
ACIAD_RS04645	*betA*	choline dehydrogenase
ACIAD_RS04650	*betB*	betaine-aldehyde dehydrogenase
ACIAD_RS04655	*betI*	transcriptional regulator BetI
ACIAD_RS04660		BCCT family transporter
ACIAD_RS04665		choline transporter
ACIAD_RS04670		MarC family protein
ACIAD_RS04675		transcriptional regulator
ACIAD_RS04680	*lipA*	lipoyl synthase
ACIAD_RS04685		thiamine pyrophosphate-dependent dehydrogenase E1 component subunit alpha
ACIAD_RS04690		alpha-ketoacid dehydrogenase subunit beta
ACIAD_RS04695		2-oxo acid dehydrogenase subunit E2
ACIAD_RS04700	*lpdA*	dihydrolipoyl dehydrogenase
ACIAD_RS04705		2,3-butanediol dehydrogenase
ACIAD_RS04710		acetoin reductase
ACIAD_RS04715		acyl-CoA thioesterase
ACIAD_RS04720		hypothetical protein
ACIAD_RS04725		GTP-binding protein
ACIAD_RS04730	*dctA*	C4-dicarboxylate transporter DctA
ACIAD_RS04735		YaeQ family protein
ACIAD_RS04740	*fbp*	FKBP-type peptidyl-prolyl cis-trans isomerase
ACIAD_RS04745	*pgaB*	poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase PgaB
ACIAD_RS04750	*pgaC*	poly-beta-1,6 N-acetyl-D-glucosamine synthase
ACIAD_RS04755	*pgaD*	poly-beta-1,6-N-acetyl-D-glucosamine biosynthesis protein PgaD

ACIAD_RS04760		acyltransferase
ACIAD_RS04765		nicotinamide mononucleotide transporter
ACIAD_RS04770		hypothetical protein
ACIAD_RS04775		dienelactone hydrolase family protein
ACIAD_RS04780		MFS transporter
ACIAD_RS04785		hypothetical protein
ACIAD_RS04790		HAD family hydrolase
ACIAD_RS04795	*def*	peptide deformylase
ACIAD_RS04800		glutaminase
ACIAD_RS04805		helix-turn-helix transcriptional regulator
ACIAD_RS04810		MFS transporter
ACIAD_RS04815		cold shock domain-containing protein
ACIAD_RS04820	*metH*	methionine synthase

Acknowledgments: Melissa Tumen-Velasquez, Lindsey Nilsen, Lauren Slarks, Chantel Duscent-Maitland, and Alyssa Baugh for contributions to this work. Chris Johnson and Gregg Beckham for helpful discussion. Funding from National Science Foundation (grant no. MCB-1615365), U.S. Department (grant no. 2018-67009-27926), Department of Energy (grant no. DE-SC0022220).

References

1. Penney WG. 1934. The theory of the stability of the benzene ring and related compounds. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character 146:223-238.
2. Díaz E. 2004. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility.
3. Seaton SC, Neidle EL. 2018. Using aerobic pathways for aromatic compound degradation to engineer lignin metabolism. Lignin Valorization: Emerging Approaches 19:252.
4. Kamimura N, Sakamoto S, Mitsuda N, Masai E, Kajita S. 2019. Advances in microbial lignin degradation and its applications. Current opinion in biotechnology 56:179-186.

5. Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT. 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proceedings of the National Academy of Sciences 111:12013-12018.

6. Bugg TD, Rahmanpour R. 2015. Enzymatic conversion of lignin into renewable chemicals. Current opinion in chemical biology 29:10-17.

7. Salvachúa D, Karp EM, Nimlos CT, Vardon DR, Beckham GT. 2015. Towards lignin consolidated bioprocessing: simultaneous lignin depolymerization and product generation by bacteria. Green Chemistry 17:4951-4967.

8. Salmela M, Lehtinen T, Efimova E, Santala S, Santala V. 2019. Alkane and wax ester production from lignin-related aromatic compounds. Biotechnol Bioeng 116:1934-1945.

9. Santala S, Efimova E, Santala V. 2018. Dynamic decoupling of biomass and wax ester biosynthesis in *Acinetobacter baylyi* by an autonomously regulated switch. Metabolic Engineering Communications 7:e00078.

10. Santala S, Santala V, Liu N, Stephanopoulos G. 2021. Partitioning metabolism between growth and product synthesis for coordinated production of wax esters in *Acinetobacter baylyi* ADP1. Biotechnology and Bioengineering 118:2283-2292.

11. Jiang X, Palazzotto E, Wybraniec E, Munro LJ, Zhang H, Kell DB, Weber T, Lee SY. 2020. Automating cloning by natural transformation. ACS Synthetic Biology 9:3228-3235.

12. Biggs BW, Bedore SR, Arvay E, Huang S, Subramanian H, McIntyre EA, Duscent-Maitland CV, Neidle EL, Tyo KEJ. 2020. Development of a genetic toolset for the highly

engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. Nucleic Acids Research doi:10.1093/nar/gkaa167.

13. Suarez GA, Dugan KR, Renda BA, Leonard SP, Gangavarapu LS, Barrick JE. 2020. Rapid and assured genetic engineering methods applied to *Acinetobacter baylyi* ADP1 genome streamlining. Nucleic Acids Research doi:10.1093/nar/gkaa204.
14. Tumen-Velasquez MP, Laniohan NS, Momany C, Neidle EL. 2019. Engineering CatM, a LysR-type transcriptional regulator, to respond synergistically to two effectors. Genes 10:421.
15. Jin Luo EM, Stacy Bedore, Ville Santala, Ellen Neidle, and Suvi Santala. 2021. Characterization of highly ferulate-tolerant *Acinetobacter baylyi* ADP1 isolates by a rapid reverse-engineering method. Applied and Environmental Microbiology.
16. Shingler V. 2003. Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. Environmental Microbiology 5:1226-1241.
17. Cain RB. 2018. The uptake and catabolism of lignin-related aromatic compounds and their regulation in microorganisms, p 21-60, Lignin biodegradation: microbiology, chemistry, and potential applications. CRC Press.
18. Brinkrolf K, Brune I, Tauch A. 2006. Transcriptional regulation of catabolic pathways for aromatic compounds in Corynebacterium glutamicum. Genetics and Molecular Research 5:773-789.
19. Davis JR, Sello JK. 2010. Regulation of genes in *Streptomyces* bacteria required for catabolism of lignin-derived aromatic compounds. Applied Microbiology and Biotechnology 86:921-929.

20. Zimmermann T, Sorg T, Siehler SY, Gerischer U. 2009. Role of *Acinetobacter baylyi* Crc in catabolite repression of enzymes for aromatic compound catabolism. Journal of Bacteriology 191:2834-2842.

21. Gerischer U. 2002. Specific and global regulation of genes associated with the degradation of aromatic compounds in bacteria. Journal of Molecular Microbiology and Biotechnology 4:111-121.

22. Harwood CS, Parales RE. 1996. The β-ketoadipate pathway and the biology of self-identity. Annual Review of Microbiology 50:553-590.

23. Brzostowicz PC, Reams AB, Clark TJ, Neidle EL. 2003. Transcriptional cross-regulation of the catechol and protocatechuate branches of the β-ketoadipate pathway contributes to carbon source-dependent expression of the *Acinetobacter sp.* strain ADP1 pobA gene. Applied and Environmental Microbiology 69:1598-1606.

24. Bleichrodt FS, Fischer R, Gerischer UC. 2010. The beta-ketoadipate pathway of *Acinetobacter baylyi* undergoes carbon catabolite repression, cross-regulation and vertical regulation, and is affected by Crc. Microbiology 156:1313-1322.

25. Fischer R, Bleichrodt FS, Gerischer UC. 2008. Aromatic degradative pathways in *Acinetobacter baylyi* underlie carbon catabolite repression. Microbiology 154:3095-3103.

26. Kasai D, Fujinami T, Abe T, Mase K, Katayama Y, Fukuda M, Masai E. 2009. Uncovering the protocatechuate 2, 3-cleavage pathway genes. Journal of Bacteriology 191:6758-6768.

27. Kamimura N, Masai E. 2014. The protocatechuate 4, 5-cleavage pathway: overview and new findings. Biodegradative bacteria:207-226.

28. Tumen-Velasquez M, Johnson CW, Ahmed A, Dominick G, Fulk EM, Khanna P, Lee SA, Schmidt AL, Linger JG, Eiteman MA. 2018. Accelerating pathway evolution by increasing the gene dosage of chromosomal segments. Proceedings of the National Academy of Sciences 115:7105-7110.

29. Pardo I, Jha RK, Bermel RE, Bratti F, Gaddis M, McIntyre E, Michener W, Neidle EL, Dale T, Beckham GT. 2020. Gene amplification, laboratory evolution, and biosensor screening reveal MucK as a terephthalic acid transporter in *Acinetobacter baylyi* ADP1. Metabolic Engineering 62:260-274.

30. Luo J, McIntyre EA, Bedore SR, Santala V, Neidle EL, Santala S. 2021. Characterization of highly ferulate-tolerant *Acinetobacter baylyi* ADP1 isolates by a rapid reverse-engineering method. bioRxiv.

31. Trautwein G, Gerischer U. 2001. Effects exerted by transcriptional regulator PcaU from *Acinetobacter sp. s*train ADP1. Journal of Bacteriology 183:873-881.

32. Eby DM. 2002. Genetic and biochemical studies of aromatic dioxygenase substrate specificity in *Acinetobacter sp.* strain ADP1University of Georgia.

33. Singh A, Bedore SR, Sharma NK, Lee SA, Eiteman MA, Neidle EL. 2019. Removal of aromatic inhibitors produced from lignocellulosic hydrolysates by *Acinetobacter baylyi* ADP1 with formation of ethanol by Kluyveromyces marxianus. Biotechnol Biofuels 12:91.

34. Cuff LE, Elliott KT, Seaton SC, Ishaq MK, Laniohan NS, Karls AC, Neidle EL. 2012. Analysis of IS1236-mediated gene amplification events in *Acinetobacter baylyi* ADP1. J Bacteriol 194:4395-405.

35. Renda BA, Dasgupta A, Leon D, Barrick JE. 2015. Genome instability mediates the loss of key traits by *Acinetobacter baylyi* ADP1 during laboratory evolution. Journal of Bacteriology 197:872-81.

36. Gralton EM, Campbell AL, Neidle EL. 1997. Directed introduction of DNA cleavage sites to produce a high-resolution genetic and physical map of the *Acinetobacter sp.* strain ADP1 (BD413UE) chromosome. Microbiology 143:1345-1357.

37. D'Argenio DA, Segura A, Coco WM, Bünz PV, Ornston LN. 1999. The physiological contribution of *Acinetobacter* PcaK, a transport system that acts upon protocatechuate, can be masked by the overlapping specificity of VanK. Journal of Bacteriology 181:3505-3515.

38. Lu P, Wang W, Zhang G, Li W, Jiang A, Cao M, Zhang X, Xing K, Peng X, Yuan B. 2020. Isolation and characterization marine bacteria capable of degrading lignin-derived compounds. PloS one 15:e0240187.

39. Ni B, Zhang Y, Chen D-W, Wang B-J, Liu S-J. 2013. Assimilation of aromatic compounds by Comamonas testosteroni: characterization and spreadability of protocatechuate 4, 5-cleavage pathway in bacteria. Applied Microbiology and Biotechnology 97:6031-6041.

40. Barry KP, Ngu A, Cohn EF, Cote JM, Burroughs AM, Gerbino JP, Taylor EA. 2015. Exploring allosteric activation of LigAB from *Sphingobium sp*. strain SYK-6 through kinetics, mutagenesis and computational studies. Archives of Biochemistry and Biophysics 567:35-45.

41. Maddocks SE, Oyston PC. 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology 154:3609-3623.

42. Barbe V, Vallenet D, Fonknechten N, Kreimeyer A, Oztas S, Labarre L, Cruveiller S, Robert C, Duprat S, Wincker P. 2004. Unique features revealed by the genome sequence of *Acinetobacter sp.* ADP1, a versatile and naturally transformation competent bacterium. Nucleic Acids Research 32:5766-5779.

43. Durot M, Le Fèvre F, de Berardinis V, Kreimeyer A, Vallenet D, Combe C, Smidtas S, Salanoubat M, Weissenbach J, Schachter V. 2008. Iterative reconstruction of a global metabolic model of *Acinetobacter baylyi* ADP1 using high-throughput growth phenotype and gene essentiality data. BMC systems biology 2:85.

44. Million-Weaver S, Samadpour AN, Moreno-Habel DA, Nugent P, Brittnacher MJ, Weiss E, Hayden HS, Miller SI, Liachko I, Merrikh H. 2015. An underlying mechanism for the increased mutagenesis of lagging-strand genes in Bacillus subtilis. Proceedings of the National Academy of Sciences 112:E1096-E1105.

45. Prentki P, Krisch HM. 1984. In vitro insertional mutagenesis with a selectable DNA fragment. Gene 29:303-313.

46. Eraso JM, Kaplan S. 1994. prrA, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*. Journal of Bacteriology 176:32-43.

47. Reams AB, Neidle EL. 2004. Gene amplification involves site-specific short homology-independent illegitimate recombination in *Acinetobacter sp*. strain ADP1. Journal of Molecular Biology 338:643-656.

48. Reams AB, Neidle EL. 2003. Genome plasticity in *Acinetobacter*: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. Molecular Microbiology 47:1291-1304.

49. Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq, p 165-188, Engineering and analyzing multicellular systems. Springer.

50. Aravind L, Ponting CP. 1997. The GAF domain: an evolutionary link between diverse phototransducing proteins. Trends in biochemical sciences 22:458-459.

51. Schmid FX. 1993. Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. Annual Review of Biophysics and Biomolecular Structure 22:123-143.

52. Quistgaard EM, Weininger U, Ural-Blimke Y, Modig K, Nordlund P, Akke M, Löw C. 2016. Molecular insights into substrate recognition and catalytic mechanism of the chaperone and FKBP peptidyl-prolyl isomerase SlyD. BMC biology 14:1-25.

53. Ishii N. 2017. GroEL and the GroEL-GroES complex. Macromolecular Protein Complexes:483-504.

54. Belas R, Bartlett D, Silverman M. 1988. Cloning and gene replacement mutagenesis of a Pseudomonas atlantica agarase gene. Applied and Environmental Microbiology 54:30-37.

55. Higgins BP, Carpenter CD, Karls AC. 2007. Chromosomal context directs high-frequency precise excision of IS492 in Pseudoalteromonas atlantica. Proceedings of the National Academy of Sciences 104:1901-1906.

56. Higgins BP, Popkowski AC, Caruana PR, Karls AC. 2009. Site-specific insertion of IS 492 in *Pseudoalteromonas atlantica*. Journal of Bacteriology 191:6408-6414.

57. Arora DP, Boon EM. 2012. Nitric oxide regulated two-component signaling in *Pseudoalteromonas atlantica*. Biochemical and biophysical research communications 421:521-526.

58. Ray S, Vigouroux J, Bouder A, Allami MF, Geairon A, Fanuel M, Ropartz D, Helbert W, Lahaye M, Bonnin E. 2019. Functional exploration of *Pseudoalteromonas atlantica* as a source of hemicellulose-active enzymes: evidence for a GH8 xylanase with unusual mode of action. Enzyme and Microbial Technology 127:6-16.

59. Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols 2010:pdb. prot5448.

60. Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Genomics, proteomics & bioinformatics 13:278-289.

61. Maitra RD, Kim J, Dunbar WB. 2012. Recent advances in nanopore sequencing. Electrophoresis 33:3418-3428.

62. Borchert AJ, Henson WR, Beckham GT. 2022. Challenges and opportunities in biological funneling of heterogeneous and toxic substrates beyond lignin. Current Opinion in Biotechnology 73:1-13.

63. Deutscher J, Aké FMD, Derkaoui M, Zébré AC, Cao TN, Bouraoui H, Kentache T, Mokhtari A, Milohanic E, Joyet P. 2014. The bacterial phosphoenolpyruvate: carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. Microbiology and Molecular Biology Reviews 78:231-256.

64. Sánchez-Cañizares C, Prell J, Pini F, Rutten P, Kraxner K, Wynands B, Karunakaran R, Poole PS. 2020. Global control of bacterial nitrogen and carbon metabolism by a PTSNtr-regulated switch. Proceedings of the National Academy of Sciences 117:10234-10245.

65. F. Titgemeyer JW, X. Cui, J. Reizer and M.H. Saier, Jr. . 1994. Proteins of the phosphoenoipyruvate:sugar phosphotransferase system in Streptomyces:possible involvement in the regulationof antibiotic production.

66. Postma PW, Lengeler JW, Jacobson GR. 1993. Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. Microbiological Reviews 57:543-594.

67. Pflüger-Grau K, Görke B. 2010. Regulatory roles of the bacterial nitrogen-related phosphotransferase system. Trends in Microbiology 18:205-214.

68. Powell BS, Inada T, Nakamura Y, Michotey V, Cui X, Reizer A, Saier MH, Reizer J. 1995. Novel proteins of the phosphotransferase system encoded within the rpoN operon of *Escherichia coli* enzyme IIANtr affects growth on organic nitrogen and the conditional lethality of an erats mutant. Journal of Biological Chemistry 270:4822-4839.

69. Martinez SE, Beavo JA, Hol WG. 2002. GAF domains: two–billion–year–old molecular switches that bind cyclic nucleotides. Molecular Interventions 2:317.

70. Speiser Y, Zusman T, Pasechnek A, Segal G. 2017. The *Legionella pneumophila* incomplete phosphotransferase system is required for optimal intracellular growth and maximal expression of PmrA-regulated effectors. Infection and immunity 85:e00121-17.

71. Deana A, Celesnik H, Belasco JG. 2008. The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. Nature 451:355-358.

72. Bedoya-Pérez LP, Muriel-Millán LF, Moreno S, Quiroz-Rocha E, Rivera-Gómez N, Espín G. 2018. The pyrophosphohydrolase RppH is involved in the control of RsmA/CsrA expression in *Azotobacter vinelandii* and *Escherichia coli*. Microbiological Research 214:91-100.

73. Suarez GA, Renda BA, Dasgupta A, Barrick JE. 2017. Reduced Mutation Rate and Increased Transformability of Transposon-Free *Acinetobacter baylyi* ADP1-ISx. Applied Environmental Microbiology 83.

74. Bacher JM, Metzgar D, de Crécy-Lagard V. 2006. Rapid evolution of diminished transformability in *Acinetobacter baylyi*. Journal of Bacteriology 188:8534-8542.

75. Leong CG, Bloomfield RA, Boyd CA, Dornbusch AJ, Lieber L, Liu F, Owen A, Slay E, Lang KM, Lostroh CP. 2017. The role of core and accessory type IV pilus genes in natural transformation and twitching motility in the bacterium *Acinetobacter baylyi.* PloS one 12:e0182139.

76. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press.

77. Neidle E, Hartnett C, Ornston L. 1989. Characterization of *Acinetobacter calcoaceticus* catM, a repressor gene homologous in sequence to transcriptional activator genes. Journal of Bacteriology 171:5410-5421.

78. Seaton SC, Elliott KT, Cuff LE, Laniohan NS, Patel PR, Neidle EL. 2012. Genome□ wide selection for increased copy number in *Acinetobacter baylyi* ADP1: locus and context□dependent variation in gene amplification. Molecular Microbiology 83:520-535.

79. Stoudenmire JL, Schmidt AL, Tumen-Velasquez MP, Elliott KT, Laniohan NS, Whitley SW, Galloway NR, Nune M, West M, Momany C. 2017. Malonate degradation in Acinetobacter baylyi ADP1: Operon organization and regulation by MdcR. Microbiology 163:789.

80. Juni E, Janik A. 1969. Transformation of *Acinetobacter calco-aceticus* (Bacterium anitratum). Journal of Bacteriology 98:281-288.

81. Vaneechoutte M, Dijkshoorn L, Tjernberg I, Elaichouni A, de Vos P, Claeys G, Verschraegen G. 1995. Identification of *Acinetobacter* genomic species by amplified ribosomal DNA restriction analysis. Journal of Clinical Microbiology 33:11-15.

82. Norrander J, Kempe T, Messing J. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26:101-106.

83. Jones RM, Williams PA. 2003. Mutational analysis of the critical bases involved in activation of the AreR-regulated σ54-dependent promoter in *Acinetobacter* sp. strain ADP1. Applied and Environmental Microbiology 69:5627-5635.

CHAPTER 3

REGULATION OF ASPARTATE TRANSPORT AND METABOLISM IN *ACINETOBACTER BAYLYI* ADP1

Introduction

The role of D-aspartate (D-Asp) as a physiological signaling molecule in diverse vertebrates and invertebrates is evidenced by its wide range of neuroendocrine effects (1). In humans, a variety of D-amino acids arise from bacterial sources and impact health and disease (2). In bacteria, D-amino acids are key components of peptidoglycan. Additionally, they affect processes such as sporulation, biofilm formation, and some types of symbioses (3). An increased awareness of their biological importance has led to renewed attention to the microbial production of D-amino acids, especially by lactic acid bacteria (4).

The biodegradation of D-amino acids has also been investigated. Improved analytical techniques confirm the natural presence and bioavailability of these compounds (5). Several studies demonstrate that D-amino acids are catabolized by soil and marine bacteria (6, 7). Nevertheless, the details of D-amino acid uptake and regulated bacterial consumption remain unclear. It was shown that a genetically malleable, saprophytic, soil bacterium, *Acinetobacter baylyi* ADP1, requires a LysR-type transcriptional regulator (LTTR), DarR, to grow on D-Asp as the sole carbon source (8). This regulator responds to D-Asp to activate transcription of a gene encoding an Aspartate racemase, RacD, and a membrane transport protein, AspT. This transporter belongs to a sodium-dicarboxylate symporter family (SDF, InterPro classification

IPR023954). The genomic clustering of *aspT*-like genes in diverse bacteria with homologs of *racD* and *darR* suggests AspT mediates D-Asp uptake (8), although this function has not been demonstrated.

This study aimed to clarify how metabolism of D-Asp is coordinated with that of its enantiomer, L-Asp. Growth of *A. baylyi* on D-Asp as the carbon source depends on RacD, suggesting that L-Asp is formed and then cleaved by aspartate ammonia lyase (AspA, **Figure 3.1: Panel A**). In the *A. baylyi* chromosome, *aspA* is located upstream and in the opposite direction as a gene encoding an LTTR. This LTTR is predicted to affect transcription of its own gene as well as *aspA* (**Figure 3.1: Panel B**) (8, 9). The encoded LTTR, designated AalR (for aspartate ammonia lyase regulator) share sequence homology to DarR. Other researchers showed that an LTTR similar to AalR regulates *aspA* transcription in *Pseudomonas aeruginosa*, a bacterium that does not grow on D-Asp as the carbon source (10). Similarities between AalR and DarR in *A. baylyi* raise questions about whether each regulator responds to D-Asp and/or L-Asp as effector molecules and whether the response is specific to one enantiomer. Additional questions arise as to how these regulators distinguish different operator-promoter sequences considering that the N-terminal DNA-binding domains of DarR and AalR share 83% amino acid sequence similarity.

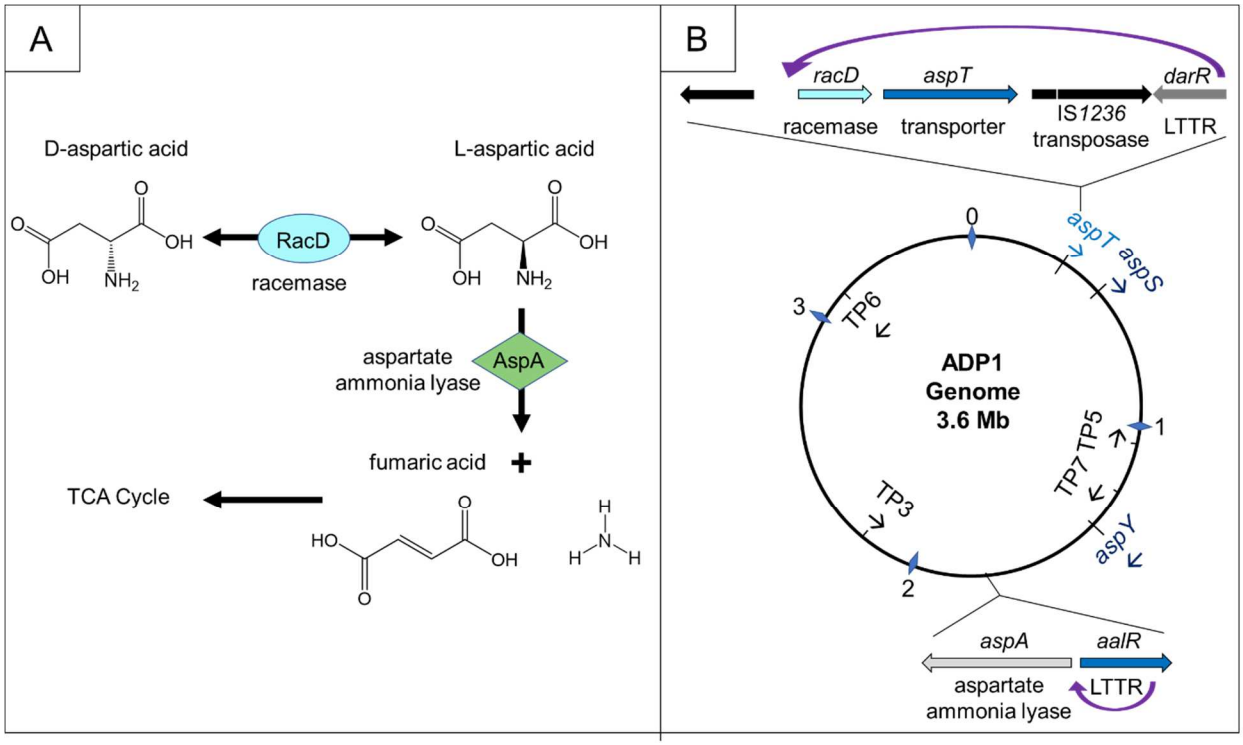


Figure 3.1: Aspartate catabolism pathway and genes in *A. baylyi* ADP1

Panel A: D-Asp is converted to L-Asp by a racemase, RacD. Next, L-Asp is cleaved by a lyase, AspA. Fumarate is further metabolized in the tricarboxylic acid (TCA) cycle. Panel B: The *racD* gene (ACIAD_RS01510) is adjacent to a gene (ACIAD_RS01515), designated *aspT,* predicted to encode a D-Asp transporter. The chromosomal positions and orientations of *aspT* and six genes encoding paralogous transport proteins (TP) are indicated (described further in **Figure 3.3**). Two of the TP genes (TP2 and TP4 were renamed *aspY* and *aspS*, respectively, based on evidence of their roles in Asp transport. Studies in this report support the following regulatory model. Co-transcription of *racD* and *aspT* can be activated by an LTTR, DarR, in response to D-Asp. A DarR paralog, AalR, activates transcription of a divergently oriented gene, *aspA*, in response to L-Asp. Curved arrows point to regulatory regions involved in DNA-LTTR binding.

To determine whether one regulator can substitute for the loss of the other, *darR* and *aalR* mutants were constructed. Additional genes involved in aspartate metabolism were also modified or deleted. Growth phenotypes were tested using several amino acids as carbon or nitrogen sources, including both isomers of aspartate, glutamate, and asparagine. Regulatory effects were investigated using transcriptional fusions, transcript characterization, and electrophoretic mobility shift assays (EMSAs). To test the role of transporters in D-Asp uptake, *A. baylyi* proteins were expressed in *Escherichia coli*, which naturally consumes L-Asp but not D-Asp.

These studies revealed redundancy in genes and functions related to aspartate transport and regulation in *A. baylyi* ADP1. Such redundancy was explored further through the isolation and characterization of suppressor mutations that altered regulation and enhanced the ability of one paralog to take the place of another. Collectively these results, which identified two inducible transport proteins for D-Asp uptake, support the conclusion that D-amino acids can be a significant source of nutrients for bacteria in natural environments.

Results

Potential Asp transporters in ADP1

In previous studies, researchers demonstrated that loss of *darR* or *racD* in ADP1 prevents mutants from growing on D-Asp, but not L-Asp, as the sole carbon source (8). These results support a regulatory model in which DarR responds to D-Asp to activate transcription of *racD* and *aspT* (**Figure 3.1**). To investigate this model further, *aspT* (ACIAD_RS01515) was deleted to test its predicted role in D-Asp uptake. The resulting Δ*aspT* mutant (ACN2035) grew comparably to the wild-type strain (ADP1) on L-Asp (**Figure 3.2 Panel A**: mutant patches in region 1 compared to wild-type patches in region 4). On D-Asp, the mutant without AspT was

able to grow, although there was a delay relative to wild-type. As shown in **Figure 3.2: panel B,** the mutant patches (in region 1) were visible only after two days of incubation compared to wild-type patches (region 4) that are visible after one day. Thus, while AspT appears to contribute to growth on D-Asp, it is not required.

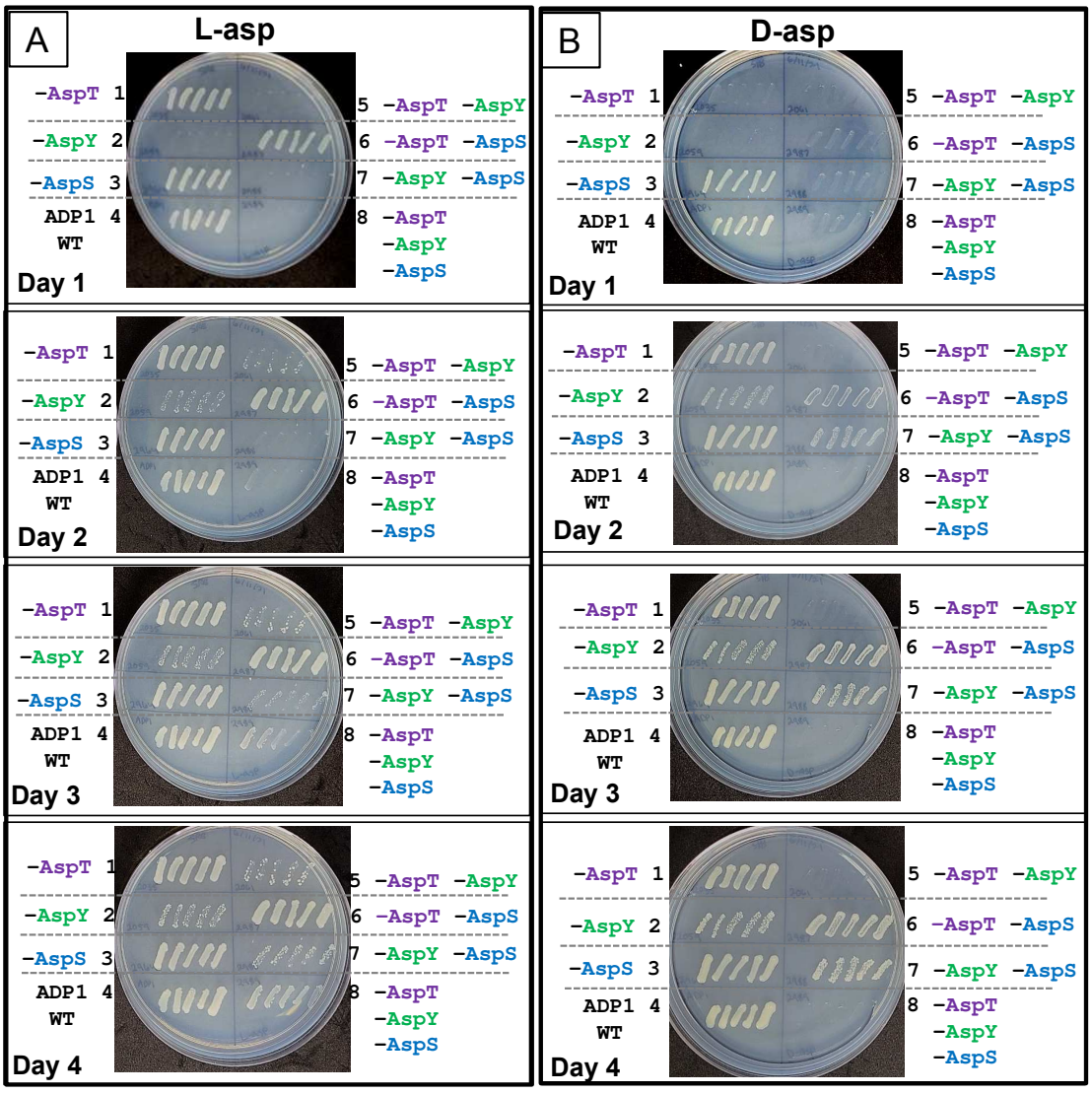


Figure 3.2: Patched colonies of the wild type (ADP1) and *aspT*, *aspY*, and *aspS* mutants

Each colony was patched in the same position on plates with L-Asp as the carbon source, D-Asp as the carbon source (or LB medium as a positive control, not shown). Five pyruvate-grown

colonies from a single plate were patched for each of eight strains: ACN2035 (no AspT) in region 1, ACN2059 (no AspY) in region 2, ACN2964 (no AspS) in region 3, ADP1 (wild type) in region 4, ACN2061 (no AspT or AspY) in region 5, ACN2987 (no AspT or AspS) in region 6, ACN2988 (no AspY or AspS) in region 7, and ACN2967 (no AspT, AspY, or AspS) in region 8. All pyruvate-grown colonies were of comparable size, and all patches were comparable on the LB control plate (not shown). Plates were incubated at 37 °C. Panel A: The L-Asp plate was photographed after 1, 2, 3, and 4 days of incubation (top to bottom). Panel B: The D-Asp plate was similarly photographed after 1, 2, 3, and 4 days of incubation (top to bottom).

Since a redundant D-Asp transporter in *A. baylyi* ADP1 could mask the impact of deleting *aspT,* we searched for homologs by sequence similarity. Six putative *A. baylyi* proteins were identified that belong to the same family of sodium-dicarboxylate symporters as AspT (InterPro IPR023954). Except for *aspT*, genomic location gave no clues about the roles of the transport proteins (TP). These paralogs, which have not been studied in ADP1, were designated TP2-TP7, in descending order of similarity to AspT (which corresponds to TP1; **Figure 3.3**). In pairwise alignments with AspT, TP2 (designated AspY), had 61% sequence identity, whereas the others (TP3-TP7) were less similar, with sequence identities of 38% for TP3, 36% for TP4 (designated AspS), 32% for TP5, 30% for TP6, and 24% for TP7.

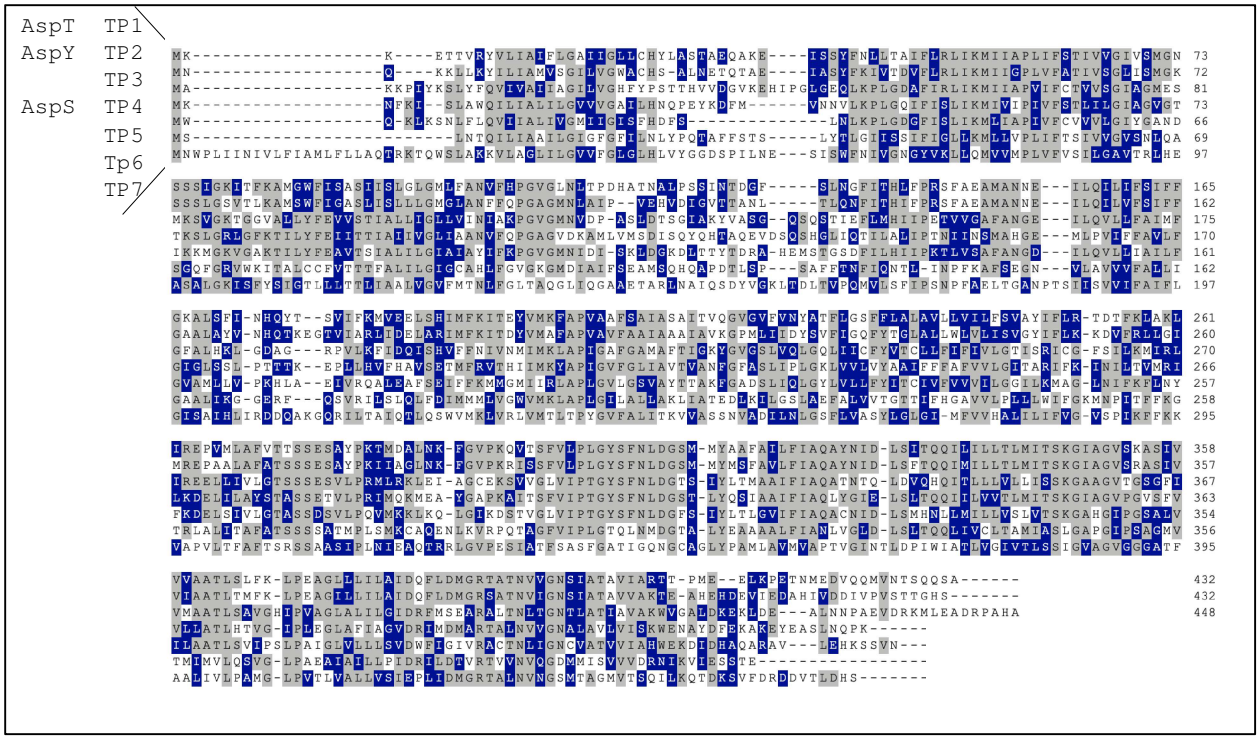


Figure 3.3: Alignment of AspT paralogs from ADP1

Sequence alignment of the six paralogs of AspT encoded by genes in different chromosomal locations. In this sequence alignment, the predicted transport proteins (TP) are numbered TP2-TP7 according to decreasing sequence similarity to AspT (TP1, ACIAD_RS01515). TP2 (ACIAD_RS06160) and TP4 (ACIAD_RS02190, annotated as GltP) were designated AspY and AspS, respectively, based on discovered roles in Asp uptake. Locus tags for TP3, TP5 (sometimes annotated as DctA), TP6, and TP7 are ACIAD_RS10205, ACIAD_RS04730, ACIAD_RS14645, and ACIAD_RS05475, respectively.

Insertional inactivation of *aspY*, the gene with the greatest similarity to *aspT,* impaired growth on aspartate. Disruption of this gene (ACIAD_RS06160), unlike deletion of *aspT*, caused poor growth on L-Asp (**Figure 3.2: Panel A**, patches in region 1 compared to region 2). Because patched colonies of this mutant took several days to grow, we hypothesize a role for AspY in L-

Asp transport. Growth of the *aspY* mutant was also delayed on D-Asp relative to that of ADP1, with a similar pattern to that of the *aspT* mutant (**Figure 3.2: Panel B**, patches in regions 1, 2, and 4). A mutant unable to produce either AspT or AspY failed to grow on D-Asp, even after prolonged incubation (**Figure 3.2: Panel B**, region 5). Yet, this strain, ACN2061 was able to use L-Asp as a sole carbon source, although it grew slowly (**Figure 3.2: Panel A**, region 5). Collectively, these results suggest that AspY transports L-Asp and D-Asp, whereas AspT transports D-Asp. AspY and AspT together may account for all D-Asp transport, as the loss of both precludes growth on this carbon source.

Selection of a spontaneous mutant that grows on D-Asp without AspY or AspT

To explore transport further, spontaneous mutants were isolated from ACN2061, missing *aspY* and *aspT*, that were D-Asp$^+$. These mutants were obtained after ACN2061 was grown on different dicarboxylic acids, concentrated, and spread on plates with D-Asp as the carbon source. The rationale for growing the starting populations on different carbon sources was to induce transport proteins that might provide some initial advantage when cells were transferred to D-Asp medium. Each ACN2061 culture was grown overnight with a single dicarboxylic acid as the carbon source (L-Asp, L-Glu, adipate, or succinate). An additional culture was grown on a non-dicarboxylic acid (pyruvate). D-Asp$^+$ colonies arose (at a frequency of 6×10^{-10}), but only when the overnight culture was grown on L-Asp. One such mutant, designated ACN2921, was characterized further.

We used a transformation assay to determine whether ACN2921 growth on D-Asp resulted from mutations in genes encoding the AspT and AspY paralogs (TP3-TP7, **Figure 3.4**). This assay is based on the high efficiency of natural transformation and homologous

recombination in *A. baylyi*. This approach has long been used for mutational analysis of ADP1-derived mutants (11). Linear DNA amplified from different chromosomal regions of ACN2921 is dropped directly onto a lawn of ACN2061. We then demand growth on D-Asp as only cells that have acquired a compensatory mutation should grow on this carbon source. As shown in **Figure 3.4**, the ACN2061 culture that was spread across the entire plate did not grow across most of the surface since D-Asp is the only carbon source. However, transformants grew in spots where donor DNA replaced the corresponding region of the ACN2061 recipient to introduce a compensatory genetic change. DNA from the chromosomal region of ACN2921 that encompasses the TP4 gene, *aspS*, (ACIAD_RS02190) transformed ACN2061 to confer a D-Asp$^+$ phenotype (**Figure 3.4: Panel A**). As expected, *aspS* DNA amplified from the original strain (ACN2061) did not allow cells to grow on D-Asp, suggesting that an *aspS* mutation is responsible for the growth on D-Asp in strains missing AspY and AspT. A single point mutation in this region of DNA from ACN2921 was found upstream of the coding sequence, in a position that could affect *aspS* transcription (**Figure 3.4 Panel B**).

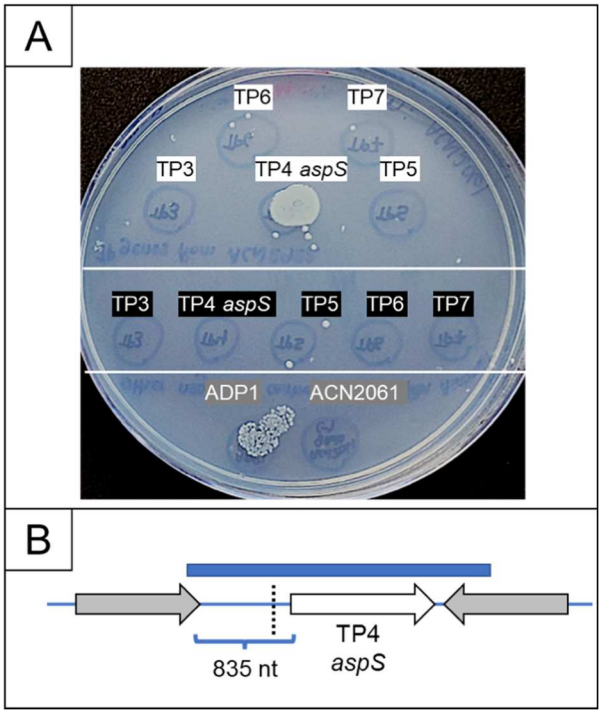


Figure 3.4: Transformation assay

Panel A: A lawn of the recipient strain, ACN2061 (lacking AspY and AspT), was spread on a plate with D-Asp as the sole carbon source. Linear DNA was dropped in marked spots. PCR fragments, made with genomic DNA as template, encompassed the sequences of genes for TP3, TP4, TP5, or TP6 from ACN2921, a D-Asp+ spontaneous mutant derived from ACN2061. Each PCR product was dropped in a spot as shown (labeled with black text on white background). Comparable fragments, made with genomic template DNA from ACN2061, were dropped in spots indicated by labels with white text on black background. At the bottom, transformation with genomic DNA from ADP1 (able to re-introduce *aspT* and *aspY*) or DNA from ACN2061 served as positive and negative controls, respectively. Panel B: Schematic showing the relative positions of the TP4 coding sequence (*aspS*, white arrow), the PCR product (blue rectangle) used in the transformation assay, and a mutation (dotted line) identified in ACN2921. At this position,

200 nt upstream of the *aspS* coding sequence, the wild-type sequence (GTTTA) is altered by the insertion of one extra T (GTTTTA).

High expression of *aspS* enables growth on D-Asp without AspY or AspT

A *lacZ* cassette was inserted into *aspS,* in strains with different genetic backgrounds. The cassette serves as a transcriptional reporter and, additionally, it prevents synthesis of AspS. An insertional inactivation of *aspS* in ACN2921 caused this spontaneous D-Asp$^+$ mutant to lose the ability to grow with D-Asp as the carbon source. This observation confirms that growth of the spontaneous mutant on D-Asp requires both AspS and the mutation indicated in **Figure 3.4: Panel B**. AspS can compensate for the loss of AspT and AspY, but only if there is a mutation that appears to affect its expression. This mutation was designated allele *aspS52921*, and it was assumed (and later shown) to be in the *aspS* promoter.

When comparing the growth of ADP1-derived strains on L-Asp or D-Asp as the carbon source, there is no observable difference caused by the loss of AspS alone or in combination with AspY and/or AspT (**Figure 3.2**). Therefore, the primary function of AspS is not evident. Nevertheless, spontaneous D-Asp$^+$ mutants derived fromACN2061 were only isolated when this strain was grown on L-Asp, raising the possibility that AspS is induced during growth on L-Asp. This possibility was tested with a strain, ACN2964, that has the *aspS*::*lacZ* transcriptional fusion inserted in the chromosome of an otherwise wild-type background. When ACN2964 was grown with L-Asp as the carbon source, LacZ activity that is controlled by the wild-type *aspS* promoter was 5-fold higher than when pyruvate was used as the carbon source for this strain (paired T-test value: $P < 0.0001$) (**Figure 3.5**). This increase in expression suggests some role for AspS in Asp uptake and metabolism.

Expression of the *aspS*::*lacZ* reporter was also compared in two pyruvate-grown strains lacking AspT and AspY. As shown in **Figure 3.5**, LacZ activity was high for ACN2969, which was engineered to carry the *aspS52921* promoter mutation (shown as P_{aspS*} in **Figure 3.5**). The expression of LacZ controlled by this promoter mutation was significant higher than it was for ACN2967, which has the wild-type *aspS* operator-promoter region. These results support the conclusion that a regulatory mutation in ACN2091 causes high AspS expression that enables growth on D-Asp despite the absence of AspT and AspY.

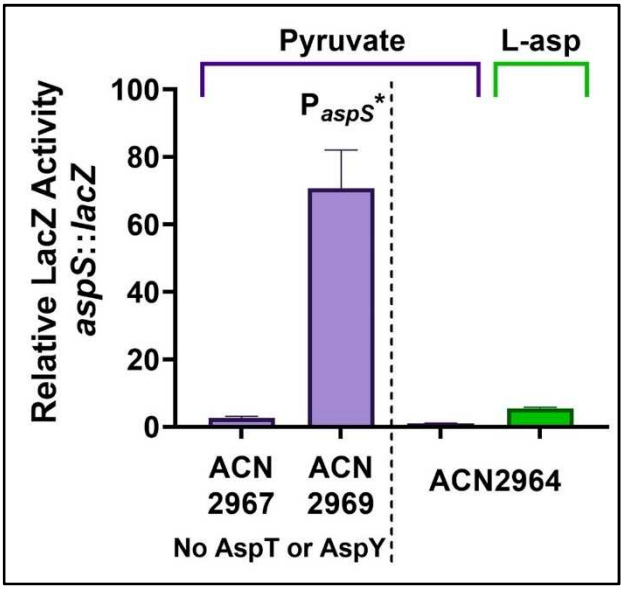


Figure 3.5: LacZ expression from the *aspS* promoter (P_{aspS})

Each strain has a chromosomal transcriptional reporter, *aspS*::*lacZ*. LacZ activity is shown relative to that of ACN2964, which is otherwise wild type, grown on pyruvate (150 Miller units). Purple bars indicate that strains were grown on pyruvate as the sole carbon source. The green bar corresponds LacZ activity when ACN2964 was grown on L-Asp as the sole carbon source. P_{aspS}* represents the mutated allele (*aspS52921*) with a T insertion in the *aspS* promoter. Error bars

show standard deviation of at least three biological replicates done in two technical replicates each on three separate days.

Assessing the roles of D-Asp transporters from *A. baylyi* ADP1 in *E. coli*.

We investigated whether the aspartate racemase (RacD) and/or transport proteins from *A. baylyi* ADP1 (TP1-TP7) would enable *E. coli* to grow on D-Asp. Plasmid-borne genes from ADP1 were introduced into *E. coli* strain DH5α, and optical density at 595 nm was measured after incubation with L-Asp or D-Asp as the sole carbon source (**Figure 3.6**). With no ADP1 proteins being expressed, cell density in D-Asp medium was comparable to that with no carbon source, confirming that this *E. coli* strain cannot consume D-Asp. In contrast, this strain, DH5α(pJPK13), grew well on L-Asp. Two plasmids (pBAC1673 and pBAC1941) conferred growth on D-Asp. Both encode RacD, a racemase not encoded by *E. coli*, and either AspT or AspY. Neither AspT nor AspY alone conferred growth. These results support the conclusion that AspT and AspY transport D-Asp and that racemization is needed for its consumption in *E. coli*.

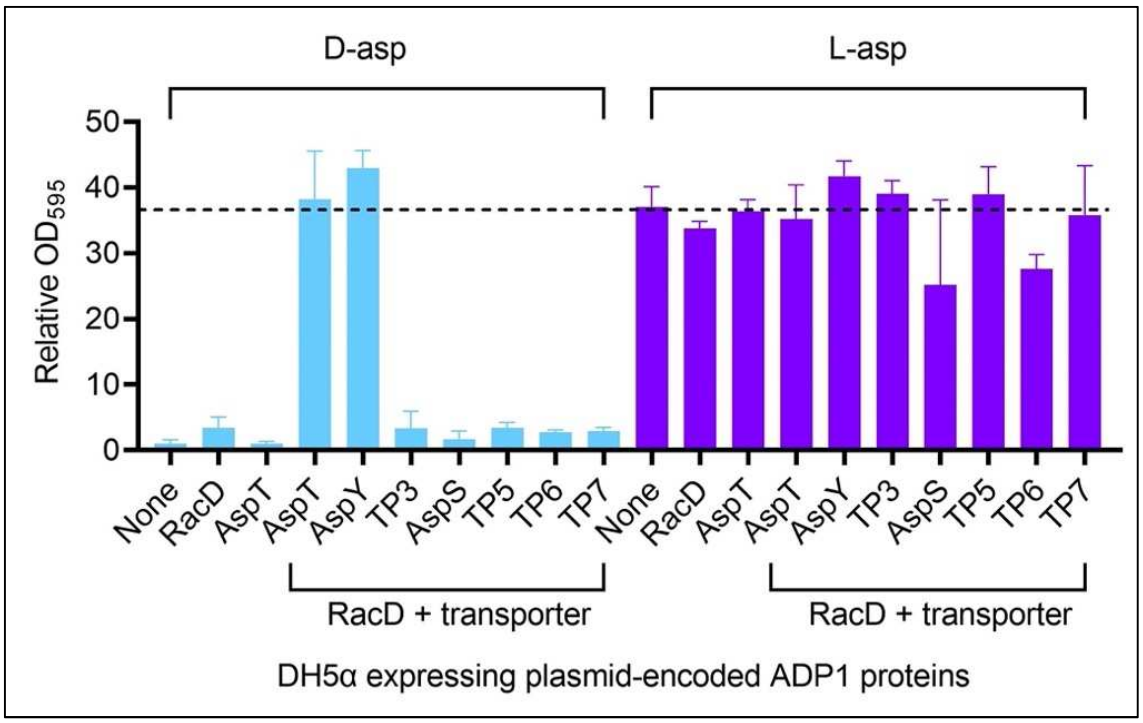


Figure 3.6: Heterologous expression of *A. baylyi* ADP1 proteins in *E. coli* DH5α

Growth is shown relative to the OD_{595} of DH5α(pJPK13) in D-Asp medium. This strain grew to OD_{595}= 0.03 on D-Asp and OD_{595} = 1 on L-Asp. The latter level is indicated by a dashed line. Plasmid-containing DH5α was grown on D-Asp (left side) or L-Asp (right side) as the carbon source. Each strain has either pJPK13 with no ADP1 DNA (None) or a derivative of this plasmid with DNA encoding ADP1 proteins. The following plasmids encode the ADP1 protein(s) indicated in parentheses: pBAC1680 (RacD), pBAC1683 (AspT), pBAC1673 (RacD and AspT), pBAC1941 (RacD and AspY), pBAC1942 (RacD and TP3), pBAC1943 (RacD and AspS), pBAC1944 (RacD and TP5), pBAC1945 (RacD and TP6), or pBAC1946 (RacD and TP7). Measurements (OD_{595}) were taken 72 h after inoculation. Error bars represent standard deviation of at least 5 biological replicates grown on three separate days.

Regulation of aspartate catabolism in ADP1

Having established that *aspT* and *aspY* encode D-Asp transporters, we wanted to understand how their transcriptional regulation is coordinated with that of genes needed for further catabolism, *racD* and *aspA* (**Figure 3.1**). Transcription of *racD* depends on DarR, which appears to respond to D-Asp (8). We predicted that the DarR homologue, AalR, regulates *aspA* (8, 9). DarR and AalR sequences were aligned in **Figure 3.7.**

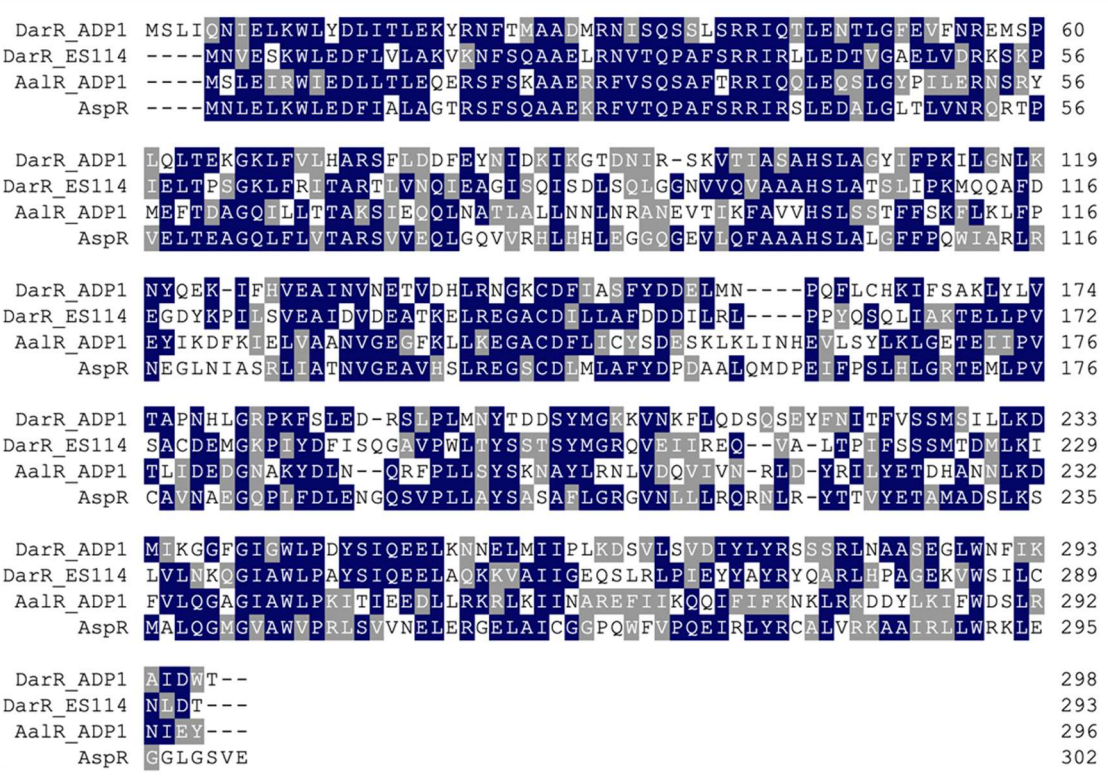


Figure 3.7: Alignment of LTTRs involved in aspartate metabolism

DarR protein sequences from ADP1 (Uniprot ID Q6FF76) and from *Vibrio fischeri* strain ES114 (Uniprot ID Q5E4K6) were aligned with AalR homologs from ADP1 (Uniprot ID Q6FBH7) and from *Pseudomonas aeruginosa* (AspR, Uniprot ID Q9HTD8). If two or more aligned residues

are identical, they are highlighted in blue. Two or more similar residues at the same position in the alignment are highlighted in grey. In a position that has blue highlighting, residues that are similar to those that are identical (in blue) are shown in grey.

When the coding sequence of *aalR* was deleted, the corresponding mutant (ACN1280) failed to grow on L-Asp or D-Asp as the carbon source. Similar results were obtained with another strain in which *aalR* was deleted and replaced with a drug-resistance cassette (ACN1260). These results are consistent with an inability to express AspA without AalR. After extended incubation on Asp plates, some growth was evident (**Figure 3.8**). Asp^+ isolates missing *aalR* were later streak purified and shown to be spontaneous mutants, as described below.

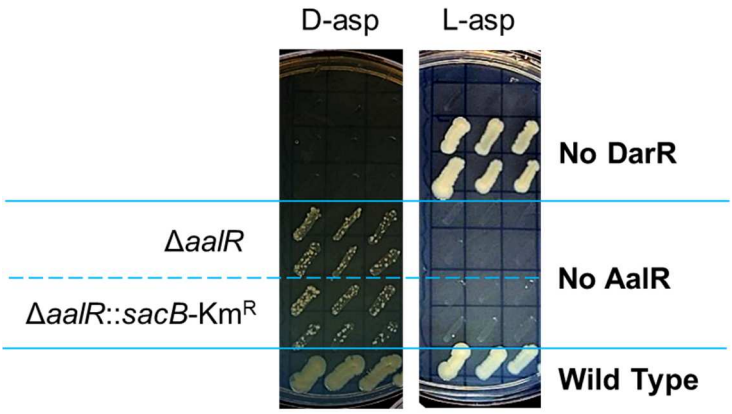


Figure 3.8: Growth of mutants missing *aalR* or *darR* on D-Asp and L-Asp

Colonies were patched to plates with D-Asp (left) or L-Asp (right) as the carbon source. Mutants lacking *darR* or *aalR* grew differently on the two plates. Photos were taken after 5 days of incubation, when some growth was evident within the patches of the *aalR* mutants. Uniform growth in other patches was observed after one day of incubation.

Mutants derived from ACN1260 or ACN1280 were selected on L-Asp, D-Asp, or a mixture of both isomers. Independently isolated spontaneous Asp^+ mutants were characterized further. In one isolate (ACN2180), whole genome sequencing revealed a mutation in *clpA* (ACIAD_RS06285). Based on this result, the *clpA* regions of other *aalR*-deleted Asp^+ mutants were sequenced. In this way, five different *clpA* alleles were identified in D-Asp^+ or L-Asp^+ mutants derived from ACN1260 or ACN1280 (**Table 3.1**).

Table 3.1: *clpA* alleles of spontaneous mutants missing *aalR* that grow on aspartate

Asp^+ mutant	*aalR* allele of parent strain	Carbon source(s) for selection	Position of mutation in 2277-bp coding sequence of *clpA*	Effect on 758- residue ClpA protein product
ACN2180	Δ*aalR51260*	L-Asp	11 nucleotide repeat (nucleotides 1138-1148)	Frame shift after amino acid 383: encodes 622-residue protein
ACN2917	Δ*aalR51260*	L-Asp	G -> T (position 82, creates TAA stop signal)	Truncates protein product after 27 residues
ACN2926	Δ*aalR*::*sacB*-K^R*51260*	Mix of L- and D-Asp	GG insertion (duplication of nucleotides 1600 and 1601)	Frame shift after amino acid 533: encodes 547-residue protein
ACN2927	Δ*aalR*::*sacB*-K^R*51260*	Mix of L- and D-Asp	ΔC (nucleotide 899)	Frame shift after amino acid residue 299: encodes 304-residue protein
ACN2928	Δ*aalR*::*sacB*-K^R*51260*	Mix of L- and D-Asp	T insertion (after nucleotide 560)	Frame shift after residue 190: encodes 194 residue protein

To determine if the *clpA* mutations were sufficient to confer growth on Asp in the absence *aalR*, we used a transformation assay in which a recipient strain (ACN1260 or ACN1280) was spread directly on a plate with D-Asp as the carbon source. Mutated *clpA* DNA when dropped directly on the plate was found to confer growth (one example is shown in **Figure 3.9**). In this example, the truncated allele of ACN2917 was used as donor DNA, suggesting that the loss of *clpA* is responsible for growth. To confirm this inference, two new strains without

aalR were constructed in which *clpA* was specifically inactivated (ACN2246 and ACN2247). These strains were D-Asp$^+$, consistent with the selected mutations causing loss of ClpA function.

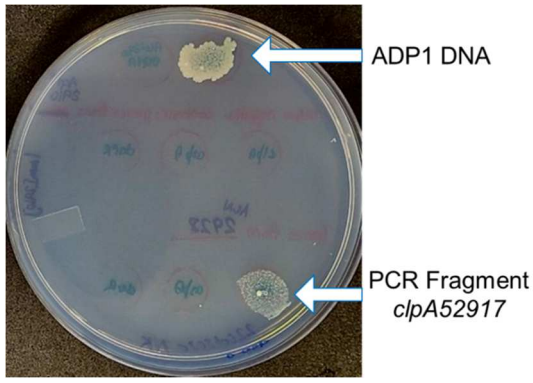


Figure 3.9: *clpA* mutation transforms Δ*aalR* recipient, enabling D-Asp$^+$ growth

Transformation assay with an ACN1280 culture spread on the surface of a D-Asp plate. ADP1 DNA dropped on the bacterial lawn enabled D-Asp$^+$ growth, presumably by allelic replacement that restores *aalR*. A PCR product carrying the mutated *clpA* allele of ACN2917 also enabled D-Asp$^+$ growth. Additional DNA dropped on the plate in different spots failed to confer growth: PCR products carrying DNA of *clpA, darR, aspA* using template DNA from the recipient, ACN1280. PCR products carrying DNA of *darR* and *aspA* from the D-Asp$^+$ *clpA* mutant (ACN2917) were also dropped on as donor DNA and failed to confer D-Asp$^+$ growth by allelic replacement.

Role of ClpA in Asp$^+$ growth of *aalR* mutants

ClpA functions as a chaperone leading to protein degradation mediated by the ClpAP protease (12). Since the *clpA* mutations appear to abolish ClpA mediated proteolysis, we considered whether increased stability of another protein might compensate for the loss of AalR

and considered DarR as a candidate. Although this paralog does not typically enable AalR-independent growth on Asp, the loss of ClpA could prevent DarR degradation. An increase in DarR concentration/stability might allow it to substitute for AalR.

To test this hypothesis, we engineered a Δ*aalR* mutant that is also unable to make DarR. In such a strain we could test whether *clpA* deletion allows growth on Asp. However, since DarR is required to transcribe *racD* and *AspT* (**Figure 3.1**), loss of DarR would prevent growth on D-Asp regardless of a possible role in *aspA* transcription. To address this issue, the *racD* promoter was replaced with a modified *trc* promoter that allows constitutive expression of chromosomal genes in *A. baylyi* (6). Without DarR, this strain (ACN2947) grows on D-Asp, demonstrating that the synthetic promoter allows D-Asp to be converted to L-Asp. When *aalR* was then removed, while leaving the other mutations of ACN2947 in the chromosome, the resulting strain (ACN2954) failed to grow on D- or L-Asp. Transformation assays were conducted using this mutant as the recipient (with no AalR, no DarR, and a constitutive promoter upstream *racD* and *AspT*). In these experiments, *clpA* mutations failed to confer a D-Asp$^+$ phenotype by allelic replacement. These results indicate that mutants can grow on D-Asp without AalR when ClpA is absent, but only if DarR is present.

Transcriptional regulation of *aspA*

Since DarR appears to activate transcription of *aspA* when cells have no AalR or ClpA, we tested whether DarR could substitute for AalR under a different condition. We hypothesized that putting the *darR* coding sequence exactly in place of *aalR* might improve its ability to activate *aspA* transcription. Two strains were constructed to test that possibility (ACN2140 and ACN2141). In both strains, *darR* was removed from its native locus and inserted exactly in place

of the *aalR* coding sequence. The only differences in the sequences were several nucleotides at the 5' region to account for two different potential start codons for the *darR* coding sequence. Both strains failed to grow on Asp. However, a spontaneous D-Asp$^+$ mutant of ACN2140 was isolated and found to have a single nucleotide change (ACN2153). This mutation (designated (P_{aspA}*52153*) was in the 136-bp region between the coding sequences of *aalR* and *aspA*.

To determine the position of this mutation relative to the *aspA* promoter, we determined the transcriptional start site for this gene. RNA was isolated from L-Asp-grown ADP1 cultures, and cDNA was synthesized with an *aspA*-specific primer. The transcriptional start (+1) site was identified by 5' rapid amplification of cDNA ends (5' RACE) (13). This start site was found to be 25-bp upstream of the start of translation. The mutation in the D-Asp$^+$ mutant (ACN2153) lies in the -10 region of the *aspA* promoter and changes it to match the consensus sequence for *E. coli* (ATAAT, **Figure 3.10: Panel A**).

This *aspA* operator-promoter sequence was aligned with comparable regions upstream of *aspA* coding sequences in different *Pseudomonas* and *Acinetobacter* strains (**Figure 3.10: Panel B**). A notably conserved sequence, ATGC-N_7-GCAT, exactly matches the sequence identified as a DarR binding site (8, 14). Consistent with the amino acid sequence similarity in the N-terminal DNA-binding domain of AalR and DarR (**Figure 3.7**), it appears that DarR and AalR recognize the same operator sequence. Electrophoretic mobility shift assays (EMSAs) were used in the MIBO4600L class to demonstrate that purified AalR was able to bind specifically to *aspA* operator-promoter DNA carrying the ATGC-N_7-GCAT sequence. Since these experiments were conducted before my involvement in this project, these data are not included in my thesis.

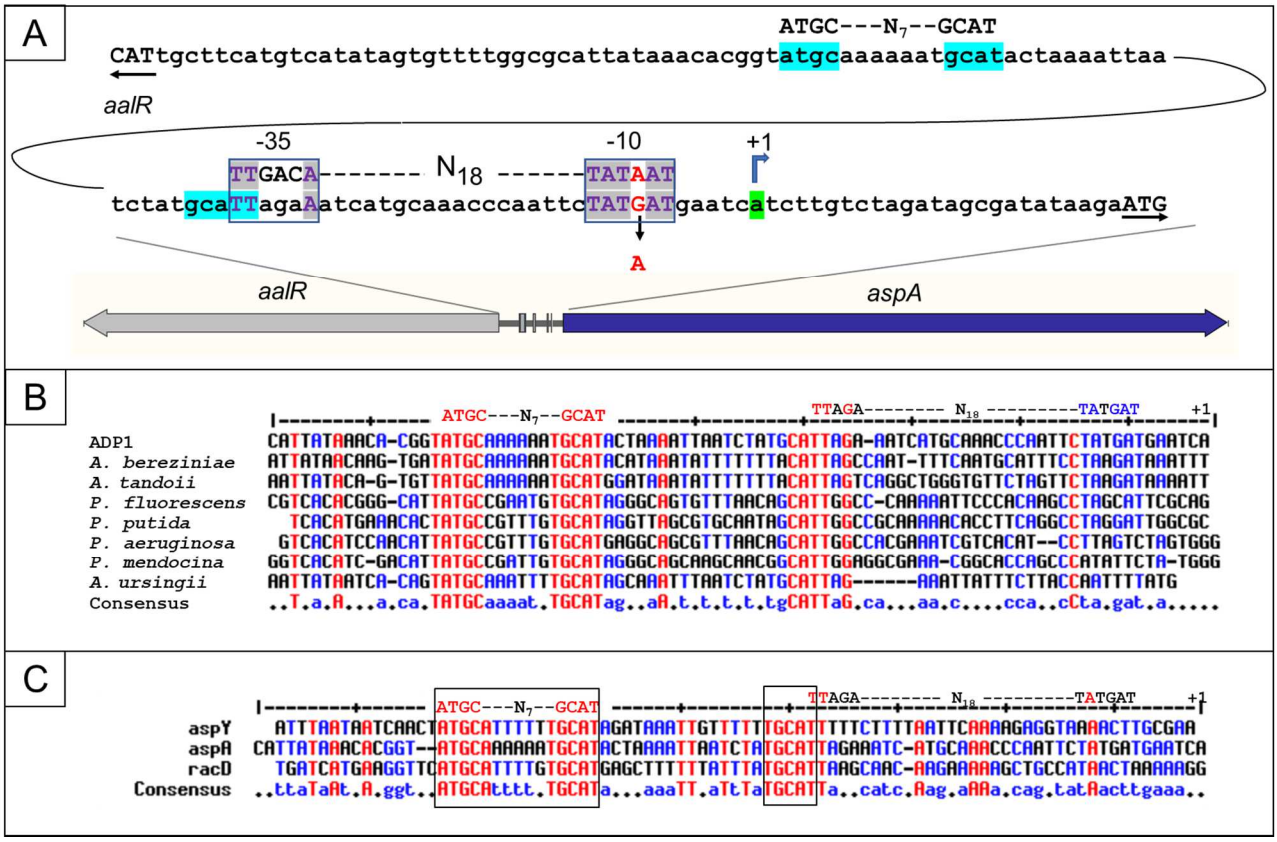


Figure 3.10: Operator-promoter sequences of *aspA*, *aspY*, and *racD*

Panel A: Key features are discussed in the text: the *aspA* transcriptional start site (+1) was determined experimentally; the selected promoter mutation (P_{AspA}*52153)* increases the resemblance to a consensus σ70 *E. coli* promoter, shown above the ADP1 sequence; blue highlights regions predicted to be important for AalR-binding/regulation. Panel B: Alignment of the ADP1 *aspA* promoter region (its +1 site indicated at the right-side) with other regions upstream of predicted a*AspA* genes in different *Acinetobacter* and *Pseudomonas* strains. Conserved sequences are marked in red or blue, depending on how many sequences are identical. Many of the red sequences correspond to the teal highlighted regions in Panel A. The sequence of the ADP1 promoter is shown above the alignment. Panel C: The *aspA* sequence from panel B is aligned with regions upstream of *aspY* and *racD* in ADP1. The right-most nucleotide shown for *aspY* (top line) is the experimentally determined transcriptional start site for this gene (+1).

Mutation allows DarR to regulate *aspA* transcription

To determine the effect of the promoter mutation ($P_{aspA}52153$), we used a chromosomal transcriptional fusion (*aspA*::*gfp*). The disruption of *aspA* allows Asp to be used as a non-metabolizable effector. Fluorescence for a strain with the transcriptional fusion and no *aalR* (ACN1895) was comparable with or without D-Asp added to the medium. This basal level was also comparable to a strain with no transcriptional fusion (ACN1280). Therefore, data are displayed for strains with different genetic backgrounds relative to this basal level, **Figure 3.11**.

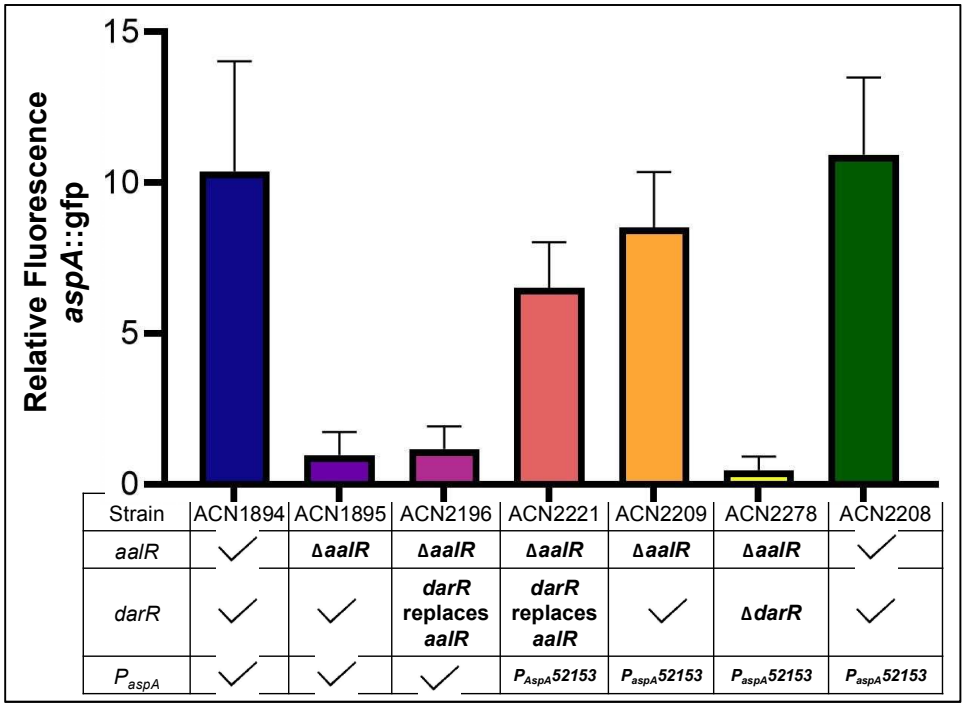


Figure 3.11: Relative fluorescence of strains with an *aspA*::gfp transcriptional fusion grown with D-Asp added as an inducer

All strains have a chromosomal transcriptional fusion controlled by an *aspA* promoter (P_{AspA}). Differences in genetic backgrounds are indicated for each strain in the table. A checkmark indicates the wild-type allele for *aalR*, *darR,* or $P_{aspA.}$ Fluorescence is shown per OD_{595} relative to

that of ACN1895 (average florescence normalized to OD_{595} = 10). Error bars represent standard deviation of at least 6 biological replicates done on three days. All strains were grown on pyruvate with D-Asp added as an inducer.

The *aspA* promoter mutation was selected in a strain in which *darR* replaced *aalR*. In this set of experiments, we focused on regulation in the presence of D-Asp, which can be converted into L-Asp, as an inducer. When *aalR* was present, fluorescence in response to D-Asp was more than 10-fold higher than when *aalR* was deleted (comparison of ACN1894 and ACN1895), consistent with AalR serving as the transcriptional activator of *aspA*. Moving the chromosomal location of *darR* to replace *aalR* did not enable D-Asp mediated regulation of GFP from the wild-type P_{aspA} (comparison of ACN2196 to ACN1895, **Figure 3.11**). However, the promoter mutation (P_{aspA}*52153)* increased transcription in the presence of D-Asp when *darR* was in its native location and when it was in the native *aspA* locus (ACN2221 and ACN2209, **Figure 3.11**). Transcription from the mutated promoter was not constitutive, but rather depended on DarR, as evidenced by low level gene expression when *darR* was deleted (ACN2278, **Figure 3.11**). The promoter mutation did not appear to affect AalR-mediated transcription, since strains with P_{aspA} or P_{aspA}*52153* resulted in nearly identical fluorescence when *aalR* was present (comparison of ACN1984 and ACN2208, **Figure 3.11**). P_{aspA}*52153* specifically improved the ability of DarR to regulate *aspA*.

Transcriptional regulation of *AspY*

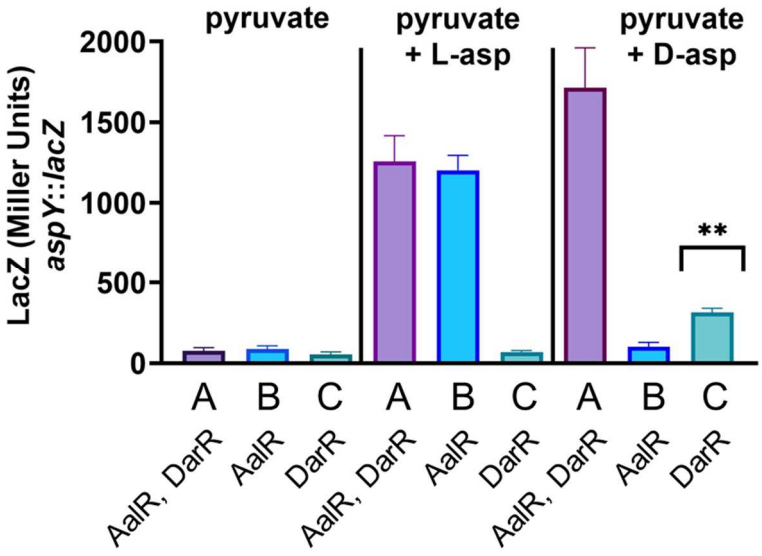
The observations that DarR and AalR both recognize the same LTTR binding site (ATGC-N_7-GCAT) and the discovery that under some circumstances DarR can substitute for

AalR, raised questions about *aspY* regulation. An alignment of the promoter regions of AalR-regulated *aspA* and DarR-regulated *racD* showed significant similarity to *aspY* regulatory sequences (**Figure 3.10: Panel C**). Using the same approach as was used for *aspA*, we determined the transcriptional start site of *aspY* by 5' RACE (13). With this method, the transcriptional start site for *aspY* was determined to correspond to the nucleotide shown at the 3' end of the *aspY* sequence shown in **Figure 3.10: Panel C**.

Since AspY appears to transport both D-Asp and L-Asp, it wasn't clear whether DarR, AalR, or both proteins regulate this gene. An *aspY*::*lacZ* fusion was integrated in the chromosome as a transcriptional reporter. Strains were constructed with this chromosomal *aspY*::*lacZ* construct in genetic backgrounds that maintained AalR alone (ACN2984), DarR alone (ACN2989), or both regulators (ACN2144). When investigating the role of L-Asp or D-Asp as effectors of AalR or DarR we wanted to prevent Asp interconversion by the RacD isomerase. Therefore, in the strains lacking *darR* (ACN2984) or *aalR* (ACN2989), *racD* was also deleted. To enable *aspT* to be transcribed constitutively, the *trc* promoter (described earlier) was used to replace the regulated promoter normally controlling the *racD* and *aspT*.

As shown in Figure **3.12**, when both DarR and AalR can be expressed, transcription from the *aspY* promoter increases in response to L-Asp or D-Asp (ACN2144, strain A). In contrast, when only AalR can be expressed, L-Asp, but not D-Asp, serves as an effector to increase transcription (ACN2984, strain B). This result suggests that AalR specifically responds to only the L-Asp isomer. In contrast, L-Asp does not induce increased expression when the DarR, but not AalR, is present (ACN2989, strain C). When DarR is present, D-Asp can induce a small but statistically significant increase in LacZ expression from the *aspY* promoter. These results

indicate that both AalR and DarR regulate *aspY* transcription and that each regulator responds only to one isomer of aspartate.



Strain:	A ACN2144	B ACN2984	C ACN2989
aalR	✓	✓	Δ***aalR***
darR	✓	Δ***darR***	✓
P_{racD}	✓	$\boldsymbol{P_{racD}}$ ***trc52947***	$\boldsymbol{P_{racD}}$ ***trc52947***

Figure 3.12: LacZ expression from the *aspY* promoter

All strains have a chromosomal transcriptional fusion controlled by the *aspY* promoter. Differences in genetic backgrounds are indicated for each strain in the table. A checkmark indicates the wild-type allele for *aalR*, *darR,* or the *racD* promoter $P_{racD.}$ (15). All strains were grown on pyruvate. D-Asp or L-Asp was added as inducer where noted. ** indicates significance in un-paired T-test of ACN2989 to ACN2984 grown on pyruvate + D-Asp (p-value > 0.0001)

L- and D-amino acids as carbon and nitrogen sources

In additional studies, we compared the ability of *A. baylyi* ADP1 to consume aspartate to its ability to utilize two related amino acids. We tested whether *A. baylyi* and various mutants use L- and D- isomers of aspartate, glutamate, or asparagine as either the sole carbon source or the sole nitrogen source. In the patching experiment shown in **Figure 3.13**, the colonies tested for growth on L-Asp and D-Asp were also tested for growth on each enantiomer of the other three amino acids as the sole carbon source. *A. baylyi* ADP1 failed to grow on D-Glu as the carbon source. The other amino acids tested (L-Asp, D-Asp, L-Asn, D-Asn, and L-Glu), could each serve as the carbon source. Results with the transport mutants concerning growth on L-Asp, L-Asn, and L-Glu were similar. Moreover, these strains grew similarly on D-Asp and D-Asn, However, in patches of mutants lacking AspY, there were some differences in the ability to grow on different amino acids. For example, the strain missing AspT, AspY, and AspS did not grow on L-Asn or L-Glu, but it grew slowly on L-Asp (region 8 of the plates). These differences may reflect different substrate specificities of the transporters.

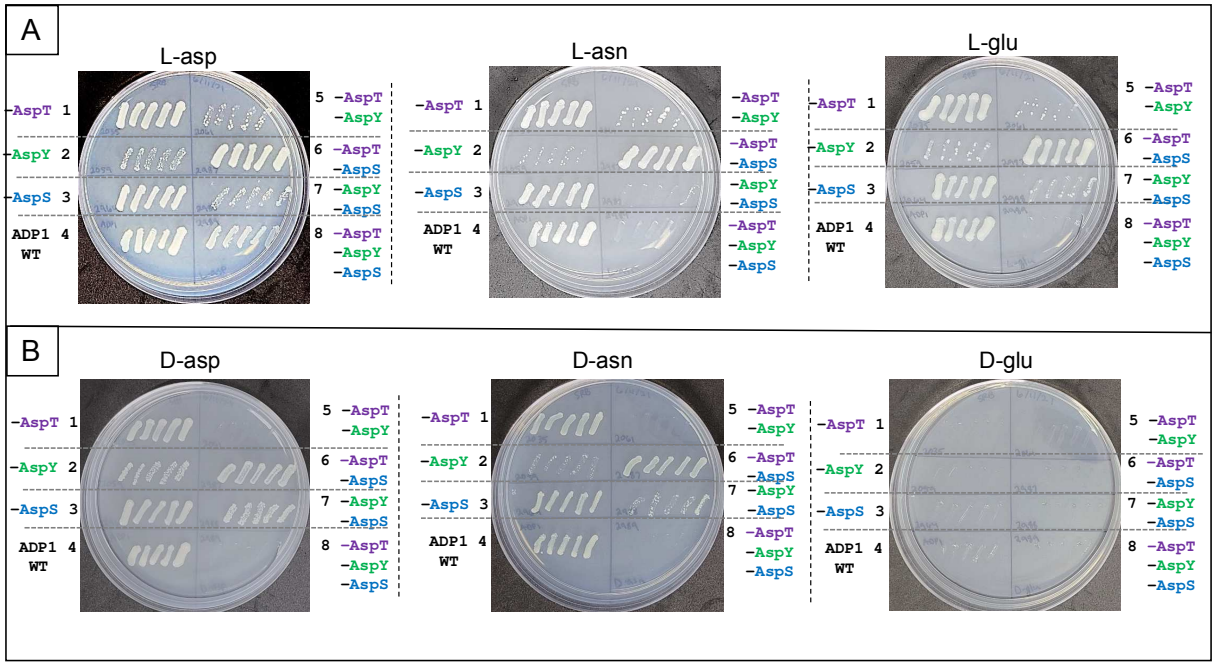


Figure 3.13: Growth of wild type and transport mutants on L- and D-isomers of Asp, Glu, or Asn as the sole carbon source

Colonies were patched as described in **Figure 3.2.** Images shown were taken after 4 days of incubation. Strains are: ACN2035 (no AspT) in region 1, ACN2059 (no AspY) in region 2, ACN2964 (no AspS) in region 3, ADP1 (wild type) in region 4, ACN2061 (no AspT or AspY) in region 5, ACN2987 (no AspT or AspS) in region 6, ACN2988 (no AspY or AspS) in region 7, and ACN2967 (no AspT, AspY, or AspS) in region 8

To assess the ability to use these three amino acids as sole nitrogen sources, the wild-type strain (ADP1) and a mutant lacking AspA (ACN1894) were grown in medium with the indicated compound as sole nitrogen source. *A. baylyi* could use each amino acid tested, except D-Glu, as the sole nitrogen source. Although the loss of AspA in ACN1894 prevents the use of Asp and Asn as sole carbon sources, AspA was not required for Asp, Asn, or L-Glu to serve as the sole nitrogen source.

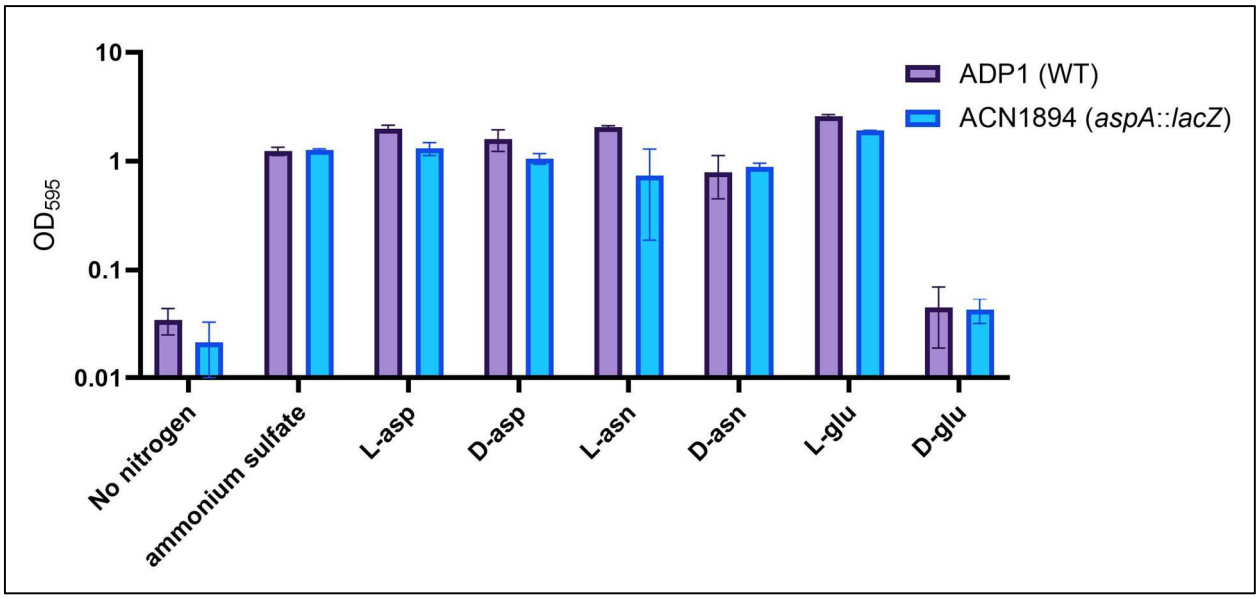


Figure 3.14: OD_{595} of *A. baylyi* strains grown with indicated sole nitrogen source

OD_{595} was measured at 24 hours post inoculation of strains ADP1 (WT), purple bars, and ACN1894 (*aspA*::*lacZ*), blue bars. The nitrogen source provided is indicated on the x-axis. OD_{595} is graphed on a $\log_{10}$ scale. Error bars represent standard deviation of six biological replicates measured on three separate days.

Discussion

Despite increased interest in the metabolism of D-amino acids, the regulated bacterial uptake and consumption of such compounds remains largely unexplored. Our studies show that two paralogous transcriptional regulators in ADP1, DarR and AalR, regulate the transcription of multiple genes including those encoding aspartate ammonia lyase (AspA), aspartate racemase (RacD), and three aspartate transporters (AspT, AspY, and AspS). These transporters belong to a set of seven paralogs of glutamate/aspartate proton symporters in strain ADP1. This type of

transporter is common in bacteria, archaea, and in eukaryotes (15). Mammalian glutamate transporters are the subject of intensive study because of their roles in neuronal signaling (16, 17). As the importance of D-Asp as a neurotransmitter has become appreciated, interest has also increased in how glutamate transporters recognize D-Asp and how they can bind and act on isomeric forms of amino acids. A recent structural study of an archaeal homolog of these seven transporters in ADP1 addresses some of the questions (18). The focus of this study, Glt_{Tk} from *Thermococcus kodakarensis*, is a protein that transports D-Asp and L-Asp with similar affinities for both substrates.

While members of this family have long been known to exist in bacteria, the many paralogs and differing growth capabilities of diverse bacteria precludes conclusions about function to be drawn based strictly on homology (19). Thus, no studies have previously demonstrated the physiological relevance of specific transport proteins for bacterial consumption of D-aspartate. Using simple genetic approaches, appropriate for an undergraduate teaching lab, we were able to demonstrate that two proteins, designated AspY and AspT, are responsible for the ability of ADP1 to take up D-Asp for use as a sole carbon source. Consistent with this conclusion, each protein from ADP1 conferred a D-Asp^+ phenotype to *E. coli* when expressed together with the aspartate racemase. Moreover, the genes encoding these transporters were shown to be regulated in response to D-Asp as an effector for the DarR regulator.

Overlapping functions of AspY, AspT, and AspS

The large number of paralogs of these transporters suggests the importance of amino acids as nutrient sources in the natural environment where *A. baylyi* is typically found. While ADP1 encodes seven such proteins, *E. coli* appears to encode four of them. The AspS protein,

which at high expression levels can confer D-Asp$^+$ growth in *A. baylyi* when AspT and AspY are absent did not enable *E. coli* to grow on D-Asp. It is possible that this protein was not expressed in *E. coli* at sufficiently high levels. Alternatively, it could be that the heterologous host did not properly produce or localize the protein. We did not investigate expression levels of the proteins in *E. coli*.

The patching studies indicated that despite some redundancy, the effects of deleting *aspY*, *aspT*, and/or *aspS* were not identical for impacting growth on L-Asp and D-Asp (**Figures 3.2** and **3.12**). Consistent with the localization of *aspT* with *racD* and its transcriptional regulation by DarR, loss of AspT was more detrimental when D-Asp was the carbon source than when L-Asp was the carbon source. Thus, AspT seems to be relatively specific for D-Asp.

The transport mutants had similar growth patterns on L-Asp, L-Glu, and L-Asn. Similar results on L-Asp and L-Glu may reflect the known overlap in specificity of glutamate/aspartate transport proteins (19). For Asn, the similarity in effects of the transport mutants on growth for L-Asn and L-Asp may indicate that Asn is converted to Asp in the periplasm so that transport across the cell membrane would be mediated by an Asp transporter. There are at least two asparaginases encoded by ADP1 (UniProt IDS Q6FEV1 and Q6FAL6). Some asparaginases are periplasmic, and in some bacteria the genes for Asp and Asn catabolism are linked (20, 21).

For the catabolism of D-Asn, we could not identify a racemase by bioinformatic analysis that seemed a likely candidate for the formation of L-Asn. However, some asparaginases act on D-Asn (22). If ADP1 degrades D-Asn via conversion to Asp, then the similar growth patterns of the mutants on D-Asp and D-Asn might be explained. Although D-Glu was not used as a sole carbon source, it seems unlikely, but possible, that transport is responsible. ADP1 has a D-Glu

racemase, MurI. Further work is needed to determine whether, transport, MurI activity, or issues with regulated protein expression prevent D-Glu consumption.

In comparing patch plates of the mutants between Asp and the other amino acids, the loss of AspY had a bigger impact on growth with D-Asn as the carbon source than with D-Asp (region 2). When growth on the L-amino acids was compared, the patch 2 region of growth was also stronger on L-Asp than the others. Notably, there was growth on L-Asp for the mutant with all three transporters missing (region 8). Additional research is needed to determine the significance of these observations.

Regulatory specificity of DarR and AalR

In comparison to some of the transport proteins that appear to use both L-Asp and D-Asp as substrates, DarR and AalR appear to distinguish L-Asp from D-Asp. This discrimination has not previously been demonstrated. Since both DarR and AalR regulate the transcription of *aspY*, the transcriptional fusion at this locus illustrated the difference in effector specificity (**Figure 3.12**). Although DarR can activate transcription from this promoter, it does not do so in response to L-Asp, as demonstrated in a strain missing the racemase. Conversely, AalR did not respond to D-Asp in the absence of RacD.

The sequence alignments of DarR and AalR demonstrate that their DNA-binding domains (at the N-terminus) are nearly identical, consistent with their ability to recognize the same operator sequence. There are greater differences in the sequences of the effector binding domains, which could explain their ability to discriminate L-Asp and D-Asp as effectors. A recent structural study of DarR from *V. fischeri* (14) may lay the foundation for further examination of potential differences in these regulators (DarR and AalR).

The transcriptional start sites were determined by 5' RACE for *aspA* and *aspY*. An alignment of these regions with the *racD* promoter highlighted striking similarity. This alignment (**Figure 3.10**) highlights conserved features. The typical dyad symmetry of the LTTR binding site is sufficient to bind a DNA-binding domain from each of two subunits. Since LTTRs typically function as a tetramer, it is not clear how the oligomer would bind and activate transcription. The sequence alignment highlights a conserved sequence downstream of the ATGCA-N5-TGCAT binding site that matches a "half-site" of this symmetry: TGCAT. These conserved regions are boxed in **Panel C** of **Figure 3.10**. A DNA-binding domain of a tetramer could bind to this TGCAT sequence which overlaps the -35 region of the promoter. Thus, the DarR or AalR could interact with RNA polymerase, perhaps with the sigma factor, as in a Class II promoter (23, 24).

Nitrogen metabolism in *A. baylyi* ADP1

Studies here showed that while the utilization of Asp as a carbon source in *A. baylyi* is dependent on aspartate-ammonia lyase (AspA). AspA catalyzes the cleavage of L-aspartate into ammonium and fumarate, which enters central carbon metabolism. However, *A. baylyi*'s use of Asp as a nitrogen source is not dependent on AspA. Both wild type *A. baylyi* and an *aspA* mutant, ACN1894, used L-Asp, D-Asp, L-Asn, D-Asn, and L-Glu as the sole nitrogen source. Because the *aspA* mutant cannot be using AspA cleavage of aspartate for nitrogen assimilation in this strain, an alternative pathway for L-aspartate as a nitrogen source must be used. L-aspartate aminotransferase, *aspC* (ACIAD_RS11695), catalyzes the conversion of L-aspartate and 2-oxo-glutarate to L-glutamate and oxaloacetate (25). It is likely that *A. baylyi* utilizes this pathway for nitrogen assimilation as it generates L-glutamate, which along with L-glutamine, represent the

primary nitrogen donors for biosynthetic pathways in bacteria (26). This finding stands in contrast to recent studies of *E. coli* nitrogen assimilation that determined that nitrogen assimilation in *E. coli* from Asp is dependent on the Aspartate-ammonia lyase (26).

Metabolic and genetic studies for undergraduate laboratory courses

In designing experiments for a semester long microbiology laboratory course for undergraduate students, it is important to choose projects and experiments that can be done and analyzed in this short time frame. Because *A. baylyi* has such a facile genetic system, genetic changes can be quickly made and/or analyzed through basic techniques. For example, the transformation assays with linear DNA dropped onto a selective plate with a lawn of cells quickly and clearly demonstrates bacterial transformation to students.

During my doctoral studies, I was the TA for multiple semesters of MIBO4600L (Experimental Microbiology) and MIBO2500L (Microbiology and Health Care) as well as an instructor of record for Microbiology and Health Care. In both classes, I used *A. baylyi*'s powerful natural transformation system to teach students about modes of horizontal gene transfer. In Microbiology and Health Care, experiments taught future healthcare professionals about the importance of antibiotic resistance. Students preformed a natural transformation assay to see how a Streptomycin resistant strains of ADP1 can confer the resistance to a Streptomycin sensitive strain. Antibiotic resistance is a major concern in our modern healthcare system as more and more bacteria become resistant to our modes of treatment, and it is important that students going into healthcare professions understand how antibiotic resistance can be spread by.

In Experimental Microbiology, students are able to do real, publishable research in *A. baylyi*. One article resulting from this course was published in 2017 about another LTTR in *A.*

baylyi, MdcR (27). Each class was able to make new and interesting discoveries during their semester that have been incorporated into the research described here. Without the ease of genetic manipulation provided by *A. baylyi*, these studies would likely not be accomplished in such a short timeframe. We plan to publish this work in the Journal of Bacteriology.

Experimental procedures

Media and growth conditions

Defined ADP1 medium was used for culturing all *A. baylyi* strains and *E. coli* DH5α for heterologous expression experiments as previously described (28). This medium was supplemented with 20 mM pyruvate, L-Asp, D-Asp, L-Asn, D-Asn, L-Glu, and/or D-Glu unless otherwise indicated. For experiments using the amino acids as nitrogen sources, media contained 20 mM pyruvate for a carbon source and 5 mM of one amino acid as a carbon and nitrogen source. For plasmid cloning, E. coli XLI Blue or DH5α (Stratagene) was grown in Luria-Bertani medium (LB) (29). All cultures were incubated at 37 °C with shaking at 250 rpm unless otherwise indicated. Antibiotics were used at final concentrations of 25 μg/ml for kanamycin, 12.5 μg/ml each for streptomycin and spectinomycin, and 150 μg/ml for ampicillin.

Strain and plasmid construction

DNA used for plasmid constructed or allelic replacement experiments was PCR-amplified by using the high-fidelity polymerases PrimeSTAR Max (Takara Biosciences) or Phusion (New England Biosciences). All primers used can be found in the primer list (**Table 3.4**). Plasmids were constructed by overlapping sequence assembly in E. coli XL1-Blue or DH5α competent cells (Stratagene) as previously described (30). Standard restriction digest followed by

ligation (Quick Ligation Kit, New England Biosciences) was also used for plasmid construction. Plasmids were confirmed by DNA restriction mapping and/or regional DNA sequencing. Plasmids are described further in the supplemental plasmid list (**Table 3.3**).

All chromosomal changes in *A. baylyi* ADP1 were constructed by allelic replacement as previously described (31). Strains were confirmed by PCR and/or regional DNA sequencing. All strains are further described in the strain table (**Table 3.2**).

Patching of *A. baylyi* strains

For patching experiments, a single colony from each strain was taken with a toothpick from a non-selective plate of minimal media with 20 mM pyruvate. It was struck in a centimeter long line to the same spot on each plate used in the experiment with the last plate being a non-selective plate to ensure that enough cells were deposited onto each previous plate. All plates were incubated at 37 °C overnight. 5 individual replicates were done for each strain. Then, the plates were incubated at room temperature to prevent the plates from drying out too fast but to let the cells continue to grow for up to 5 days.

Computational search for aspartate transporter homologues

NCBI's Basic Local Alignment Search Tool, BLAST, was used to search the A. baylyi genome for homologs to the predicted Aspartate transporter, AspT (32). Six paralogues were found with 24% or greater similarity to AspT and were investigated further.

Selection D-Asp$^+$ spontaneous mutant without AspY or AspT (ACN2921)

The *A. baylyi* strain missing AspY and AspT, ACN2061, was grown overnight on 5 ml minimal media with 20 mM pyruvate and 20 mM L-Asp, D-Asp, L-Asn, D-Asn, L-Glu, or D-Glu. The 5 ml cultures were then pelleted at 5 x g for 3 min. Cells were washed with 1 ml minimal media and pelleted again. Cells were resuspended in 100 µl minimal media, plated to minimal media with 20 mM D-Asp and incubated overnight.

A. baylyi transformation assays

AspT paralogs were PCR amplified from ACN2061, strain missing AspY and AspT, or ACN2921, strain missing AspY and AspT selected to grow on D-Asp (ACN2921) and cleaned and concentrated using Monarch PCR & DNA Cleanup Kit (New England BioLabs Inc. # T1030S). An overnight culture of ACN2061 was spiked with 100 ul 20 mM pyruvate and incubated shaking at 37 °C 250 rpm for another 30 min to bring the cultures into log growth phase. 100 ul of culture was plated on minimal media with 20 mM L-Asp or D-Asp and then 100 µg of each PCR product was dropped top of cells and incubated overnight.

clpA alleles were PCR amplified from spontaneous mutants growing on Aspartate without AalR and cleaned and concentrated using Monarch PCR & DNA Cleanup Kit (New England BioLabs Inc. # T1030S). An overnight culture of ACN1260 or ACN1280 was spiked with 100 ul 20 mM pyruvate and incubated shaking at 37 °C 250 rpm for another 30 min to bring the cultures into log growth phase. 100 ul of culture was plated on minimal media with 20 mM L-Asp or D-Asp and then 100 µg of each PCR product was dropped top of cells and incubated overnight.

β-Galactosidase assays to assess *AspY::lacZ* and *AspS::lacZ* transcriptional fusions

Either an *AspY::lacZ* or an *AspS::lacZ* fusion was introduced into the *A. baylyi* chromosome. Liquid cultures were grown overnight on minimal media with either pyruvate, pyruvate and L-Asp, or pyruvate and D-Asp. Cells were diluted between 1 in 10 and 1 in 40 with Z-buffer (16.1 g Na2HPO4·7H2O, 5.5 g NaH2PO4·H2O, 0.75 g KCl, and 0.246 g MgSO4·7H2O in 1 L H2O with 27 µl β-mercaptoethanol added fresh to each 10 ml Z-buffer used) and a final concentration of 0.001% SDS and chloroform were added. The reaction, at 37 °C, was started with the addition of 250 µl o-nitrophenyl-β-D-galactoside (ONPG) for every 1 ml of cell mixture (4 mg/ml ONPG in 0.1 M potassium phosphate buffer, pH 7.0). The reaction was stopped when a yellow color appeared with 500 µl stop buffer (106 gm Na2CO3 in 1 L H2O). Two replicates were done for each sample with different reactions times to ensure the enzyme reaction was taking place in the linear range. Samples were pelleted at 5 x g for 3 min and the supernatant was measured at 420 nm. Each culture was also diluted 1 in 5 in Z-buffer and measured on the spectrophotometer at 650 nm. Miller Units were calculated as Miller units = (1000 × A420) / (t × v × A650). Each strain was done in at least 4 replicates over 2 days.

Heterologous expression of *A. baylyi* proteins in *E. coli*

Plasmids, listed in **Table 3.3**, expressing *A. baylyi* RacD, AspT, or RacD plus one of the 7 Aspartate transporter homologs were introduced into *E. coli* DH5α. Cells were then grown on minimal media with kanamycin, for plasmid retention, and either 20 mM L-Asp or D-Asp overnight. Cell density was measured after 72 hours at 595 nm on spec. Each strain was done in at least 6 replicates over 3 days.

Isolation of L-Asp$^+$ mutants of *A. baylyi* strains missing AalR

Spontaneous mutants were selected for ACN1280 and ACN1260, strain missing *aalR*, by growth on L-Asp as the sole carbon source. Multiple mutations in the *clpA* coding sequence were discovered by next generation sequencing and regional Sanger sequencing, Eton Biosciences. A mutation in the *aspA* promoter region was located via regional Sanger sequencing, Eton Biosciences.

Next generation sequencing and analysis

Genomic DNA was isolated from ADP1 derived strains and sonicated to approximately 500 bp size fragments in a total of 500 ng. Genomic libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) including end-repair, adaptor ligation, and addition of an individual i5/i7 index primer. Sequencing was performed on the Illumina NextSeq500 instrument at the Georgia Genomics and Bioinformatic Core at the University of Georgia.

Sequence analysis was performed using Geneious software with default settings. A reference genome was constructed by mapping reads from ADP1 to GenBank CR543861. Paired-end reads were mapped to this refence genome and a minimum variant frequency of 0.8 was used to call variants and identify single nucleotide polymorphisms.

Transcript analysis

A. baylyi cell cultures were harvested while in late logarithmic phase and entering stationary. RNA was preserved using RNAprotect (Qiagen) and isolated using the RNAeasy kit (Qiagen). RNA was further treated with RQ1 DNase (Promega) and confirmed to be free of DNA through PCR using LongAmp Taq (NEB) and assessment via agarose gel electrophoresis.

The 5' RACE System for Rapid Amplification of cDNA Ends (Invitrogen) was used to determine the 5' end of transcripts. cDNA for *aspA* was generated using primer ALS60. Primers ALS61 and ALS62 were used for internally amplification of 5' RACE products which were then cloned into pCR4-TOPO vector (Invitrogen) and sequenced. For generation of *AspY* cDNA, primer LES9 was used, with primers LES10 and LES11 were used for internal amplification. For generation of *aalR* cDNA, primer ALS114 was used. Primers ALS66 and ALS115 were used for internal amplification of *aalR* cDNA. For generation of *racD* cDNA, primer LES12 was used. Primer used for internal amplification were either ALS18 or ALS129, and ALS123
The transcript of *AspA* was evaluated with two independently isolated RNA samples from ADP1 grown on L-Asp. Both *AspY* and *racD* transcripts were evaluated from at least two RNA samples from ADP1 grown on D-Asp. The *aalR* transcripts was evaluated with three independently isolated RNA samples from ACN2029 grown on L-Asp.

Fluorescence assays to assess *AspA::gfp* transcriptional fusions

Liquid cultures were grown on minimal media with pyruvate with and without 5 mM L-Asp or D-Asp, inoculated 1:10 from cultures with no Aspartate. Cell cultures were assayed at an OD600 of 0.86-1.2. Green fluorescence (excitation 485, emission 528) and OD600 were measured using a Synergy 2 plate reader (Biotek). Fluorescence measures are reported as green fluorescence relative to OD600. Relative fluorescence was then normalized to an *A. baylyi* strain without *gfp*.

Table 3.2: Strains used in Chapter 3 study

Strain	Relevant Characteristics	Source
ADP1	Wild-type strain (BD413)	(33)
ACN1260	Δ*aalR*::*sacB*-K^R*51260* *sacB*-K^R cassette from pRMJ1 replacing *aalR* (**1,753,790 - 1,754,674**) pBAC1051 XmnI X ADP1	This study
ACN1261	Δ*mdcR*::*sacB*-K^R*51261* *sacB*-K^R*52161* cassette from pRMJ1 replacing *mdcR* pBAC1052 X ADP1	(27)
ACN1280	Δ*aalR51280* Deletion of *aalR* inside of start and stop codons (**1,753,790 - 1,754,674**) pBAC1032 EcoRI X ACN1260	This study
ACN1720	Δ*racD*::*sacB*-K^R*51720* *sacB*-K^R cassette replacing *racD* (**321,876 - 324,768**) pBAC1289 SphI/SacI X ADP1	(8)
ACN1721	Δ*darR*::*sacB*-K^R*51721* *sacB*-K^R cassette replacing *darR* (**323,700 – 322,810**) pBAC1297 BglI/ BspQI X ADP1	(8)
ACN1726	Δ*darR51726* XhoI recognition site replacing *darR* (**323,700 – 322,810**) inside of start and stop codons pBAC1296 SapI/ SacI X ACN1721	(8)
ACN1862	Δ*mdcR*::*mCherry51862* *mCherry* interrupting *mdcR* pBAC1425 PvuI x ACN1261	This study
ACN1864	Δ*mdcR*::*mCherry51862;* Δ*aalR*::*sacB*-K^R*51260* ACN1260 lysate x ACN1862	This study
ACN1876	Δ*mdcR*::*mCherry51862*; Δ*aalR51280* ACN1280 lysate X ACN1864	This study
ACN1883	Δ*mdcR*::*mCherry51862*; *AspA*::*gfp*-ΩK^R*51883* *gfp*-ΩKm^R*51883* interrupting *AspA* pBAC1468 XmnI X ACN1862	This study
ACN1884	Δ*mdcR*::*mCherry51862*; *AspA*::*gfp*-ΩK^R*51883;* Δ*aalR51280* pBAC1468 PsiI X ACN1876	This study
ACN1894	*AspA*::*gfp*-ΩK^R*51883* ADP1 lysate X ACN1883	This study
ACN1895	Δ*aalR51280*; *AspA*::*gfp*-ΩK^R*51894* ADP1 lysate X ACN1884	This study
ACN1924	Δ*aalR*::*sacB*-K^R*51260*; Δ*darR51726* ACN1260 lysate x ACN1726	This study
ACN1964	Δ*darR51726*; Δ*racD*::*sacB*-K^R*51720* pBAC1289 SphI/SacI X ACN1726	This study
ACN2035	Δ*AspT52035* Deletion of *AspT* (**320,208 – 321,532**) replaced with XhoI recognition site pBAC1541 PvuII X ACN1720	This study
ACN2059	*AspY*::ΩS^R*52059* Interruption of *AspY* with ΩS^R pBAC1569 AatII/SapI X ADP1	This study
ACN2061	Δ*AspT52035; AspY*::ΩS^R*52059* pBAC1569 AatII/SapI X ACN2035	This study

ACN2140	Δ*aalR*::*darR52140*; Δ*darR51726* *darR* (**322,807 - 323,697**) replacing *aalR* (**1,753,787 – 1,754,677**) pBAC1510 BamHI/EcoRV X ACN1924	This study
ACN2141	Δ*aalR*::*darR52141*; Δ*darR51726* Longer *darR* (**322,807 - 323,703**) replacing *aalR* (**1,753,787 – 1,754,677**) pBAC1607 BamHI/EcoRV X ACN1924	This study
ACN2144	*AspY*::*lacZ*-K^R*52144* *lacZ*-K^R*52144* inserted into *AspY* at XhoI recognition site in same orientation. Expression should be controlled by *AspY* promoter pBAC1611 AatII X ADP1	This study
ACN2152	Δ*aalR*::*darR52140*; Δ*darR51726* ACN2140 spontaneous mutant using 1 mM D-Aspartate as sole carbon source	This study
ACN2153	Δ*aalR*::*darR52141*; *AspA52153*; Δ*darR51726* spontaneous mutations in the promoter region of *AspA* (C -> T ; **1,753,683**) ACN2141 spontaneous mutant using 1 mM D-Aspartate as sole carbon source	This study
ACN2179	*AspY*::*lacZ*-K^R*52179* *lacZ*-K^R*52179* inserted into *AspY* at XhoI recognition site in opposite orientation as *AspY*. Expression should not be controlled by *AspY* promoter pBAC1612 AatII X ADP1	This study
ACN2180	Δ*aalR51260*; *clpA52180* 11 nt repeat in *clpA* (**1,359,239 – 1,359,249**) ACN1260-derived L-Asp+ spontaneous mutant	This study
ACN2195	Δ*aalR51280*; *AspA52153* pBAC1637 AatII/NdeI X ACN1260	This study
ACN2196	Δ*aalR*::*darR52140*; *AspA*::*gfp*-ΩK^R*51883; Δdar51726* pBAC1468 PsiI X ACN2140	This study
ACN2197	Δ*aalR*::*darR52141; AspA*::*gfp*-ΩK^R*51894; Δdar51726* pBAC1468 PsiI X ACN2141	This study
ACN2208	*AspA52153; AspA*::*gfp*-ΩK^R*51894* pBAC1638 PmeI/AatII X ADP1	This study
ACN2209	Δ*aalR1280*; *AspA52153*; *AspA*::*gfp*-ΩK^R*51894* pBAC1638 PmeI/AatII X ACN2195	This study
ACN2220	Δ*aalR51280*; *AspA52153; Δdar51726* pBAC1637 AatII/SapI X ACN1924	This study
ACN2221	Δ*aalR*::*darR51943*; *AspA52153; AspA*::*gfp*-ΩK^R*51894; Δdar51726* pBAC1645 AatII X ACN2140	This study
ACN2222	Δ*aalR*::*darR52141*; *AspA52153; AspA*::*gfp*-ΩK^R*51894; Δdar51726* pBAC1646 AatII X ACN2141	This study
ACN2246	Δ*aalR*::*sacB*-K^R*51260; clpA*::ΩS^R*52246* ΩS^R from pUI1638 interrupting *clpA* pBAC1634 AatII/NdeI X ACN1260	This study
ACN2247	Δ*aalR51280; clpA*::ΩS^R*52246* pBAC1634 AatII/NdeI X ACN1280	This study
ACN2252	Δ*aalR51280*; *AspA52153*; Δ*darR*::*sacB*-K^R*51721* ACN1721 lysate X ACN2195	This study
ACN2253	Δ*aalR51280*; *AspA52153*; Δ*darR51726* ACN2220 L-Asp$^+$ spontaneous mutant	This study
ACN2265	Δ*aalR*::*darR52265* (G->T, **323,095**); Δ*darR51726* *darR* replacing *aalR* with G -> T point mutation isolated from Stabb Lab experiments pBAC1658 PsiI/KpnI X ACN1924	This study

ACN2266	Δ*aalR*::*darR52266* (G->T, 323,095); Δ*darR51726* longer *darR* replacing *aalR* with G -> T point mutation isolated from Stabb Lab experiments pBAC1659 PsiI/KpnI X ACN1924	This study
ACN2277	Δ*aalR51280*; *AspA*::*gfp*-ΩK^{R}*51894*; Δ*darR51726* pBAC1468 PsiI X ACN2220	This study
ACN2278	Δ*aalR51280*; *AspA52153; AspA*::*gfp*-ΩK^{R}*51894*; Δ*darR51726* pBAC1638 PsiI X ACN2220	This study
ACN2291	Δ*aalR*::*darR52265*; *AspA52153*; Δ*darR51726* pBAC1665 AatII/SapI X ACN1924	This study
ACN2292	Δ*aalR*::*darR52266*; *AspA52153*; Δ*darR51726* pBAC1666 AatII/SapI X ACN1924	This study
ACN2293	Δ*aalR*::*darR52265*; *AspA*::*gfp*-ΩK^{R}*51883*; Δ*darR51726* pBAC1468 PsiI X ACN2265	This study
ACN2294	Δ*aalR*::*darR52266*; *AspA*::*gfp*-ΩK^{R}*51883*; Δ*darR51726* pBAC1468 PsiI X ACN2266	This study
ACN2295	Δ*aalR*::*darR52265*; *AspA52153*; *AspA*::*gfp*-ΩK^{R}*51883*; Δ*darR51726* pBAC1638 PsiI X ACN2291	This study
ACN2296	Δ*aalR*::*darR22665*; *AspA52153*; *AspA*::*gfp*-ΩK^{R}*51894*; Δ*darR51726* pBAC1638 PsiI X ACN2292	This study
ACN2910	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (FM-B for BM (students)) pBAC1051 AatII X ADP1	This study
ACN2911	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (DW-B for AC (students)) pBAC1051 AatII X ADP1	This study
ACN2912	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (FM-A (students)) pBAC1051 AatII X ADP1	This study
ACN2913	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (JS-B (students)) pBAC1051 AatII X ADP1	This study
ACN2914	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (JI-A (students)) pBAC1051 AatII X ADP1	This study
ACN2915	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (AW-B (students)) pBAC1051 AatII X ADP1	This study
ACN2916	Δ*aalR*::*sacB*-K^{R}*51260* Remade ACN1260 to check that phenotype is correct (DW-A (students)) pBAC1051 AatII X ADP1	This study
ACN2917	Δ*aalR51280*; *clpA52917* *clpA*, G -> T (**1,358,183**) SNP in 82nd nucleotide resulting in a premature stop codon ACN1280 L-Asp^{+} suppressor mutant isolated on 20 mM L-Asp as the sole carbon source (JS-student).	This study

ACN2918	Δ*aalR51280* ACN1280 L-Asp^{+} suppressor mutant isolated on 20 mM L-Asp as the sole carbon source (AW-student).	This study
ACN2919	Δ*aalR51280* ACN1280 L-Asp^{+} suppressor mutant isolated on 20 mM L-Asp as the sole carbon source (AC-student).	This study
ACN2921	Δ*AspT52035; AspY*::ΩS^{R}*52059; AspS52921* T insertion in promoter region of *AspS* (after **474,090**) ACN2061 D-Asp^{+} suppressor mutant on 20 mM D-Asp as the sole carbon sourse after overnight growth on 20 mM L-Asp as the sole carbon source (DW-student).	This study
ACN2922	Δ*AspT52035; AspY*::ΩS^{R}*52059; AspS52921* T insertion in promoter region of *AspS* (after **474,090**) ACN2061 D-Asp^{+} suppressor mutant on 20 mM D-Asp as the sole carbon sourse after overnight growth on 20 mM L-Asp as the sole carbon source (DW-student).	This study
ACN2923	Δ*AspT52035; AspY*::ΩS^{R}*52059; AspS52921* T insertion in promoter region of *AspS* (after **474,090**) ACN2061 D-Asp^{+} suppressor mutant on 20 mM D-Asp as the sole carbon sourse after overnight growth on 20 mM L-Asp as the sole carbon source (DW-student).	This study
ACN2924	Δ*AspT52035; AspY*::ΩS^{R}*52059; AspS52921* T insertion in promoter region of *AspS* (after **474,090**) ACN2061 D-Asp^{+} suppressor mutant on 20 mM D-Asp as the sole carbon sourse after overnight growth on 20 mM L-Asp as the sole carbon source (DW-student).	This study
ACN2925	Δ*aalR*::*sacB*-K^{R}*51260* ACN2911 L-Asp^{+}D-Asp^{+} suppressor mutant isolated on 20 mM L-Asp and 20 mM D-Asp as the carbon sources.	This study
ACN2926	Δ*aalR*::*sacB*-K^{R}*51260*; *clpA52926* GG insertion after 1,601st nucleotide (**1,359,702**) in *clpA* CDS causing alternative end to protein ACN2915 L-Asp^{+}D-Asp^{+} suppressor mutant isolated on 20 mM L-Asp and 20 mM D-Asp as the carbon sources.	This study
ACN2927	Δ*aalR*::*sacB*-K^{R}*51260*; *clpA52927* Deleted C (**1,359,000**) in codon 300 causing premature stop codon in *clpA*. ACN2913 L-Asp^{+}D-Asp^{+} suppressor mutant isolated on 20 mM L-Asp and 20 mM D-Asp as the carbon sources.	This study
ACN2928	Δ*aalR*::*sacB*-K^{R}*51260*; *clpA52928* *sacB*-K^{R} cassette from pRMJ1 replacing *aalR* (**1,753,790 - 1,754,674**). T insertion in 188th codon (1,358,664) causing premature stop codon in *clpA*. ACN2910 L-Asp^{+}D-Asp^{+} suppressor mutant isolated on 20 mM L-Asp and 20 mM D-Asp as the carbon sources.	This study
ACN2947	Δ*darR51726*; *racD52947; P$_{racD}$trc52947* *P$_{racD}$trc52947* = 437 nt from *racD* promoter (**318,942-319,379**) replaced with *trc* promoter (305 nt of *trc* construct from (6) downstream *tonB* terminator*). *racD52947* = *racD* encoding K3N replacement (CDS nt 9 change A->C). Selected for growth on 20 mM D-Asp as the sole carbon source. pBAC1950 AhdI X ACN1726	This study
ACN2948	Δ*darR51726*; *racD52947; P$_{racD}$trc52947* *P$_{racD}$trc52947* = 437 nt from *racD* promoter (**318,942-319,379**) replaced with *trc* promoter (305 nt of *trc* construct from (6) downstream *tonB* terminator*). *racD52947* = *racD* encoding K3N replacement (CDS nt 9 change A->C). Selected for growth on 20 mM D-Asp as the sole carbon source. pBAC1950 AhdI X ACN1726	This study
ACN2949	Δ*AspS*::*sacB*-S^{R}*52949; AspS52921* *AspS* (**474,293-475,597**) removed with ACN2921 and replaced with *sacB*-S^{R} and extra T (after **474,090**) in promoter region (P$_{AspS}$*) pBAC1949 AhdI X ADP1	This study

ACN2950	Δ*aalR*::*darR52140*; Δ*darR51726* Re-made ACN2140. pBAC1510 BamHI/EcoRV X ACN1924	This study
ACN2951	Δ*aalR*::*darR52141*; Δ*darR51726* Re-made ACN2141. pBAC1607 BamHI/EcoRV X ACN1924	This study
ACN2952	Δ*darR51726; P_{racD}trc52947* Removed native *racD* promoter (**318,942-319,379**) replaced with trc promoter. pBAC1952 AhdI/NdeI X ACN1964	This study
ACN2953	Δ*darR51726*; Δ*racD52953; P_{racD}trc52947* Removed native *racD* promoter and CDS (**318,942-320,087**) replaced with trc promoter. pBAC1953 AhdI X ACN1964	This study
ACN2954	Δ*darR51726*; *racD52947; ΔaalR*::*sacB*-K^R*51260; P_{racD}trc52947* pBAC1051 AatII X ACN2947	This study
ACN2955	Δ*darR51726*; *racD52947; ΔaalR*::*sacB*-K^R*51260; P_{racD}trc52947* pBAC1051 AatII X ACN2948	This study
ACN2956	Δ*darR51726*; *racD52947; P_{racD}trc52947* pBAC1952 AhdI/NdeI X ACN1726	This study
ACN2957	Δ*darR51726, P_{racD}trc52947; ΔaalR*::*sacB*-K^R*51260* pBAC1051 AhdI/AatII X ACN2952	This study
ACN2958	Δ*darR51726*; Δ*racD52953; ΔaalR*::*sacB*-K^R*51260; P_{racD}trc52947* pBAC1051 AhdI/AatII X ACN2953	This study
ACN2959	Δ*darR51726; P_{racD}trc52947; ΔaalR*::*sacB*-K^R*51260* pBAC1051 AhdI/AatII X ACN2956	This study
ACN2964	*AspS*::*lacZ*-K^R*52964* *lacZ*-K^R interrupting *AspS* after 47th codon (**474,433**) pBAC1954 AhdI X ADP1	This study
ACN2965	*AspS*::*lacZ*-K^R*52964; Δdar51726* pBAC1954 AhdI X ACN1726	This study
ACN2966	*AspS*::*lacZ*-K^R*52964; ΔaalR51280* pBAC1954 AhdI X ACN1280	This study
ACN2967	*AspS*::*lacZ*-K^R*52964; ΔAspT52035; AspY*::ΩS^R*52059* pBAC1954 AhdI X ACN2061	This study
ACN2968	*AspS*::*lacZ*-K^R*52964; ΔAspT52035; AspY*::ΩS^R*52059* pBAC1954 AhdI X ACN2921	This study
ACN2969	*AspS*::*lacZ*-K^R*52969; ΔAspT52035; AspY*::ΩS^R*52059* *lacZ*-K^R interrupting *AspS* after 47th codon (**474,433**) with T insertion in promoter region (after **474,090**). (P_{AspS}*) pBAC1955 AhdI X ACN2061	This study
ACN2970	*AspS*::*lacZ*-K^R*52969; ΔAspT52035; AspY*::ΩS^R*52059* pBAC1955 AhdI X ACN2921	This study
ACN2971	*AspS*::*lacZ*-K^R*52969; ΔaalR51280* pBAC1955 AhdI X ACN1280	This study
ACN2972	*AspS*::*lacZ*-K^R*52969; Δdar51726* pBAC1955 AhdI X ACN1726	This study
ACN2978	Δ*darR51726*; Δ*racD52953; P_{racD}trc52947* *aalR* PCR (KTE121/124) DpnI X ACN2958	This study
ACN2979	Δ*darR51726*; Δ*racD52953; ΔaalR51280; P_{racD}trc52947* pBAC1032 AhdI X ACN2958	This study

ACN2980	Δ*racD*::*sacB*-K^R*51720; ΔracD52953; P$_{racD}$trc52947* pBAC1953 AhdI X ACN1720	This study
ACN2981	Δ*racD52953; ΔaalR51280; ΔdarR*::*sacB*-K^R*51721; P$_{racD}$trc52947* pBAC1297 AhdI X ACN2979	This study
ACN2982	Δ*racD52953; ΔaalR51280; P$_{racD}$trc52947* darR PCR (ALS7/8) DpnI X ACN2981	This study
ACN2983	*AspY*::*lacZ*-K^R*52144* pBAC1611 AhdI X ADP1	This study
ACN2984	*AspY*::*lacZ*-K^R*52144; ΔdarR51726*; Δ*racD52953; P$_{racD}$trc52947* pBAC1611 AhdI X ACN2978	This study
ACN2986	*racD*-ΩK^R*52986* ΩK^R inserted downstream stop codon of *racD* (**320,090**) to determine if *AspT* has is own promoter pBAC1956 PvuI X ADP1	This study
ACN2987	*AspS*::*lacZ*-K^R*52964; ΔAspT52035* pBAC1954 AhdI X ACN2035	This study
ACN2988	*AspS*::*lacZ*-K^R*52964; AspY*::ΩS^R*52059* pBAC1954 AhdI X ACN2059	This study
ACN2989	*AspY*::*lacZ*-K^R*52144; ΔracD52953; ΔaalR51280; P$_{racD}$trc52947* pBAC1611 AhdI X ACN2982	This study

* *tonB* terminator sequence: AGTCAAAAGCCTCCGGTCGGAGGCTTTTGACT

a. *A. baylyi* strains were derived from ADP1, previously known as *Acinetobacter calcoaceticus* or *Acinetobacter sp.* (7)

b. Bold numbers correspond to positions on the ADP1 chromosome in NCBI entry NC_005966.

c. Underlined text indicates the donor DNA and, where relevant, the restriction enzyme used to linearize a plasmid (pBAC number/Enzyme). The donor DNA was used to transform (X) the indicated recipient strain.

Table 3.3: Plasmids used in Chapter 3 study

Plasmid	**Relevant Characteristics**	**Source**
pUC18	Ap^R; cloning vector	(34)
pUC19	Ap^R; cloning vector	(34)
pminiT2.0	Ap^R; cloning vector	New England Biolabs
pJPK13	K^R; expression vector	(35)

pUI1637	Ap^R, K^R; source of ΩK^R cassette	(36)
pUI1638	Ap^R, S^R; source of ΩS^R cassette	(36)
pKOK6	Ap^R K^R; source of *lacZ*-K^R cassette	(15)
pRMJ1	Ap^R, K^R; source of *sacB*-K^R cassette	(37)
pJLS27	Cm^R K^R; source of *gfp*	(38)
pBAC1032	Ap^R; Δ*aalR51280* DNA fragments generated by PCR upstream *aalR* with KTE121 and KTE122 (**1,753,277 - 1,753,789**) and downstream *aalR* with KTE123 and KTE124 (**1,754,675 - 1,755,236**) including an XhoI recognition site engineered in frame with *aalR* in between the start and stop codons (**1,753,790 - 1,754,674**) assembled by SOEing. DNA was inserted into pUC19 by digestion with BamHI and EcoRI	This study
pBAC1033	Ap^R; Δ*mdcR51281* (**1,766,508–1,766,937**; **1,767,853–1,768,371**) in pUC19 vector	(27)
pBAC1051	Ap^R K^R; Δ*aalR*::*sacB*-K^R*51260* *sacB*-K^R cassette from pRMJ1 digested with SalI inserted in XhoI site of pBAC1032 (compatible sticky ends)	This study
pBAC1052	Ap^R, K^R; Δ*mdcR*::*sacB*-K^R*51261*; *sacB*-K^R cassette from pRMJ1 inserted in XhoI site of pBAC1033	(27)
pBAC1289	Ap^R, K^R; Δ*racD*::*sacB*- K^R*51720*; *sacB*- K^R cassette inserted in XhoI site of pBAC1291	(8)
pBAC1291	Ap^R ; *ΔracD51725* in pUC19 XhoI site replacing *racD*	(8)
pBAC1296	Ap^R; Δ*darR51726* XhoI site replacing *darR*	(8)
pBAC1297	Ap^R, K^R; *ΔdarR*::*sacB*- K^R*51720*; *sacB*- K^R cassette inserted in XhoI site of pBAC1296	(8)
pBAC1299	Ap^R; *darR* (**322,810 – 323,700**) inserted into pBAC1296 linearized with XhoI by gap repair with ADP1	This study
pBAC1300	Ap^R; site directed mutagenesis of pBAC1299 with ALS9 and ALS10 to generate change to nucleotide 603 of *darR* (G -> T; M -> I; **323,095**)	This study
pBAC1407	Ap^R; PCR fragment KTE121 and ALS52 carrying part of *AspA* and *aalR* (**1,753,277-1,754,674**) introduced into pBAC1032 by restriction digest with SphI and HindIII	This study
pBAC1419	Ap^R ; *AspA*::*gfp51894, gfp* from pJLS27 amplified with ALS69 and ALS70 inserted into SphI site in *aalR* in pBAC1407	This study
pBAC1421	Ap^R ; *AspA*::*gfp51894,* additional ADP1 sequence (**1,749,977 - 1,753,341**) amplified from WT with ALS72 and ALS75 inserted into pBAC1419 KpnI and EcoRI sites	This study
pBAC1424	Ap^R, K^R; Δ*mdcR*::*sacB-mCherry* mCherry from pJLS27 digested with XhoI and XbaI ligated pBAC1052 digested with XhoI and SpeI	This study
pBAC1425	Ap^R ; Δ*mdcR*::*mCherry51862* pBAC1424 cut with EcoRV and SnaBI and ligated to remove *sacB*	This study
pBAC1466	Ap^R K^R; ΩK^R cassette from pUI1637 inserted into EcoRI site in pUC18	This study
pBAC1468	Ap^R K^R; *AspA*::*gfp*-ΩK^R*51883* ΩK^R cassette from pBAC1466 inserted into KpnI site in pBAC1421	This study
pBAC1504	Ap^R; Δ*aalR*::*darR52140, darR* (**322,807 - 323,697**) amplified with ALS110 and ALS116 and sequence downstream *aalR* (**1,754,678 – 1,755,236**) amplified with ALS117 and KTE124 assembled with Splicing by Overlap Extension PCR with ALS116 and KTE124 and inserted into BamHI and PsiI sites of pBAC1032	This study
pBAC1510	Ap^R; Δ*aalR*::*darR52140*, sequence near PsiI site in pBAC1504 restored to match pBAC1032 by replacing DNA from KpnI site to PsiI site (**1,753,757 – 1,753,336**)	This study
pBAC1541	Ap^R; Δ*AspT52035*; DNA upstream *AspT* (**318,384 - 320,207**) amplified with MTV399 and ALS134 and DNA downstream *AspT* (**321,533 - 324,768**) amplified with MTV396 and ALS135 with an engineered XhoI site in place of *AspT* and 26 bp DNA downstream assembled with	This study

	Splicing by Overlap Extension PCR with MTV399 and MTV396 and inserted into SacI and PstI sites in pUC19	
pBAC1567	Ap^R; *AspY,* DNA from WT (**1,334,426 - 1,337,336**) amplified with ALS141 and ALS142 inserted into XbaI and HindIII sites in pUC18	This study
pBAC1569	Ap^R, S^R; *AspY*::ΩS^R*52059*, ΩS^R cassette from pUI1638 inserted into PstI site of pBAC1567	This study
pBAC1607	Ap^R; Δ*aalR*::*darR52141* Site directed mutagenesis of pBAC1510 with ALS165 and ALS166 introducing 6 nucleotides (**323,698 – 323,703**) upstream *darR* adding the alternative start site	This study
pBAC1611	Ap^R, K^R; *AspY*::*lacZ*-K^R*52144*, *lacZ*-K^R cassette from pKOK6 inserted in XhoI site of pBAC1567 disrupting *AspY*	This study
pBAC1612	Ap^R, K^R; *AspY*::*lacZ*-K^R*52179*, *lacZ*-K^R cassette from pKOK6 inserted in XhoI site of pBAC1567 in opposition orientation to pBAC1611 disrupting *AspY*	This study
pBAC1629	Ap^R; WT sequence (**1,357,557-1,360,899**) including *clpA* amplified with ALS143 and ALS144 inserted into pminiT	This study
pBAC1630	Ap^R; ACN2180 sequence (**1,357,557-1,360,899**) including *clpA52180* amplified with ALS143 and ALS144 inserted into pminiT	This study
pBAC1634	Ap^R, S^R; *clpA*::ΩS^R*52246*; ΩS^R cassette from pUI1638 digested with EcoRV inserted into SnaBI site of pBAC1629	This study
pBAC1636	Ap^R ; *AspA*::*gfp51894*; *AspA52153* Site directed mutagenesis of pBAC1468 with SRB113 and SRB114 introducing *AspA* promoter mutation C->T at position **1,753,683**	This study
pBAC1637	Ap^R; Δ*aalR51280*; *AspA52153* Site directed mutagenesis of pBAC1032 with SRB113 and SRB114 introducing *AspA* promoter mutation C->T at position **1,753,683**	This study
pBAC1638	Ap^R K^R; *AspA*::*gfp52153;* ΩK^R*51894,* ΩK^R from pBAC1468 inserted into KpnI site of pBAC1636 with *gfp* interrupting *aalR* and *AspA* promoter mutation C->T at position **1,753,683**	This study
pBAC1645	Ap^R K^R; Δ*aalR*::darR51943; AspA::*gfp52153;* ΩK^R*51894,* *AspA::gfp* region from pBAC1638 digested with PsiI inserted into PsiI and SfoI sites in pBAC1510	This study
pBAC1646	Ap^R K^R; Δ*aalR*::darR52141; AspA::*gfp52153,* ΩK^R*51894, AspA::gfp* region from pBAC1638 containing *AspA* promoter mutation digested with PsiI inserted into PsiI and SfoI sites in pBAC1607	This study
pBAC1658	Ap^R; Δ*aalR*::*darR52265*, pBAC1300 and pBAC1510 digested with NcoI and EcoRV and ligated to replace *aalR* with shorter *darR* (**322,807 - 323,697)** with point mutation G->T in *darR* at position **323,095** resulting in M201I	This study
pBAC1659	Ap^R; Δ*aalR*::*darR52266*, pBAC1300 and pBAC1607 digested with NcoI and EcoRV and ligated to replace *aalR* with longer *darR* (**322,807 - 323,703)** with point mutation G->T in *darR* at position **323,095** resulting in M203I	This study
pBAC1665	Ap^R; Δ*aalR*::*darR52265*; *AspA52153* *AspA52153* promoter mutation from pBAC1637 digested with PsiI and KpnI and inserted into PsiI and KpnI sites of pBAC1658 containing *aalR* replaced with shorter *darR* (**322,807 - 323,697)** with point mutation G->T in *darR* at position **323,095** resulting in M201I	This study
pBAC1666	Ap^R; Δ*aalR*::*darR52266*; *AspA52153* *AspA52153* promoter mutation from pBAC1637 digested with PsiI and KpnI and inserted into PsiI and KpnI sites of pBAC1659 containing *aalR* replaced with longer *darR* (**322,807 - 323,703)** with point mutation G->T in *darR* at position **323,095** resulting in M203I	This study
pBAC1673	K^R; *racD* and *AspT* sequences (**319,380-321,518**) in pJPK13, under control of the *lac* promoter assembled by NEBuilder of ALS185 and ALS186 pJPK13 fragment and ALS187 and ALS188 fragment with *racD* and *AspT* sequences	This study
pBAC1680	K^R; *racD* under control of *lac* promoter in pJPK13 made by digest with NcoI and re-ligation to remove *AspT* sequences (**320,463 – 321,056**)	This study

pBAC1683	K^R; *AspT* under control of *lac* promoter in pJPK13 assembled by NEBuilder of ALS185 and ALS186 pJPK13 fragment and ALS187 and ALS196 fragment with *AspT* sequences (**320,208 – 321,518**)	This study
pBAC1935	K^R; *AspY* under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB754 and ALS185 pBAC1683 fragment and SRB751 and SRB754 fragment with *AspY* sequences (**1,335,452 – 1,336,750**)	This study
pBAC1936	K^R; ACIAD_RS10205 (TP3) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB754 and ALS185 pBAC1683 fragment and SRB756 and SRB758 fragment with ACIAD_RS10205 sequences (**2,119,881 – 2,201,277**)	This study
pBAC1937	K^R; *AspS* under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB754 and ALS185 pBAC1683 fragment and SRB760 and SRB762 fragment with *AspS* sequences (**474,293 – 475,597**)	This study
pBAC1938	K^R; ACIAD_RS04730 (TP5) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB754 and ALS185 pBAC1683 fragment and SRB764 and SRB767 fragment with ACIAD_RS04730 sequences (**1,017,738 – 1,019,009**)	This study
pBAC1939	K^R; ACIAD_RS14645 (TP6) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB754 and ALS185 pBAC1683 fragment and SRB768 and SRB771 fragment with ACIAD_RS14645 sequences (**3,143,711 – 3,144,952**)	This study
pBAC1940	K^R; ACIAD_RS05475 (TP7) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB754 and ALS185 pBAC1683 fragment and SRB772 and SRB775 fragment with ACIAD_RS05475 sequences (**1,189,971– 1,191,359**)	This study
pBAC1941	K^R; *racD* and *AspY* under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB753 and SRB754 pBAC1673 fragment and SRB751 and SRB752 fragment with *AspY* sequences (**1,335,452 – 1,336,750**)	This study
pBAC1942	K^R; *racD* and ACIAD_RS10205 (TP3) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB753 and SRB754 pBAC1673 fragment and SRB756 and SRB757 fragment with ACIAD_RS10205 sequences (**2,119,881 – 2,201,277**)	This study
pBAC1943	K^R; *racD* and *AspS* under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB753 and SRB754 pBAC1673 fragment and SRB760 and SRB761 fragment with *AspS* sequences (**474,293 – 475,597**)	This study
pBAC1944	K^R; *racD* and ACIAD_RS04730 (TP5) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB753 and SRB754 pBAC1673 fragment and SRB764 and SRB765 fragment with ACIAD_RS04730 sequences (**1,017,738 – 1,019,009**)	This study
pBAC1945	K^R; *racD* and ACIAD_RS14645 (TP6) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB753 and SRB754 pBAC1673 fragment and SRB768 and SRB769 fragment with ACIAD_RS14645 sequences (**3,143,711 – 3,144,952**)	This study
pBAC1946	K^R; *racD* and ACIAD_RS05475 (TP7) under control of *lac* promoter in pJPK13 assembled by NEBuilder of SRB753 and SRB754 pBAC1673 fragment and SRB772 and SRB773 fragment with ACIAD_RS05475 sequences (**1,189,971– 1,191,359**)	This study
pBAC1947	Ap^R; *AspS* with P_{AspS}* in pUC19 assembly in *E. coli* DH5α of fragments make by PrimeSTAR PCR with SRB780 and SRB781 (*AspS* fragment from ACN2921 template) (**473,292 - 476,076**) and SRB776 and SRB777 (pUC19)	This study
pBAC1948	Ap^R; Δ*AspS* (**474,293 – 475,597**) with P_{AspS}*; Constructed by E. coli assembly of fragments make by PrimeSTAR PCR with SRB780 and SRB793 (upstream TP4 fragment from ADP1 template), SRB792 and SRB781 (downstream TP4 fragment from ADP1 template), and SRB776 and SRB777 (pUC19).	This study
pBAC1949	Ap^R S^R; Δ*AspS* (**474,293 – 475,597**) replaced with *sacB*::S^R with P_{AspS}*; Constructed by E. coli assembly of fragments make by PrimeSTAR PCR with SRB796 and SRB797 (pBAC1947 template), and SRB794 and SRB795 (pUI1638 template)	This study
pBAC1950	Ap^R; trc promoter controlling *racD* with K3N mutation and intergenic region upstream *racD* deleted (no DarR binding site); Constructed by *E. coli* assembly of fragments make by PrimeSTAR PCR with SRB800 and SRB802 (upstream *racD* from ADP1 template), SRB803 and SRB804 (trc promoter from pBAC1919 template(Bradley paper reference)), SRB805 and	This study

	SRB806 (*racD* from ADP1 template), and SRB807 and SRB801 (pUC19). Upon sequencing I found a mutation in the 3rd codon changing a K to and N.	
pBAC1951	Ap^R; *AspS* in pUC19; constructed by *E. coli* assembly of fragments made by PrimeSTAR PCR with SRB780 and SRB781 (*AspS* fragment from ADP1 template) and SRB776 and SRB777 (pUC19)	This study
pBAC1952	Ap^R; trc promoter controlling *racD* and intergenic region upstream *racD* deleted (no DarR binding site); Constructed by *E. coli* assembly of fragments make by PrimeSTAR PCR with SRB800 and SRB802 (upstream *racD* from ADP1 template), SRB803 and SRB804 (trc promoter from pBAC1919 template(Bradley paper reference)), SRB805 and SRB806 (*racD* from ADP1 template), and SRB807 and SRB801 (pUC19).	This study
pBAC1953	Ap^R; trc promoter upstream Δ*racD* (**319,383 – 320,087**); Constructed by *E. coli* assembly of fragments made by PrimeSTAR PCR with SRB808 and SRB809 (pBAC1952 template) to remove *racD* coding sequence in between the start and stop codons	This study
pBAC1954	Ap^R K^R; *AspS*::*lacZ*-K^R; Constructed by *E. coli* assembly of fragments made by PrimeSTAR PCR with SRB810 and SRB811 (*lacZ*-K^R from pKOK6 template) and SRB812 and SRB813 (*AspS* from pBAC1951 template).	This study
pBAC1955	Ap^R K^R; *AspS*::*lacZ*-K^R; P_{AspS}*; Constructed by *E. coli* assembly of fragments made by PrimeSTAR PCR with SRB810 and SRB811 (*lacZ*-K^R from pKOK6 template) and SRB812 and SRB813 (*AspS* from pBAC1947 template).	This study
pBAC1956	Ap^RK^R; trc promoter controlling *racD* and intergenic region upstream *racD* deleted (no DarR binding site) with ΩK^R directly downstram of RacD CDS; Constructed by *E. coli* assembly of fragments make by PrimeSTAR PCR with SRB816 and SRB817 (ΩK^R from pUI1637 template) and SRB814 and SRB815 (pBAC1952 template)	This study

Table 3.4: Primers used in Chapter 3 study

Name	**Sequence (5'-3')**
ALS5	CGTTCCCACAATGAATCTTG
ALS6	ATCAGGAGTGAGGTTCAAG
ALS7	TCTATTCAGCCGTCGTATTG
ALS8	ACCGCAATCATTATCTGAGT
ALS9	AAATTTATTGACCTTTTTGCC**A**ATATAGCTATCATCGG TATAATTCATTAATGG
ALS10	CCATTAATGAATTATACCGATGATAGCTATAT**T**GGCAA AAAGGTCAATAAA TTT
ALS18	GTATGTGTTTGTTGCTGTAATTGA
ALS52	GATATACTCGAGGTATTCGATATTTCTAAGAGAATCCC
ALS58	GAGTCTGCATGCGCTTAAGGAGATATACATATGGCTA G
ALS59	GATCATGCATGCTCAGTTGTACAGTTCATCCATG
ALS60	GGATAAACGTCATTGGTCGATTGAG
ALS61	TGGCATTCATGTTGGTAGAGGT
ALS62	AATGGGGAACTGTTCGTGGTA
ALS66	TTCAAGACTCATTGCTTCATGTC
ALS69	TTTGCTTCTGCACATGCATGTCAGTTGTACAGTTCATC
ALS70	ATTAGCGATGGTGAAGCATGTCTAGGCTTAAGGAGAT ATACATATGG
ALS72	TCAATAATATCTGACCCGCATCG

ALS73	AAACCCAGTTATTCCAGAGGC
ALS74	AGATAAGGATTCAGAGCAGTCAC
ALS75	TCCTGCACCTAAAAATCCTG
ALS92	TATCCAGAAAGCTTAAACTCACC
ALS93	ATGCTTCACCATCGCTAATG
ALS104	TTACTGGGTCTACGTGAAGT
ALS110	TTCAAAAATTTGCTATTATTTCATGTCCAGTCAATTGC
ALS114	ACGTTCTTGTTCGAGTGTGAGTAG
ALS115	GCGCATTATAAACACGGTATGC
ALS116	TTATAATGCGCCAAAACACTATATGACATGAAGCATT GATACAAAATATTGAATTAAAGTGG
ALS117	GCAATTGACTGGACATGAAATAATAGCAAATTTTTGA AAAAGC
ALS126	AACACATCTGCTCATTCTGATC
ALS132	CACTCAGTAGACGCATTCAAAC
ALS133	TGGCAATGACCGATCTTCTAAAG
ALS134	TGAAAGTAACTCCAACTAAGCTACTCTCGAGACTACA CTCACTAGATTACCGAAG
ALS135	CTTCGGTAATCTAGTGAGTGTAGTCTCGAGAGTAGCTT AGTTGGAGTTACTTTCA
ALS136	TGAAAGTAACTCCAACTAAGCTACT
ALS141	CGACAGTCAAACAGGCAA
ALS142	GCAATTCAATCGCAAGGTC
ALS143	GTACACCTGTGGTCAAGC
ALS144	CTGCTACATCAGGCTCTAG
ALS163	GGCATGTCACAGTGCAT
ALS164	AATGCACTGTGACATGCC
ALS165	ATGTCCTTGATACAAAATATTGAATTAAAGTG
ALS166	TGCTTCATGTCATATAGTG
ALS179	CTGCTCTTGCTTTTGCGAC
ALS180	CGCAATCACCGTACCTTC
ALS181	ATCATGCTCATGTGCTTCAG
ALS182	GAATGGTACGTTGCGTTGC
ALS183	GGTTAATCTGTATTTTGC
ALS184	ATACAGCCGACATTC
ALS185	AGCTGTTTCCTGTGTGAAATTG
ALS186	GGTGATGGTTCACGTAGTG
ALS187	CCACTACGTGAACCATCACCTTAAAAACTGAGTTAGG CAC
ALS188	ATTTCACACAGGAAACAGCTATGCATAAAGCAATCGG AATATTAG
ALS192	CAACAGTTGATGCGA
ALS193	TGTTATGCAAAATGTAGAT
ALS194	CATATAAAAATTTGGATAATTCTT

ALS195	CAGTTGCTGGATCTC
ALS196	ATTTCACACAGGAAACAGCTATGAAAAAGGAAACCAC AGTACGCT
LES9	ATGAAATTCTGTAATGTC
LES10	AATAAGAGAAGCACCGATGAACC
LES11	TGAGCCTAAAGATGAAGATTTACCC
LES12	ATTCTTTTTGATACAGTTC
KTE121	GATCTTGAATTCTGACCCAAATGCTCTAATG
KTE122	GGCTTTTTCAAAAATTTGCTATTATTTTACTCGAGCATT GCTTCATGTCATATAGTGTT
KTE123	AACACTATATGACATGAAGCAATGCTCGAGTAAAATA ATAGCAAATTTTTGAAAAAGCC
KTE124	TTGCGAGGATCCAATGAGATAGTAAGGAAATTCTTCT AG
MTV58	GCCATTACGCTCGTCATCAA
MTV89	GACAGCATCCTTGAACAAG
MTV396	GATCATCTGCAGTGATTCCCTGGATAACGAATTTGAC
MTV399	GAGTCAGAGCTCAAATTGACTGGAAAACATGGTGTTG
MTV475	GCAGCCGATTGTCTGTTG
SRB113	TTGTCTAGATAGCGATATAAG
SRB114	GATGATTCATTATAGAATTGGGTTTG
SRB749	GAACAAGAGTCCACTATTAAAGAACGTG
SRB750	AGCTGGCACGACAGG
SRB751	CCATCACCTTAAAAACTGAGTTATGAATGACCTGTCGT ACTCAC
SRB752	GGTAATCTAGTGAGTGTAGTATGAATCAGAAAAAATT GCTCAAATATAT
SRB753	ACTACACTCACTAGATTACCGAAGAAC
SRB754	CTCAGTTTTTAAGGTGATGGTTCAC
SRB755	ATTTCACACAGGAAACAGCTATGAATCAGAAAAAATT GCTCAAATATAT
SRB756	CCATCACCTTAAAAACTGAGTTATGCATGTGCAGGAC GATC
SRB757	GGTAATCTAGTGAGTGTAGTATGGCTAAAAAACCAAT TTATAA
SRB758	ATTTCACACAGGAAACAGCTATGGCTAAAAAACCAAT TTATAA
SRB759	TCTGAGTCTGTATTGCCACG
SRB760	CCATCACCTTAAAAACTGAGTTAAGCATCAGCTTTGGG TTGATTCAGAG
SRB761	GGTAATCTAGTGAGTGTAGTATGAAAAATTTTAAAATT AGCCTCGCATGGCAAATT
SRB762	ATTTCACACAGGAAACAGCTATGAAAAATTTTAAAAT TAGCCTCGCATGGCAAATT
SRB763	CTCAAGGATGAATTGATTTTAGCGTATTC

SRB764	CCATCACCTTAAAAACTGAGTTAGTTGACTGATGATTTATGCTCAAGTACG
SRB765	GGTAATCTAGTGAGTGTAGTATGTGGCAGAAATTAAAGTCAAACC
SRB766	CTCTTCTGATTCTGTGCTTCCAC
SRB767	ATTTCACACAGGAAACAGCTATGTGGCAGAAATTAAAGTCAAACC
SRB768	CCATCACCTTAAAAACTGAGTTACTCAGTTGAACTTTCTATTACTTTGATATTTCG
SRB769	GGTAATCTAGTGAGTGTAGTATGAGCTTAAATACGCAAATTTTAATTGC
SRB770	CATGGTGCAGTGGTCTTGC
SRB771	ATTTCACACAGGAAACAGCTATGAGCTTAAATACGCAAATTTTAATTGC
SRB772	CCATCACCTTAAAAACTGAGCTAAGAGTGATCCAGTGTGACATC
SRB773	GGTAATCTAGTGAGTGTAGTATGAATTGGCCTCTAATCATAAACATTGTTC
SRB774	CCTGTACTCACCTTTGCATTTAC
SRB775	ATTTCACACAGGAAACAGCTATGAATTGGCCTCTAATCATAAACATTGTTC
SRB778	CGACGGCCAGTGAATTCGAGCACCATAGCCACCATATTCTTCTGG
SRB779	GCTATGACCATGATTACGCCAAGAAACTTCACCTTTCAGAATCATACTAAATGC
SRB780	CGACGGCCAGTGAATTCGAGCCTACAAACCGATAAGGTATTTGAAC
SRB781	GCTATGACCATGATTACGCCAAGGGTGCTGATCCCAAGAAGAAGG
SRB782	CGACGGCCAGTGAATTCGAGGTTGTTATGCAATCAATCCAACTGG
SRB783	GCTATGACCATGATTACGCCAAGCTTAAGTTTATCTTCTGATGGCTGACC
SRB784	CGACGGCCAGTGAATTCGAGTCAGATAATCGCTTAATGCCACAC
SRB785	GCTATGACCATGATTACGCCAAGGCGAGTAATTCTGCAAAACATAGC
SRB786	CGACGGCCAGTGAATTCGAGGCAATACGAAGTTCAGGATCTGC
SRB787	GCTATGACCATGATTACGCCAAGCTAGCCAAGGTCGCATGTCC
SRB788	CTTTGCATTAAATCTTGACCAAGCTACTC
SRB789	CCACACAAAAACTGCCAAACTAATAATG
SRB790	CGTACCGTAGAAGGTGCAGATG
SRB792	GTCTTATCGCATCAATAAACCTGATTCAA
SRB793	GCGATAAGACAATCCTTATGCTTGGCAAG
SRB794	TGCTTGCCAAGCATAAGGATTGTCTTCGATACCGTCGACCTCGAG

SRB795	TATTGATGCGATAATACAGAGAATGAAAAGAAACAGATAGATTTT
SRB796	TTTCTTTTCATTCTCTGTATTATCGCATCAATAAACCTGATTCAA
SRB797	TCGAGGTCGACGGTATCGAAGACAATCCTTATGCTTGGCAAG
SRB798	CTCAATAGAATGTTGCGATTATTTTATCG
SRB799	CTTAGATTTATACCAATCTGTCGTGAATCG
SRB800	ACGGCCAGTGAATTCGAGTTAAAAGTTGGCTTCAAACGTCATTTG
SRB801	AATGACGTTTGAAGCCAACTTTTAACTCGAATTCACTGGCCGTCG
SRB802	AGCCTCCGACCGGAGGCTTTTGACTATGTCTCACTATTTACAGAGATCATTTAAAAAATC
SRB803	AGCCTCCGGTCGGAGGCTTTTGACTGCGCAACGCAATTAATGTGAG
SRB804	TAATATTCCGATTGCTTTATGCATGGTCTGTTTCCTGTGTGAAATTGTTA
SRB805	AATTTCACACAGGAAACAGACCATGCATAAAGCAATCGGAATATTAGGAG
SRB806	GAGCGCAACGCAATGAAATAGATGAGTAATAAATCCATTTAATGAAAACC
SRB807	TTCATTAAATGGATTTATTACTCATCTATTTCATTGCGTTGCGCTCACTG
SRB808	GATCACATGGTCTGTTTCCTGTGTGAA
SRB809	CCATGTGATCATGAAGGTTCATGCATTTTGTG
SRB810	GGGCCAGATCGGGGGCTAGAGAGGAAACAGCTATG
SRB811	TAAAGATCGTTGTGTCTCAAAATCTCTGATGTTAC
SRB812	CACAACGATCTTTATTAGCCTGATCAAAATGATTG
SRB813	CCTCTCTAGCCCCCGATCTGGCCCAATGGCTTCAG
SRB814	GTACCGAGCTTCATGAAGGTTCATGCATTTTGTGC
SRB815	TCTACCGGGCCTCAGCTACAAGTTTCTGGCTGTCC
SRB816	AGCCAGAAACTTGTAGCTGAGGCCCGGTAGATCCG
SRB817	GCATGAACCTTCATGAAGCTCGGTACGATCCGGTG
SRB818	TAGCTAGCATGCATTTACGTTTAAGGAAGAATGTAAGGAAGTTGTGAAAC
SRB819	TCTAGCCCCCGAGTTTTACCTGTAATCTAAGATATTTATTAGCTTTACAG
SRB820	ATAAATATCTTAGATTACAGGTAAAACTCGGGGGCTAGAGAGGAAACAGC
SRB821	TGAGCCTTTCGTTTTATTTGTCAGAATTGGTTAATTGGTTGTAACACTGG
SRB822	TGTTACAACCAATTAACCAATTCTGACAAATAAAACGAAAGGCTCAGTCG
SRB823	TCACAACTTCCTTACATTCTTCCTTAAACGTAAATGCATGCTAGCTATAG
SRB824	CGGTATGCTGTTCAAACGTATTCTG

SRB825	GGGAAATAAAATGACACTAGGTCTAGG
SRB826	GAGACATTCCTTCCGTATCTTTTACG

Acknowledgments

Alicia L. Schmidt, Lauren E. Slarks, Chantel Duscent-Maitland, Katherine T. Elliot, Silke Andreson, Flavia Costa, Sophia Worth, Melissa Tumen-Velasquez, Lindsey Nilsen, Cassandra E. Bartlett, Anna Karls, Timothy R. Hoover for contributions to this work. Undergraduate students from multiple semesters of the microbiology undergraduate research course as well as Caroline M. Dunn, summer 2017 Research Experience for Undergraduates, also contributed to this work. This research was supported by National Science Foundation Grants: DEB-1556541, MCB-1615365, and DBI-1460671 (the REU Site program). This project was also supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Number DE-SC0022220. Additional funding was provided by the UGA Microbiology Department.

References

1. Ota N, Shi T, Sweedler JV. D-Aspartate acts as a signaling molecule in nervous and neuroendocrine systems. Amino Acids. 2012;43(5):1873-86.
2. Bastings JJ, van Eijk HM, Olde Damink SW, Rensen SS. d-amino acids in health and disease: A focus on cancer. Nutrients. 2019;11(9):2205.
3. Cava F, Lam H, De Pedro MA, Waldor MK. Emerging knowledge of regulatory roles of D-amino acids in bacteria. Cellular and Molecular Life Sciences. 2011;68(5):817-31.
4. Kobayashi J. D-amino acids and lactic acid bacteria. Microorganisms. 2019;7(12):690.

5. Wang R, Zhang Z, Sun J, Jiao N. Differences in bioavailability of canonical and non-c anonical D-amino acids for marine microbes. Science of The Total Environment. 2020;733:139216.

6. Biggs BW, Bedore SR, Arvay E, Huang S, Subramanian H, McIntyre EA, et al. Development of a genetic toolset for the highly engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. Nucleic Acids Research. 2020.

7. Vaneechoutte M, Dijkshoorn L, Tjernberg I, Elaichouni A, de Vos P, Claeys G, et al. Identification of *Acinetobacter* genomic species by amplified ribosomal DNA restriction analysis. Journal of Clinical Microbiology. 1995;33(1):11-5.

8. Jones Jr RM, Popham DL, Schmidt AL, Neidle EL, Stabb EV. *Vibrio fischeri* DarR directs responses to D-aspartate and represents a group of similar LysR-type transcriptional regulators. Journal of Bacteriology. 2018;200(15):e00773-17.

9. Craven SH, Ezezika OC, Momany C, Neidle EL. LysR homologs in *Acinetobacter*: insights into a diverse and prevalent family of transcriptional regulators: Norfolk: Caister Academic Press; 2008.

10. Li G, Lu C-D. Molecular characterization and regulation of operons for asparagine and aspartate uptake and utilization in *Pseudomonas aeruginosa*. Microbiology. 2018;164(2):205-16.

11. Neidle EL, Ornston LN. Cloning and expression of *Acinetobacter calcoaceticus* catechol 1, 2-dioxygenase structural gene catA in Escherichia coli. Journal of Bacteriology. 1986;168(2):815-20.

12. Hoskins JR, Pak M, Maurizi MR, Wickner S. The role of the ClpA chaperone in proteolysis by ClpAP. Proceedings of the National Academy of Sciences. 1998;95(21):12135-40.

13. Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proceedings of the National Academy of Sciences. 1988;85(23):8998-9002.

14. Wang W, Wu H, Xiao Q, Zhou H, Li M, Xu Q, et al. Crystal structure details of Vibrio fischeri DarR and mutant DarR-M202I from LTTR family reveals their activation mechanism. International Journal of Biological Macromolecules. 2021.

15. Kokotek W, Lotz W. Construction of a lacZ-kanamycin-resistance cassette, useful for site-directed mutagenesis and as a promoter probe. Gene. 1989;84(2):467-71.

16. Grewer C, Gameiro A, Rauen T. SLC1 glutamate transporters. Pflügers Archiv-European Journal of Physiology. 2014;466(1):3-24.

17. Ji Y, Postis VL, Wang Y, Bartlam M, Goldman A. Transport mechanism of a glutamate transporter homologue GltPh. Biochemical Society Transactions. 2016;44(3):898-904.

18. Arkhipova V, Trinco G, Ettema TW, Jensen S, Slotboom DJ, Guskov A. Binding and transport of D-aspartate by the glutamate transporter homolog GltTk. Elife. 2019;8:e45286.

19. Rahman M, Ismat F, Jiao L, Baldwin JM, Sharples DJ, Baldwin SA, et al. Characterisation of the DAACS family *Escherichia coli* glutamate/aspartate-proton symporter GltP using computational, chemical, biochemical and biophysical methods. The Journal of Membrane Biology. 2017;250(2):145-62.

20. Zuo S, Zhang T, Jiang B, Mu W. Recent research progress on microbial L-asparaginases. Applied Microbiology and Biotechnology. 2015;99(3):1069-79.

21. Sun D, Setlow P. Cloning, nucleotide sequence, and expression of the *Bacillus subtilis* ans operon, which codes for L-asparaginase and L-aspartase. Journal of Bacteriology. 1991;173(12):3831-45.

22. Campbell HA, Mashburn LT. L-Asparaginase EC-2 from *Escherichia coli*. Some substrate specificity characteristics. Biochemistry. 1969;8(9):3768-75.

23. Tumen-Velasquez MP, Laniohan NS, Momany C, Neidle EL. Engineering CatM, a LysR-type transcriptional regulator, to respond synergistically to two effectors. Genes. 2019;10(6):421.

24. Browning DF, Busby SJ. The regulation of bacterial transcription initiation. Nature Reviews Microbiology. 2004;2(1):57-65.

25. Liu F, Hao J, Yan H, Bach T, Fan L. AspC-mediated aspartate metabolism coordinates the *Escherichia coli* cell cycle. PLoS One. 2014;9(3):e92229.

26. Schubert C, Zedler S, Strecker A, Unden G. L□Aspartate as a high□quality nitrogen source in *Escherichia coli:* Regulation of L□aspartase by the nitrogen regulatory system and interaction of L□aspartase with GlnB. Molecular Microbiology. 2021;115(4):526-38.

27. Stoudenmire JL, Schmidt AL, Tumen-Velasquez MP, Elliott KT, Laniohan NS, Whitley SW, et al. Malonate degradation in *Acinetobacter baylyi* ADP1: Operon organization and regulation by MdcR. Microbiology. 2017;163(5):789.

28. Singh A, Bedore SR, Sharma NK, Lee SA, Eiteman MA, Neidle EL. Removal of aromatic inhibitors produced from lignocellulosic hydrolysates by *Acinetobacter baylyi*

ADP1 with formation of ethanol by *Kluyveromyces marxianus*. Biotechnology For Biofuels. 2019;12:91.

29. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual: Cold Spring Harbor Laboratory Press; 1989.

30. Kostylev M, Otwell AE, Richardson RE, Suzuki Y. Cloning should be simple: *Escherichia coli* DH5α-mediated assembly of multiple DNA fragments with short end homologies. PloS One. 2015;10(9):e0137466.

31. Neidle E, Hartnett C, Ornston L. Characterization of *Acinetobacter calcoaceticus* catM, a repressor gene homologous in sequence to transcriptional activator genes. Journal of Bacteriology. 1989;171(10):5410-21.

32. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990;215(3):403-10.

33. Juni E, Janik A. Transformation of *Acinetobacter calco-aceticus* (Bacterium anitratum). Journal of Bacteriology. 1969;98(1):281-8.

34. Norrander J, Kempe T, Messing J. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene. 1983;26(1):101-6.

35. Peterson J, Phillips GJ. New pSC101-derivative cloning vectors with elevated copy numbers. Plasmid. 2008;59(3):193-201.

36. Eraso JM, Kaplan S. prrA, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*. Journal of Bacteriology. 1994;176(1):32-43.

37. Jones RM, Williams PA. Mutational analysis of the critical bases involved in activation of the AreR-regulated σ54-dependent promoter in *Acinetobacter sp.* strain ADP1. Applied and Environmental Microbiology. 2003;69(9):5627-35.

38. Colton DM, Stoudenmire JL, Stabb EV. Growth on glucose decreases cAMP□CRP activity while paradoxically increasing intracellular cAMP in the light□organ symbiont *Vibrio fischeri*. Molecular Microbiology. 2015;97(6):1114-27.

CHAPTER 4

MODULAR EXPANSION: INTRODUCTION OF A VERATRATE AND ISOVANILLATE DEGRADATION PATHWAY TO *ACINETOBACTER BAYLYI* ADP1

Introduction

The integration, expression, and optimization of a segment of foreign DNA in the *A. baylyi* ADP1 chromosome was demonstrated in **Chapter 2**. In those studies, the 7-kbp region of foreign DNA encoded an alternative pathway for the catabolism of an aromatic compound that *A. baylyi* already consumes, protocatechuate. I next sought to insert and optimize a different foreign pathway in the *A. baylyi* chromosome. Here, genes for the catabolism of two aromatic compounds that *A. baylyi* does not naturally consume, veratrate and isovanillate, were introduced.

Veratrate catabolism has been studied in two bacteria, *Comamonas testosteroni* strain BR6020 and *Pseudomonas putida* CSV86 (1, 2). In both organisms, researchers identified two sets of two-component O-demethylase enzymes that sequentially convert veratrate to protocatechuate. The veratrate demethylase enzyme, called VerAB (sometimes called IvaAB), is proposed to remove the methyl from the 4-methoxy group of the aromatic acid. The substrate

specificity of this enzyme also allows organisms to consume isovanillate either as an intermediate of veratrate catabolism or as the initial carbon source (**Figure 4.1**).

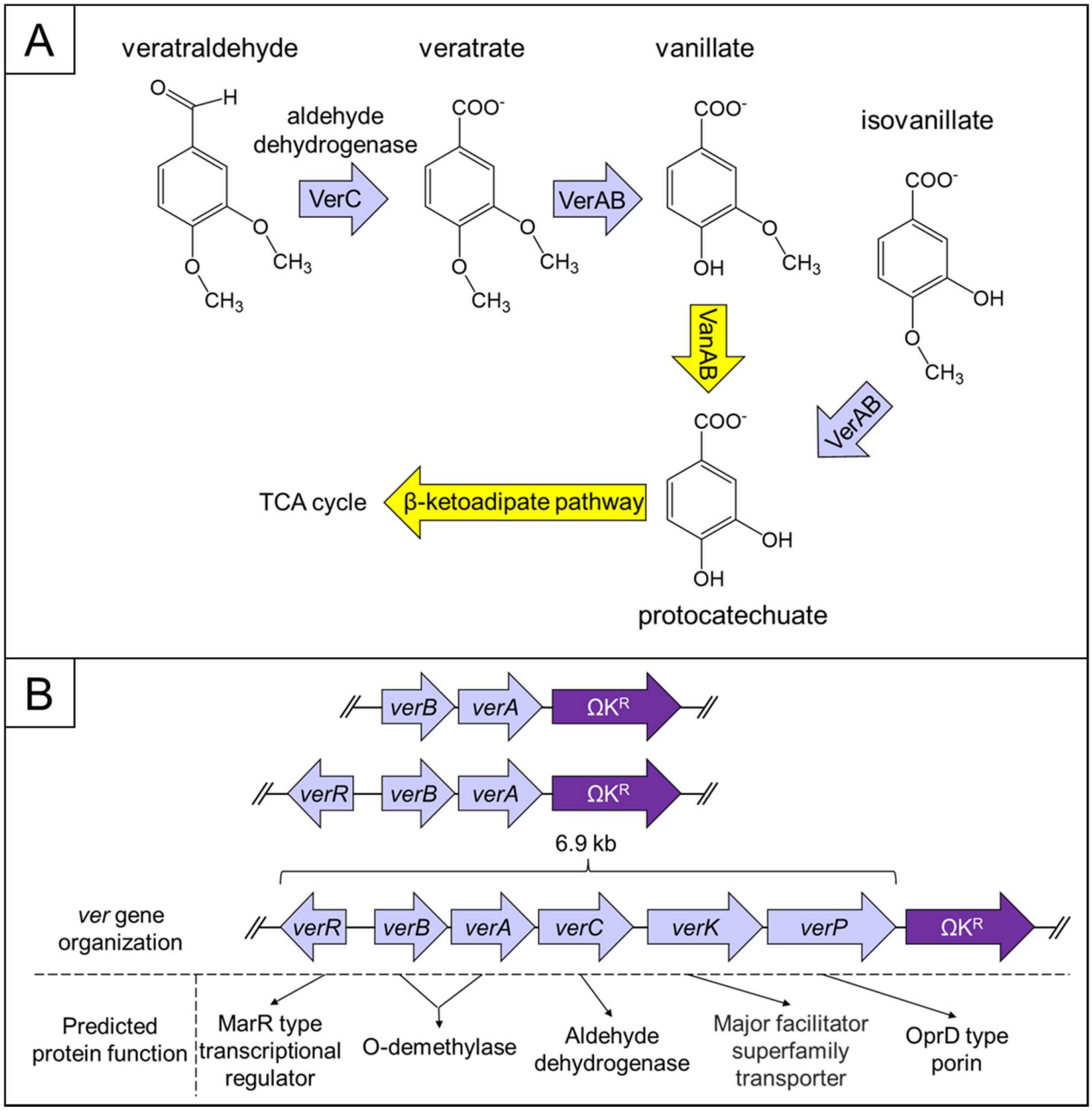


Figure 4.1: Veratrate degradation pathway and gene organization

Panel A: Predicted veratrate degradation scheme with foreign pathway in *A. baylyi* ADP1. Purple arrows represent enzymatic steps encoded by *P. putida* CSV86 *ver* genes. Yellow arrows represent enzymatic steps encoded by native ADP1 genes. Panel B: The 6.9 kb region (*verR*

through *verP*) shows the genetic organization from the native host, *Pseudomonas putida* CSV86. This organization was maintained for integrating the *ver* genes in *A. baylyi*. To select the introduction of *ver* genes into the ADP1 chromosome, the ΩK^R drug-resistance cassette was inserted downstream the *ver* genes in different constructs.

The second two-component O-demethylase enzyme is VanAB, a vanillate O-demethylase, studied in ADP1. VanAB removes the methyl from the 3-methoxy group of the aromatic acid (**Figure 4.1**) (3, 4). VanB functions as the oxidoreductase subunit of the enzyme (5). It contains FAD binding and [2Fe-2S] ferredoxin domains that transfer an electron from an electron-carrier co-enzyme to the Rieske-type [2Fe-2S] cluster of the VanA monooxygenase subunit (5-7). In VanA, the electron is then passed to an Fe^{2+} site, where oxygen binding and subsequence demethylation of the bound substrate can occur (5, 7-9). The similarity between the predicted VanAB from *P. putida* CSV86 and ADP1 is evident in a sequence alignment (**Figure 4.2**).

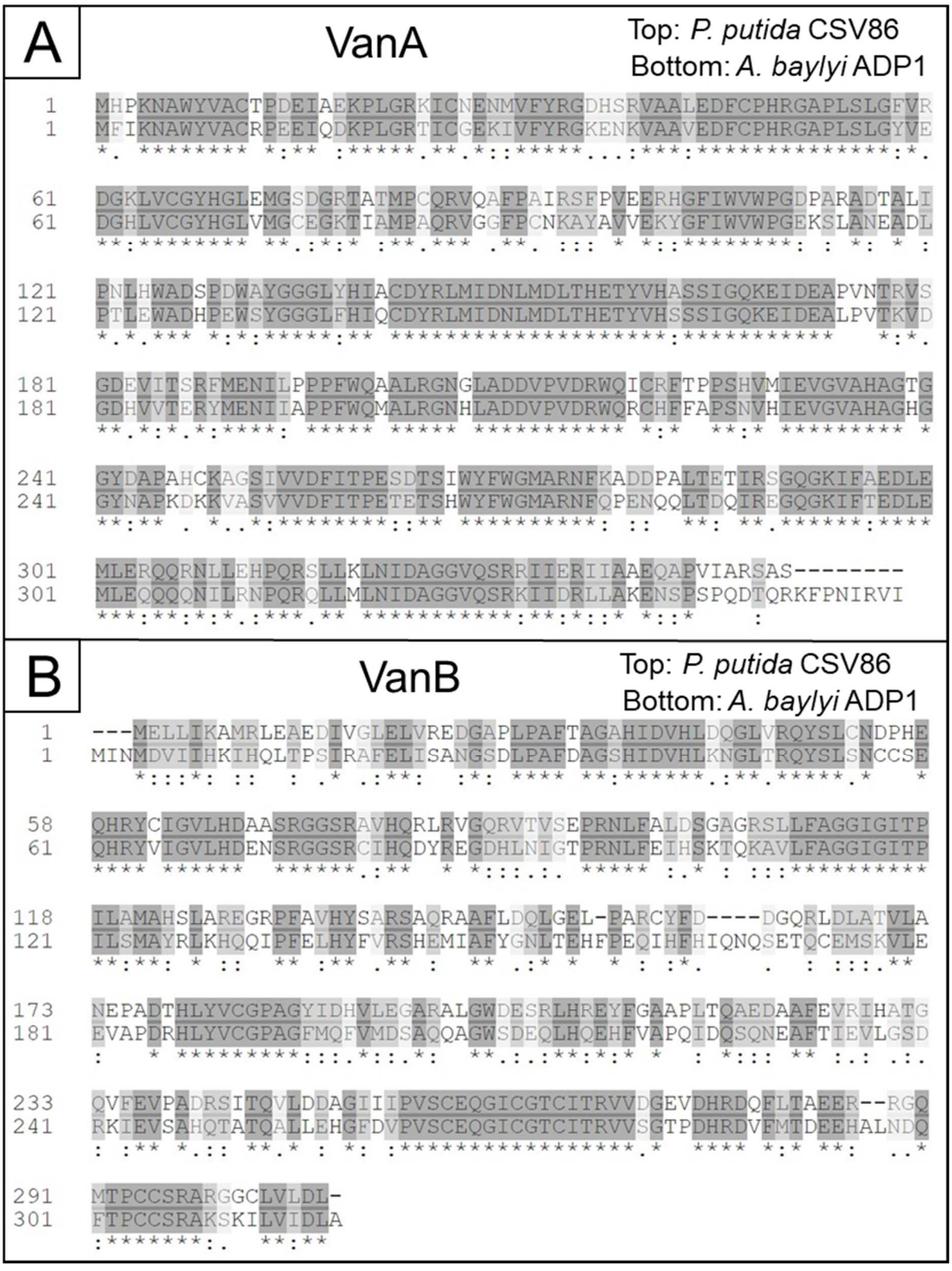


Figure 4.2: Alignment of VanA and VanB from *P. putida* CSV86 and *A. baylyi* ADP1

Panel A: Alignment of *P. putida* CVS86 VanA (UniProt ID L1LSN6) and *A. baylyi* ADP1 VanA (UniProt ID Q6FDI6) (10). Panel B: Alignment of *P. putida* CVS86 VanB (UniProt ID L1LS29) and *A. baylyi* ADP1 VanB (UniProt ID O24840) (10). Markings underneath each residue and highlighted color in the alignment indicate identical residues (* dark grey), residues with strongly similar properties (: medium grey), or residues with weakly similar properties (. light grey), as defined on https://www.uniprot.org/help/sequence-alignments (10).

The VerAB system likely functions in a similar manner since both components are homologs of their VanAB counterparts and both enzyme systems catalyze analogous demethylation reactions. Phylogenetic studies report information about the relationships and distribution of vanillate and veratrate O-demethylase components (2). Both A and B subunits cluster into two distinct groupings based on the substrate, veratrate or vanillate (2). Sequence alignments are shown in **Figure 4.3**. While the *A. baylyi* VanAB and VanAB homologs in other bacteria are reported to have wide substrate specificity (3, 11, 12), these bioinformatic and phylogenic findings point to distinct roles for the VerAB and VanAB O-demethylases.

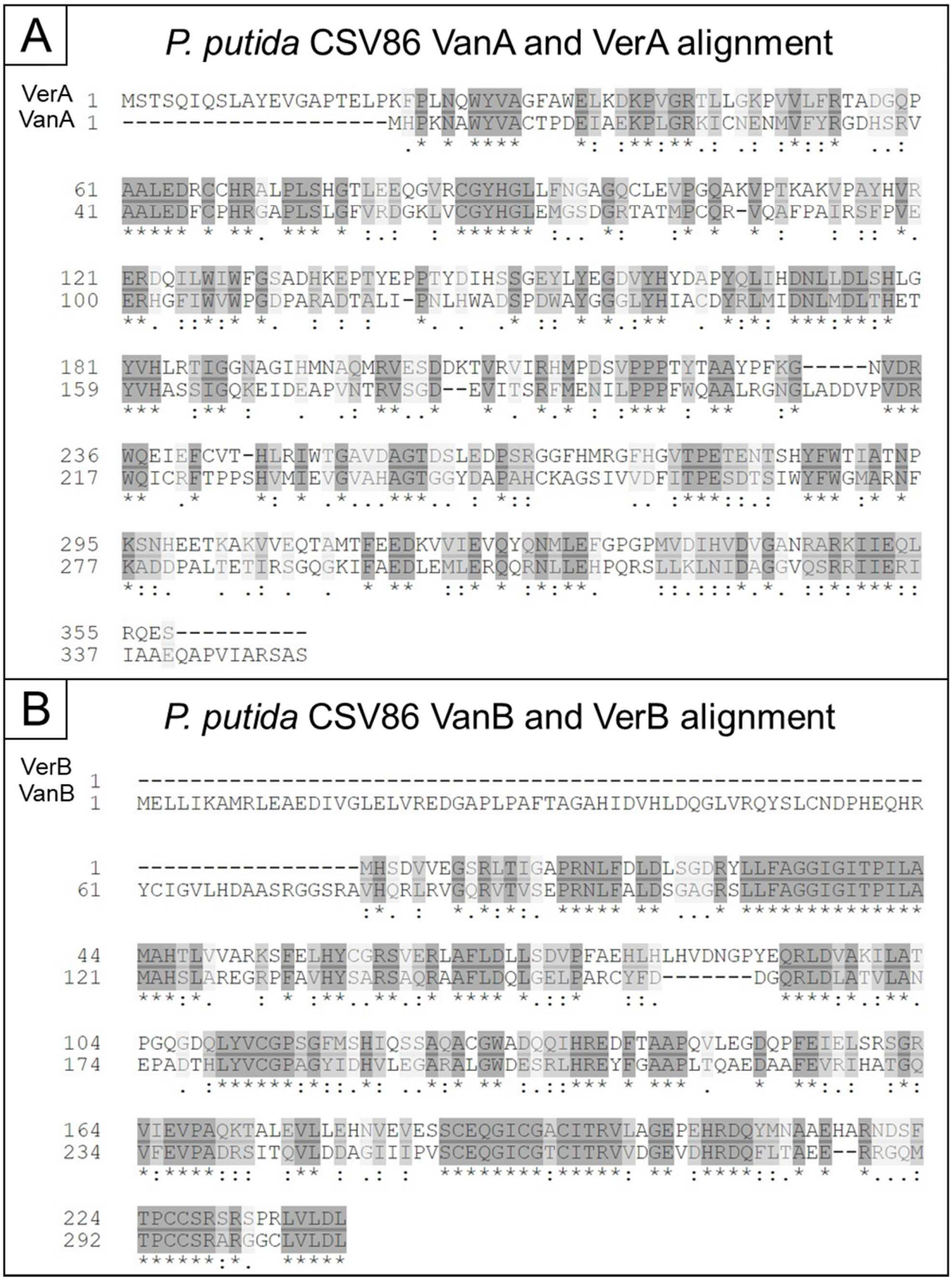

A
P. putida CSV86 VanA and VerA alignment
VerA
VanA
B
P. putida CSV86 VanB and VerB alignment
VerB
VanB

Figure 4.3: Alignment of *P. putida* CSV86 VerA to VanA and VerB to VanB

Panel A: Alignment of *P. putida* CVS86 VerA (UniProt ID L1LYA4) VanA (UniProt ID L1LSN6) (10). Panel B: Alignment of *P. putida* CVS86 VerB (UniProt ID L1LYD0) and VanB (UniProt ID L1LS29) (10). Markings underneath each residue and highlighted color in the alignment indicate identical residues (* dark grey), residues with strongly similar properties (: medium grey), or residues with weakly similar properties (. light grey), as defined on https://www.uniprot.org/help/sequence-alignments (10).

O-demethylase enzymes are particularly important for microbial degradation of lignin, a highly abundant, natural plant polymer. The individual components making up the lignin polymer are often O-methylated, and initial catabolic steps for the degradation of these lignin monomers in bacteria typically involve demethylating the aromatic ring substituents to hydroxyl groups (13). The EASy method of laboratory evolution detailed in **Chapter 1** had been used to alter one such O-demethylase, GcoAB, for guaiacol degradation (14, 15). Using EASy with foreign *gcoAB* genes from *Amycolatopsis sp*. ATCC 39116 resulted in an altered GcoAB enzyme that allowed *A. baylyi* ADP1 and *P. putida* KT2440 to grow on guaiacol as the sole carbon source (14). As described here, I sought to use EASy to enable *A. baylyi* to grow on veratrate using a pathway identified in a different organism, *P. putida* CSV86.

Transcriptional studies in *P. putida* CSV86, showed that *verBACKP* are encoded as two transcriptional units, *verBAC* (CSV86_16960, CSV86_16965, CSV86_16970) and *verKP* (CSV86_16975, CSV86_16980) (**Figure 4.1**) (2). *verC* encodes an aldehyde dehydrogenase, *verK* encodes a transporter, and *verP* encodes a porin. A MarR type transcriptional regulator is encoded by *verR* (CSV86_16955), located divergently upstream *verB* (**Figure 4.1**). The gene

organization of *verR* to *verBAC* is typical of MarR type transcriptional regulators which typically function to repress transcription of their target operons in the absence of an effector ligand (16). It has not been experimentally verified that VerR, VerK, and VerP are involved in veratrate regulation and uptake respectively, but their proximity to *verBA* suggests this connection.

In these studies, we integrated codon optimized genes based on the *Pseudomonas putida* CSV86 veratrate degradation pathway in the chromosome of *A. baylyi.* Growth on veratrate or isovanillate as the carbon source was selected using the EASy method of laboratory evolution. ADP1-derived strains thus acquired the ability to grow on two new aromatic carbon sources. Subsequent whole genome re-sequencing of evolving strains led to potential insights into the regulation of the veratrate pathway that current experiments are addressing.

Results

ADP1 does not grow on veratrate as the sole carbon source

When tested for growth on minimal medium with 1, 3, or 5 mM veratrate or isovanillate as the sole carbon source, *A. baylyi* ADP1 did not grow. This result is presumably due to the lack of veratrate catabolic genes in *A. baylyi*. We thus sought to integrate one of the known veratrate degradation pathways into the *A. baylyi* chromosome. Both sets of characterized veratrate catabolic genes have higher G+C content than *A. baylyi* genes (1, 2). To avoid issues with G+C content or codon usage, each *ver* gene from *Pseudomonas putida* CSV86 was optimized by Dr. Jan Mrázek to match those of the highly expressed ribosomal protein genes in ADP1.

Introduction of a veratrate degradation pathway from *Pseudomonas putida* CSV86

For introduction of the *ver* genes into *A. baylyi*, we considered which *ver* genes to include and where to put them in the chromosome. Maintaining the genetic organization from *P. putida* could preclude the need to alter promoter regions or include other synthetic constructs. Therefore, we chose to keep the intergenic regions of the *ver* cluster unchanged. To maximize the likelihood that the integration and adaptation of the *ver* genes is successful in *A. baylyi*, we chose to make strains containing all or some of the genes.

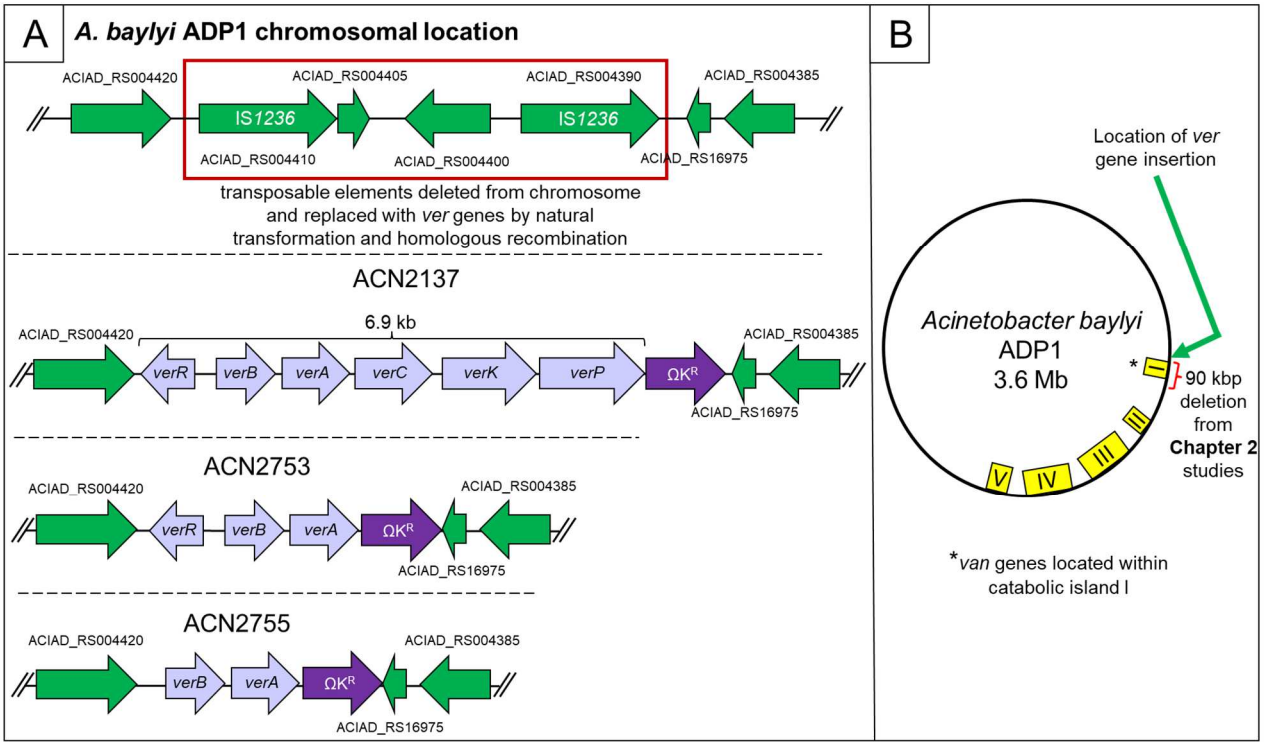


Figure 4.4: Chromosomal integration of *ver* genes into *A. baylyi*

Panel A: Chromosomal region in *A. baylyi* (green arrows) where *ver* genes and a kanamycin resistance cassette (purple arrows) were introduced. Transposable elements removed from chromosome by homologous recombination in red box. Panel B: Depiction of *A. baylyi* chromosome with relative locations of the catabolic islands (yellow boxes), location of *ver* gene

insert (green arrow), location of van genes (*), and 90 kbp deletion from **Chapter 2** (red bracket).

The *ver* genes, with an ΩK^R cassette for selection, were integrated into the *A. baylyi* chromosome through natural transformation and homologous recombination. In strain ACN2137, all six *ver* genes were introduced (*verRBACKP)* (**Figure 4.4 Panel A**). Strain ACN2755 has only the O-demethylase encoding genes (*verBA)*, and strain ACN2753 has the O-demethylase encoding genes along with the gene encoding the predicted regulator (*verRBA*) (**Figure 4.4 Panel A**).

Since the foreign VerAB and the native VanAB are likely to be required for the catabolism of veratrate, we chose to integrate the *ver* genes into the *A. baylyi* chromosome near the *van* genes in catabolic island I (5). During my studies detailed in Chapter 2, a large 90 kbp deletion of the chromosome, including the *van* genes, was discovered that likely resulted from transposable elements, IS*1236*, located next to the large deletion (17, 18). To avoid the instability in this region of the genome often caused by the transposable elements located near catabolic island I (18), the veratrate degradation genes were inserted in place of the *IS1236* transposable elements and the composite transposon they flank, ACIAD_RS004390, ACIAD_RS04400, ACIAD_RS04405, and ACIAD_RS04410 (6, 7) (**Figure 4.4 Panel B**). An ΩK^R cassette was also integrated into the chromosome with the *ver* genes for antibiotic selection during chromosomal integration (**Figure 4.4 Panel A**).

Selection of ver$^+$ and isov$^+$ *A. baylyi* ADP1-derived strains

After undergoing allelic replacement in the *A. baylyi* chromosome to introduce the *ver* genes, none of the *A. baylyi*-derived strains grew on veratrate or isovanillate as the sole carbon source. Possible reasons for the failure to grow include low-level heterologous expression, unbalanced metabolic flux, inappropriate regulation, unstable or misfolded proteins, and/or problems with transport. We therefore sought to use the previously described EASy method of laboratory evolution used for the PCA-4,5 pathway in **Chapter 2**.

Initial laboratory evolution experiments focused on strain ACN2137 containing all *ver* genes, *verRBACKP*. A summary of the EASy amplification of this strain is shown in **Figure, 4.5**.

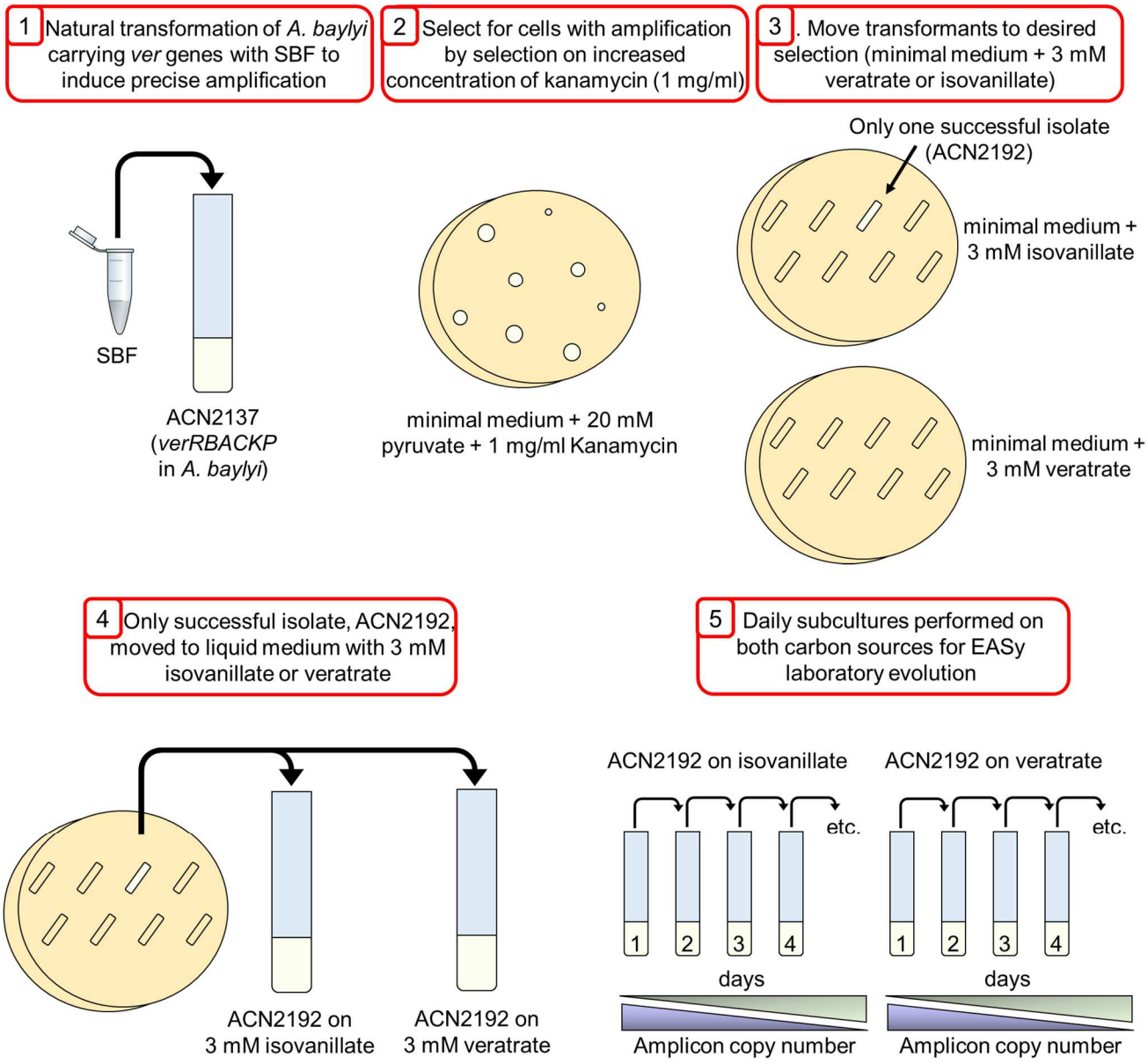


Figure 4.5: EASy laboratory evolution of *A. baylyi* strain with *ver* genes, ACN2192

Overview of EASy laboratory evolution of ACN2192 derived from ACN2137. Steps summarized in red boxes. Evolution of ACN2192 was carried out on both veratrate and isovanillate as sole carbon sources.

Briefly, a precise duplication of the amplicon including the *ver* genes was initiated by natural transformation of the strain with the synthetic bridging fragment (SBF), pBAC1604 linearized by restriction digest with AatII. Successful transformants with induced amplifications were selected for their increased resistance to kanamycin. Colonies were then moved to medium containing either 3 mM veratrate or 3 mM isovanillate as the sole carbon source. For the *A. baylyi* strain containing *verRBACKP*, ACN2137, only one successful strain gained the ability to grow on the new aromatics, ACN2192. It was first isolated growing on isovanillate, designated as isov$^+$, but after isolation, it grew on both isovanillate and veratrate, ver$^+$.

A. baylyi strains containing some of the *ver* genes, *verRBA* (ACN2753) and *verBA* (ACN2755), also were subjected to chromosomal amplification. Only three isolates were successful from ACN2753, one isolated on isovanillate and the other two isolated on veratrate. No successful ver$^+$ or isov$^+$ isolates have been discovered from the strain with only *verBA*, ACN2755. Experiments with these strains containing only some of the *ver* genes were stopped when I was no longer able to go into the laboratory during the initial COVID-19 outbreak. Characterization of only the ver$^+$ and isov$^+$ strain containing all *ver* genes, ACN2192, was continued.

Long-term culturing of a ver$^+$ and isov$^+$ *A. baylyi* strain containing *verRBACKP* (ACN2192)

The isov$^+$/ver$^+$ strain, ACN2192, was then moved to liquid culture and EASy evolution was continued by daily subcultures of the populations (**Figure 4.5**). Cells were cultured on each aromatic compound in separate experiments. The amplicon copy number was assessed periodically during long-term culturing by quantitative (q) PCR. As was done in **Chapter 2**, the copy number of a small fragment of the K^R gene in the amplicon, designated F87, was used as an

approximation of the gene dosage for the amplified region, since typically the tandem array increases or decreases by integral amplicon units (19).

The amplicon copy number of ACN2192 evolving on veratrate or isovanillate is depicted in **Figure 4.6**. On each carbon source, the population starts out with a high F87 copy number, averaging 50 on veratrate and 71 on isovanillate. Consistent with previous EASy evolution experiments, the F87 copy number quickly decreased, presumably as beneficial mutations arose and the need for extra copies of the foreign genes decreased (8, 9). After hitting a low of 10 copies at week 3, the F87 copy number increased again on both carbon sources and remained elevated through 21 weeks of evolution. In the PCA-4,5 evolution described in **Chapter 2**, even though strains didn't lose all F87 copies, they maintained a lower copy number throughout long-term culturing. However, in ACN2192 growing on veratrate or isovanillate, the F87 copy number increased again after the initial decrease in copy number. Populations grew on isovanillate or veratrate with the lowest average copy number of the *ver* genes at week 3 (10 copies). So, a single colony was selected from each population at week 3, named ACN2334 (cultured on veratrate) and ACN2439 (cultured on isovanillate), and sent for whole genome re-sequencing.

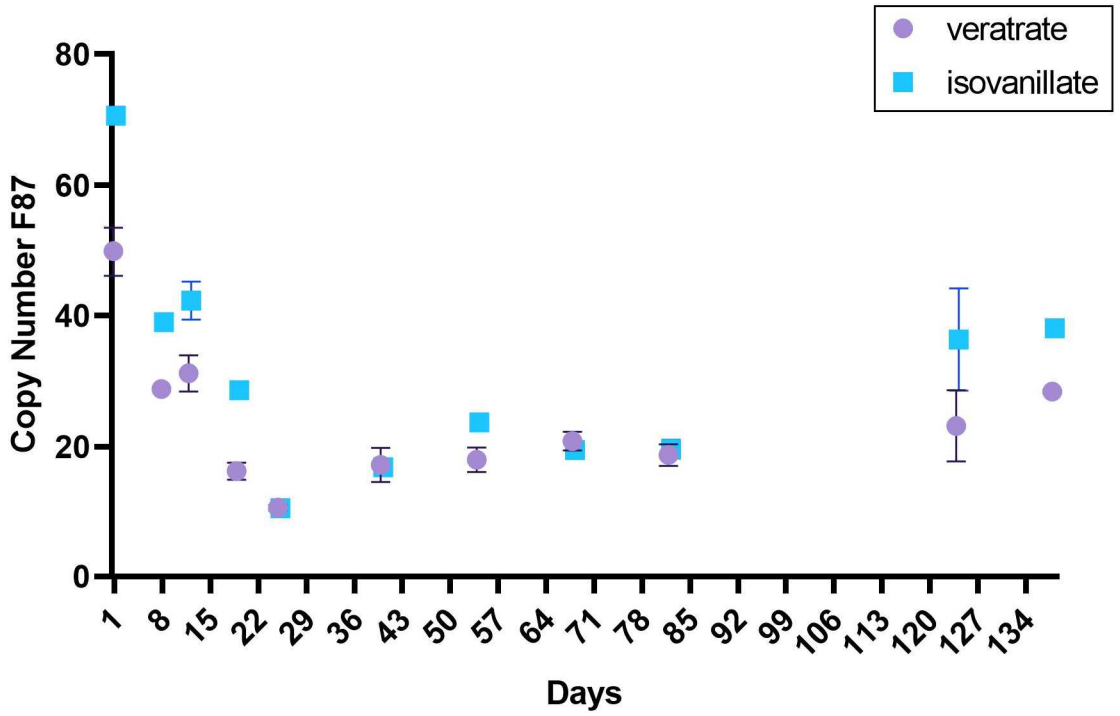


Figure 4.6: Copy number of F87 during culturing on veratrate or isovanillate

Long-term culturing of ACN2192, containing all *ver* genes, the copy number of the F87 fragment was analyzed periodically by qPCR. Purple dots represent the copy number of the population growing on veratrate as the sole carbon source, and blue squares represent the copy number of the population growing on isovanillate as the sole carbon source. Error bars represent the standard deviation for 4 technical replicates from a genomic DNA extraction of the whole evolving population.

WGS of *A. baylyi* strains after long-term culturing on either aromatic compound

Results of Illumina whole genome re-sequencing of ACN2334, *A. baylyi* with *verRBACKP* cultured on veratrate, and ACN2439, *A. baylyi* with *verRBACKP* cultured on isovanillate, are shown in **Tables 4.1 and 4.2**. WGS reads were mapped to the *A. baylyi* reference genome NCBI entry NC_005966. WGS analysis was done through the Breseq pipeline

(20). ACN2334, cultured on veratrate, had five mutations. The significance of four of these mutations is not clear. The fifth mutation was a small deletion in ACIAD_RS04520, the vanillate transporter gene, *vanK*, resulted in a frameshift. This mutation causes a premature truncation in coding sequence of VanK, which could be affecting veratrate transport.

Whole genome re-sequencing of the strain cultured on isovanillate, ACN2439, revealed none of the five mutations which occurred in the veratrate cultured strain. Even though ACN2334 and ACN2439 started from the same population, ACN2192, five mutations accrued only when the population was cultured on veratrate and not isovanillate.

Table 4.1: WGS revealed mutations in *A. baylyi* strain cultured on veratrate

Strain	Gene locus ID (name)	Position*	Predicted description	Mutation type	DNA change	Protein effect
ACN2334 (cultured on veratrate)	ACIAD_RS02750	592,433 – 592,434	glycosyl transferase	deletion	Δ2 bp	frameshift
	ACIAD_RS04520 (*vanK*)	968,748	MFS superfamily vanillate transporter	deletion	(G)6→5	frameshift
	ACIADtRNASer_34 → / ← ACIAD_RS05325	1,148,146	tRNA□Ser/efflux RND transporter periplasmic adaptor subunit	bp change	C→T	N/A
	ACIAD_RS13200 (*rpoD*)	2,854,879	RNA polymerase sigma factor RpoD	bp change	C→T	R116C
	ACIAD_RS13730 (*rho*)	2,968,318	transcription termination factor Rho	bp change	G→A	A112T

* Position refers to positions on the ADP1 chromosome in NCBI entry NC_005966.

Larger chromosomal rearrangements are reported through the Breseq pipeline by their novel junction (**Table 4.2**). A new chromosomal junction was identified in both strains after long-term culturing, ACN2334 (cultured on veratrate) and ACN2439 (cultured on isovanillate) (**Figure 4.7**). The new junction created a new amplification in the chromosome that was similar in size and location to the engineered amplification (induced by the SBF) including the

verBACKP operon and the ΩK^R cassette. However, the outer bound ended between *verB* and *verR,* excluding *verR* from the amplicon (**Figure 4.7**). When cultured on veratrate, this rearrangement took over the population leaving no copies of the original designed amplicon present. There is thus only a single copy of *verR* in the chromosome. However, when cultured on isovanillate, both junctions are still present. Because Illumina WGS uses short sequence reads, ≥300 bp, it is unclear how these amplifications are distributed within the chromosome. This result may suggest a stronger evolutionary pressure to the cells when grown on veratrate compared to isovanillate to lower the transcription of or product formation of VerR. However, the stochastic nature of spontaneous mutations makes it difficult to draw conclusions from a comparison of only two cultures.

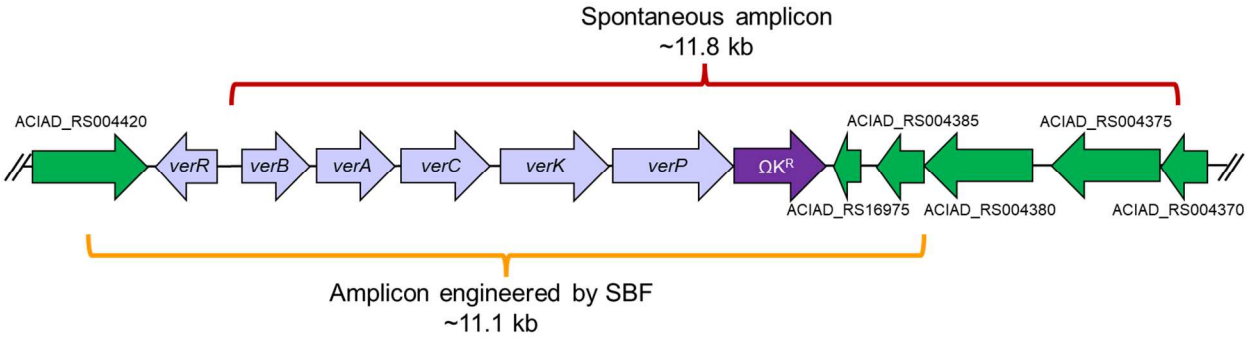


Figure 4.7: Original and spontaneous amplicons including *ver* genes

Green arrows depict *A. baylyi* loci. Purple arrows depict *ver* genes and the kanamycin resistance cassette. Orange bracket indicates the relative bounds of the amplicon engineered by the SBF (pBAC1604 linearized at AatII site). Red bracket indicates the relative bounds of the spontaneous amplification.

Table 4.2: WGS revealed junctions

Evolved strain	Predicted description	Junction/ location of insertion*
ACN2334 (cultured on veratrate)	New amplification junction: ACIAD_RS04370 and between *verR* and *verB*	TGAACATCTAGTCCATGCTCGAAGTTTGTG
ACN2439 (cultured on isovanillate)	New amplification junction: ACIAD_RS04370 and between *verR* and *verB*	TGAACATCTAGTCCATGCTCGAAGTTTGTG
	Designed junction induced by SBF: between ACIAD_RS04380 and ACIAD_RS04385 and ACIAD_RS04420	AAAATAATGACATATGTATGAAGAAACAAT

Future Directions: characterizing VerR

Based on its sequence, VerR belongs to the MarR family of transcriptional regulators. Canonically, MarR regulators suppress transcription of their target operons in the absence of their effector ligands (16, 21, 22). De-repression occurs when the effector is present and binds to the regulator, allowing transcription of the target operon to proceed (**Figure 4.8 Panel B**). However, actual regulation by MarR type regulators may vary and needs to be determined for each protein experimentally. Based on the canonical model of MarR type regulation and the results of the ACN2192 long-term culturing experiments, we hypothesize that isovanillate, and not veratrate may be the true effector of VerR (**Figure 4.8 Panel C**).

Our reasoning stands that when cells are grown on isovanillate, the true VerR effector, the effector is thus present to de-repress expression of the *ver* operon. However, when cells are grown on veratrate, de-repression is not induced since veratrate is not the true effector of VerR, resulting in minimal expression of the *ver* genes. This hypothesis can be tested by Surface Plasmon Resonance (SPR) experiments analyzing veratrate or isovanillate as potential inducers of VerR. SPR can measure receptor-ligand interactions and kinetics through very sensitive detection of small molecules binding to proteins fixed to the SPR surface (23-26).

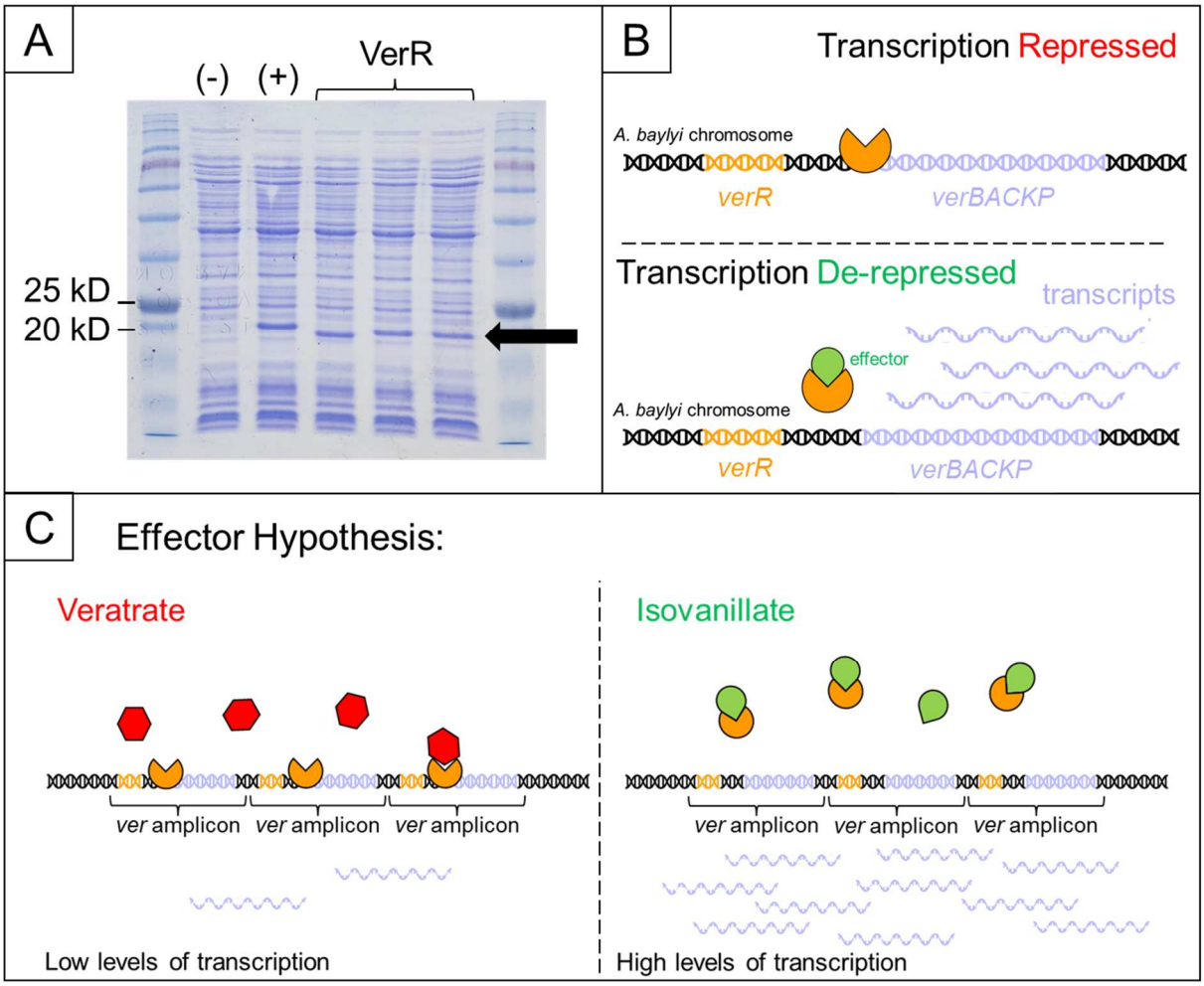


Figure 4.8: VerR regulatory hypothesis

Panel A: SDS-PAGE gel of *in vitro* expression of VerR. Black arrow points to band around the expected size, 23 kDa, for VerR only present in reactions that included pBAC1803 (for VerR expression). Panel B: Schematic of canonical MarR transcriptional regulation. VerR is represented by orange wedged circle, an effector is represented by green raindrop shapes, and DNA is represented by double helix figure (*ver* operon in purple DNA). Panel C and D: predicted VerR regulation where veratrate, red hexagon, does not bind VerR, orange wedged-

circle (Panel C), but isovanillate, green raindrop, does bind VerR, orange wedged-circle, and derepresses transcription of *ver* operon, purple DNA.

Expressing and purifying VerR

To test our hypothesis concerning the inducer specificity of VerR by SPR, I need to purify the protein so it can be bound to the SPR surface and tested for a response to veratrate or isovanillate ligand binding. The *verR* coding sequence was put into the *E. coli* expression plasmid pET-28b. Expression should result in a C-terminal his tag affixed to VerR. The his tag would be used to purify the protein and then to fix it to the SPR sensor.

I first attempted to express VerR using autoinduction medium to grow and induce expression in *E. coli* BL21 (27, 28). However, no product around the expected size of VerR, 23 kilo Daltons, could be detected on SDS-PAGE gels of the cell lysates. Next, I attempted to induce expression of *verR* in *E. coli* KRX with IPTG and rhamnose induction. However, I still could not detect any VerR product on SDS-page gel of the cell lysates. To determine that the plasmid pBAC1803 was capable of expressing VerR, I tried a commercial method of *in vitro* protein expression, NEBExpress®. Using the recommended protocol for this method of protein expression resulted in the formation of a product around the expected size of VerR (**Figure 4.8 Panel A**). This result suggests that the pBAC1803 is expressing VerR.

Discussion

Selection of *A. baylyi* strains utilizing veratrate or isovanillate as sole carbon sources

In these studies, I generated *A. baylyi* strains capable of growing on two new aromatic carbon sources, veratrate and isovanillate, through the introduction and chromosomal

amplification of foreign genes, *ver* genes, from *Pseudomonas putida* CSV86. However, only four *A. baylyi* strains capable of growing on veratrate and/or isovanillate were isolated through EASy laboratory evolution. In past EASy experiments, other researchers and I isolated many successful EASy amplification mutants growing through various foreign genes (**Chapter 2**) (Chantel Duscent-Maitland unpublished data) (29, 30). Problematic aspects with this pathway are likely to involve transport, enzyme substrate specificity, and/or regulation. These topics are discussed below along with possible directions for future research.

Transport of veratrate and isovanillate in *A. baylyi*

The consumption of aromatic acids such as those being studied here typically require specific transport proteins. This requirement is reflected in the large amount of genomic DNA in ADP1 that is devoted to encoding transporters (31). In **Chapter 3**, multiple transporters were shown to contribute to the uptake of amino acids. For this reason, the genes encoding potential transport proteins for veratrate (*verK* and *verP*) were introduced and amplified in the chromosome along with the genes encoding the demethylase. However, it is also known that problems can be caused if the level of transport proteins is too high or too low. During my doctoral studies, I worked to express transport proteins from synthetic promoters, for example for *pcaK* and for *aspT* (32). Firsthand experience illustrated that that the promoter strength was important for growth. In recent EASy studies of terephthalate degradation, the major outcomes of the evolutionary studies were changes to transport proteins. In this example, the foreign genes encoding transporters became mutated and variants of native ADP1 transporters were selected (30).

Balanced expression of foreign and native enzymes for veratrate catabolism

For cells to grow on veratrate or isovanillate, their degradation must be coordinated with vanillate degradation. Both native and foreign genes must be expressed at the right time, and metabolites must be produced at appropriate levels. If too much or too little of any intermediary compound is made it might affect induction of additional proteins, producing enough of a subsequent metabolite for growth. Additionally, in *P. putida*, the *van* genes were expressed when the strain was grown on veratrate and vanillate, compounds with a methyl group on the *meta* position of the aromatic ring. Lastly, since *A. baylyi* can catabolize many aromatic compounds, competitive inhibition could be occurring. Another enzyme acting on veratrate or isovanillate could be producing an unusable end product and preventing *A. baylyi* from growing on these aromatic compounds.

Differences in evolutionary trajectory for the same initial *A. baylyi* population when cultured on different aromatics

Because ACN2192 was the only successful amplification mutant possessing *verRBACKP*, I expected to find a mutation that was present in this population that allowed for this population to grow on veratrate and isovanillate where all other *A. baylyi* cells failed. WGS revealed that the only mutation present after the strain was evolved on isovanillate was a rearrangement of the amplicon resulting in two separate amplicons. One amplicon was the original engineered junction induced by the synthetic bridging fragment, but the second was a new spontaneous amplicon that excluded *verR*. Therefore, there were more copies of *verBACKP* than *verR* in the chromosome. When the population was cultured on isovanillate, both amplicons

occurred in about equal frequency. However, on veratrate, the cells lost all copies of the original amplicon, leaving only one copy of *verR* in the chromosome.

Insights into regulation of veratrate and isovanillate catabolism by VerR

MarR type transcriptional regulators, the family of regulators VerR belongs to, typically repress transcription from target operons by binding DNA and blocking RNA polymerase. When their effector is present and binds, the regulator's conformation changes, and it no longer binds to the target DNA. Expression is then de-repressed, and transcription can occur (**Figure 4.4 Panel B**). In these studies, I hypothesized that because ACN2192 evolved to have only one copy of *verR* on veratrate but maintained multiple copies on isovanillate, isovanillate may be the true effector of VerR (**Figure 4.4 Panel C**). This hypothesis will be tested by Surface Plasmon Resonance (SPR) by testing the binding efficiency of various effectors, likely veratrate and isovanillate, to VerR.

In summary, the research done here has provided a strong foundation for continuing to study the catabolism of veratrate in *A. baylyi*. In future experiments, the ver^+ and $isov^+$ strains with only *verBA* or *verRBA* can be submitted to long-term culturing on these carbon sources and subsequently evaluated for adaptive mutations. The regulatory hypotheses regarding veratrate catabolism in *A. baylyi* presented here can also be evaluated in future experiments.

Experimental Procedures

Media and growth conditions

Defined ADP1 medium was used for culturing all *A. baylyi* strains as previously described (33). This medium was supplemented with 2 mM veratrate or isovanillate and

incubated at 30 °C or pyruvate and incubated at 37 °C with liquid cultures shaking at 250 rpm, unless otherwise indicated. For plasmid cloning, *E. coli* XLI Blue or DH5α (Stratagene) was grown in Luria-Bertani medium (LB) (34). All cultures were incubated at 37 °C with liquid cultures shaking at 250 rpm, unless otherwise indicated. *E. coli* BL21 used for VerR protein expression was cultured on autoinduction media as previously described (27). *E. coli* KRX used for VerR protein expression was cultured on terrific broth (TB) (34). Expression was induced with final concentrations of 1 mM Isopropyl β- d-1-thiogalactopyranoside (IPTG) and 1% rhamnose. Antibiotics were used at final concentrations of 25 μg/ml for kanamycin, 12.5 μg/ml each for streptomycin and spectinomycin, and 150 μg/ml for ampicillin. Increased concentration of kanamycin used to select for EASy amplification mutants was 1 mg/ml.

Strain and plasmid construction

DNA used for plasmid constructed was PCR-amplified by using the high-fidelity polymerases PrimeSTAR Max (Takara Biosciences) or Phusion (New England Biosciences). All primers used can be found in the primer table (**Table 4.5**). Plasmids were constructed by overlapping sequence assembly by NEBuilder (New England Biosciences) or in *E. coli* XL1-Blue or DH5α competent cells (Stratagene) as previously described (cloning should be easy). Standard restriction digest followed by ligation (Quick Ligation Kit, New England Biosciences) was also used for plasmid construction. Plasmids were confirmed by DNA restriction mapping and/or regional DNA sequencing (Eton Biosciences or Eurofins Genomics). Plasmids are described further in the supplemental plasmid table (**Table 4.4**).

All chromosomal changes in ADP1 were constructed by allelic replacement as previously described (35). Strains were confirmed by PCR and/or regional DNA sequencing. All strains are further described in the strain table (**Table 4.3**).

Laboratory evolution by EASy

A. baylyi derived strains carrying the *ver* genes were evolved growth by the EASy method. Strains were transformed with donor DNA called the synthetic bridging fragment (SBF), linearized pBAC1604 in this study. This fragment induces precise, defined duplications in the chromosome that can undergo subsequent amplification (19, 36). Transformants were then exposed to 1 mg/ml of kanamycin. This concentration of antibiotic kills cells carrying a single copy of the ΩK^R cassette but not cells that underwent chromosomal amplification (14). Therefore, this step selects for strains with the desired chromosomal amplifications including ΩK^R and the *ver* genes. Strains that grew on the high concentration of kanamycin were then moved to liquid ADP1 minimal media with veratrate or isovanillate as the sole carbon source at 30 °C. Strains were then cultured with constant selective pressure. For the first week 10% inoculations were used for daily subcultures. This daily inoculation was then tapered down over multiple days to 0.1% for the rest of the evolution. The number of copies of the amplicon in evolving strains was analyzed by quantitative (q) PCR as described below.

qPCR copy number analysis

The qPCR method used to assess copy number during EASy experiments was previous described (37). Primers for qPCR are listed in the primer table (**Table 4.5**). Standard curves used genomic DNA from strains with a single copy of the amplicon (ACN2137) for both *rpoA* and

ΩK^R. The ratio of calculated concentration for ΩK^R relative to *rpoA* was presumed to equal the copy number of that fragment.

Next generation sequencing and analysis

Genomic DNA was isolated from ADP1 derived strains and sonicated to approximately 500 bp size fragments in a total of 500 ng. Genomic libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) including end-repair, adaptor ligation, and addition of an individual i5/i7 index primer. Sequencing was performed on the Illumina NextSeq500 instrument at the Georgia Genomics and Bioinformatic Core at the University of Georgia.

Breseq version 0.35.4 was used to analyze WGS data on the UGA GACRC High-Performance Computing Cluster (Sapelo2), Linux based. Version 2.4.1 of bowtie2 and version 4.0.0 of R were used in the pipeline. The consensus mode was used with the default consensus frequency cutoff of 0.8 and polymorphism frequency cutoff of 0.2. The reference genome used for ADP1 was from NCBI entry GenBank CR543861. Differences between our laboratory ADP1 sequence and the NCBI entry can be found in **Table 2.10**.

Protein purification

Plasmid pBAC1803 was used to purify VerR. Initial protein expression and purification method was followed as previously described (28). When this method failed to result in the expression of VerR, another method was attempted. In this second attempted method, pBAC1803 was first transformed into chemically competent *E. coli* KRX by heat shock. A 5 mL culture were grown on LB + kanamycin overnight then diluted 1% into 125 mL TB in a 500 ml

baffled flask and shaken at 37 °C until an OD_{595} of 0.7-0.8 is reached. Cultures are then moved to shaking at 30 °C until an OD^{595} of 1.0-1.5 is reached. Next, 25 ml of culture is removed and remains uninduced. To induce the remaining 100 ml, IPTG is added to a final concentration of 1 mM and rhamnose is added to a final concentration of 1%. Both the induced and uninduced cultures are shaken at 19 °C for 24 hours.

Table 4.3: Strains used in Chapter 4 study

Strain	Relevant Characteristics	Source
ADP1	Wild type (BD413)	(38)
ACN2137	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon (**941,301 – 944,785**[b]) pBAC1601/AatII X ADP1	This study
ACN2192	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137*; *SBF52192* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon amplified by SBF pBAC1604. Selected for growth on 3 mM isovanillate pBAC1604/AatII X ADP1	This study
ACN2334	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137*; *SBF52192* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon amplified by SBF pBAC1604. Isolated from evolving population on 3 mM veratrate week 3 for whole genome resequencing	This study
ACN2439	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137*; *SBF52192* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon amplified by SBF pBAC1604. Isolated from evolving population on 3 mM isovanillate week 3 for whole genome resequencing	
ACN2446	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137*; *vanK*::*sacB*:S^R*52446* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon. vanK interrupted after **968,779**[b] with *sacB*:S^R cassette pBAC1762/AatII X ACN2137	This study
ACN2455	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBACKP52455*; ΩK^R*52137* *ver* operon without the regulator, *verR,* from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon pBAC1776/AatII/AhdI X ADP1	This study
ACN2483	Δ*rsfA52483* Deletion of *rsfA*	Alaa Ahmed Dissertation
ACN2540	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBACKP52137*; ΩK^R*52137*; *vanK52334*	This study

	ver operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon. *vanK* allele from evolved isolate ACN2334 replacing *sacB*:S^R cassette pBAC1707/AatII X ACN2446	
ACN2751	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137;* ΔACIAD_RS1391052483 *(rsfA)* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon with deletion of *rsfA* pBAC1601AatII X ACN2483	This study
ACN2752	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBA52752*; ΩK^R*52137;* ΔACIAD_RS1391052483 *(rsfA)* *ver* operon from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon with deletion of *rsfA* pBAC1877AatII X ACN2483	
ACN2753	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBA52752*; ΩK^R*52137* *verRBA* genes from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon pBAC1877AatII X ADP1	This study
ACN2754	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBA52754*; ΩK^R*52137;* Δ ACIAD_RS1391052483 *(rsfA)* *verBA* genes from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon with deletion of *rsfA* pBAC1879AatII X ACN2483	This study
ACN2755	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBA52754*; ΩK^R*52137* *verBA* genes from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon pBAC1879AatII X ADP1	This study
ACN2768	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBACKP52455*; ΩK^R*52137;* Δ ACIAD_RS1391052483 *(rsfA)* *ver* operon without the regulator, *verR,* from *Pseudomonas putida CSV86* codon optimized for ADP1 with ΩK^R cassette inserted in place of insertion elements 1 and 2 and composite transposon with deletion of *rsfA* pBAC1776/AatII/AhdI X ACN2483	This study
ACN2800	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBACKP52455*; ΩK^R*52137*; *SBF52137* *ver* operon with no regulator, *verR*, amplified by SBF pBAC1604 selected for growth on 2 mM isovanillate pBAC1604/AatII X ACN2455	This study
ACN2801	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBACKP52455*; ΩK^R*52137*; *SBF52137* *ver* operon with no regulator, *verR*, amplified by SBF pBAC1604 selected for growth on 2 mM isovanillate pBAC1604/AatII X ACN2455	This study
ACN2802	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391052483 *(rsfA)* *ver* operon amplified by SBF pBAC1604 selected for growth on 2 mM isovanillate with deletion of *rsfA* pBAC1604/AatII X ACN2751	This study
ACN2803	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBA52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391052483 *(rsfA)* *verBA* amplified by SBF pBAC1604 selected for growth on 2 mM isovanillate with deletion of *rsfA* pBAC1604/AatII X ACN2754	This study
ACN2804	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBA52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391052483 *(rsfA)* *verBA* amplified by SBF pBAC1604 selected for growth on 2 mM isovanillate with deletion of *rsfA* pBAC1604/AatII X ACN2754	This study

ACN2805	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBA52752*; ΩK^R*52137; SBF52137* *verRBA* amplified by SBF pBAC1604 selected for growth on 2 mM isovanillate pBAC1604/AatII X ACN2753	This study
ACN2806	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391*052483 (rsfA)* *ver* operon amplified by SBF pBAC1604 selected for growth on 2 mM veratrate with deletion of *rsfA* pBAC1604/AatII X ACN2751	This study
ACN2807	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBACKP52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391*052483 (rsfA)* *ver* operon amplified by SBF pBAC1604 selected for growth on 2 mM veratrate with deletion of *rsfA* pBAC1604/AatII X ACN2751	This study
ACN2808	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBA52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391*052483 (rsfA)* *verBA* amplified by SBF pBAC1604 selected for growth on 2 mM veratrate with deletion of *rsfA* pBAC1604/AatII X ACN2754	This study
ACN2809	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verBA52137*; ΩK^R*52137; SBF52137;* ΔACIAD_RS1391*052483 (rsfA)* *verBA* amplified by SBF pBAC1604 selected for growth on 2 mM veratrate with deletion of *rsfA* pBAC1604/AatII X ACN2754	This study
ACN2810	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBA52752*; ΩK^R*52137; SBF52137* *verRBA* amplified by SBF pBAC1604 selected for growth on 2 mM veratrate pBAC1604/AatII X ACN2753	This study
ACN2811	ΔACIAD_RS004390; ΔACIAD_RS04400; ΔACIAD_RS04405; ΔACIAD_RS04410::*verRBA52752*; ΩK^R*52137; SBF52137* *verRBA* amplified by SBF pBAC1604 selected for growth on 2 mM veratrate pBAC1604/AatII X ACN2753	This study

a. *A. baylyi* strains were derived from ADP1, previously known as *Acinetobacter calcoaceticus* or *Acinetobacter sp.* (39)

b. Bold numbers correspond to positions on the ADP1 chromosome in NCBI entry NC_005966.

c. Underlined text indicates the donor DNA and, where relevant, the restriction enzyme used to linearize a plasmid (pBAC number/Enzyme). The donor DNA was used to transform (X) the indicated recipient strain.

Table 4.4: Plasmids used in Chapter 4 study

Plasmid	Relevant Characteristics	Source
pUC18	ApR; cloning vector	(40)
pUC19	ApR; cloning vector	(40)
pUI1637	Ap^R, K^R; source of ΩK^R cassette	(41)
pBAC1548	Ap^R, S^R; source of *sacB*-S^R cassette	(42)

pBAC1593	Codon optimized *ver* genes amplified as multiple fragments SRB60/61, SRB62/63, SRB64/65, SRB66/67 assembled by NEBuilder and inserted into pUC18 digested with Eco53KI and HindIII	This study
pBAC1599	Plasmid to delete *IS1236* elements and insert *ver* genes. Upstream ACIAD_RS04390 amplified from *A. baylyi* genome with SRB85/86 and downstream ACIAD_RS04410 with SRB87/88. Fragments assembled by NEBuilder and inserted into pUC18 digested with Eco53KI and SphI	This study
pBAC1600	ΩK^R inserted into pBAC1599 by restriction digest and ligation with SacI	This study
pBAC1601	pBAC1593 and pBAC1600 digested with XhoI and NotI and ligated together to flank *ver* operon with *A. baylyi* DNA for chromosomal insertion	This study
pBAC1604	Synthetic bridging fragment. Fragments connecting downstream (SRB94/95) and upstream (SRB96/97) regions of amplicon assembled by NEBuilder in pUC18 digested with KpnI and XbaI	This study
pBAC1707	*vanK* allele from ACN2334 amplified from chromosome with SRB247/248 and assembled into pUC19 fragment made with SRB249/250 assembled in *E. coli*	This study
pBAC1762	pBAC1548 digested with EcoRV and MscI and inserted into pBAC1707 MscI site by ligation	This study
pBAC1776	Delete *verR* by PCR amplification pBAC1601 with SRB253/254 and self-ligated in *E. coli*	This study
pBAC1803	VerR expression plasmid, in pET-28b backbone. *verR* amplified with SRB393/394 and pET-28b amplified with SRB395/396 assembled in *E. coli*	This study
pBAC1877	Delete *verCKP* by PCR amplification pBAC1601 with SRB600/601 and self-ligated in *E. coli*	This study
pBAC1879	Delete *verR* by PCR amplification pBAC1877 with SRB253/254 and self-ligated in *E. coli*	This study

Table 4.5: Primers used in Chapter 4 study

Primer	Sequence	Use
MTV274	GCTCGACGCCTTCTATTTCAA	pPCR for *rpoA*
MTV275	TTTACGTCGCATTCTATTGTCTTCTT	pPCR for *rpoA*
MTV302	GCGTTGGCTACCCGTGATA	pPCR for ΩK^R
MTV303	GGAAGCGGTCAGCCCATT	pPCR for ΩK^R
SRB60	CTATGACCATGATTACGAATTCGAGGCGGCCGC TCATAACTGC	Amplify *verRBACKP* and assemble into pUC18. Sits in verR and pUC18
SRB61	TTAGGGGTTTATAAGTCTAGAACCAAACGAGGT GAACG	Amplify *verRBACKP* and assemble into pUC18. Sits in *verB* and downstream *verB*
SRB62	TTGGTTCTAGACTTATAAACCCCTAACAAGAAA TTCAGATAGC	Amplify *verRBACKP* and assemble into pUC18. Sits in *verB* and downstream *verB*
SRB63	GATCTTCAGTTACGTAAAGGCCCGGGATGCC	Amplify *verRBACKP* and assemble into pUC18. Sits in *verC*
SRB64	CCCGGGCCTTTACGTAACTGAAGATCCAGGTGC	Amplify *verRBACKP* and assemble into pUC18. Sits in *verC*
SRB65	AAACCGATTAGTGAGGTACCTACGATGCCAC	Amplify *verRBACKP* and assemble into pUC18. Sits in *verK*
SRB66	ATCGTAGGTACCTCACTAATCGGTTTATTAGCG GC	Amplify *verRBACKP* and assemble into pUC18. Sits in *verK*
SRB67	CGTTGTAAAACGACGGCCAGTGCCACTCGAGCG TATGGACGCTATG	Amplify *verRBACKP* and assemble into pUC18. Sits downstream *verP* and in pUC18
SRB68	GGTTACGGGTACGCATCATCAA	PCR confirmation and sequencing. Sits in *verR*
SRB69	GAATCAGGGAGGAGGTAACAGG	PCR confirmation and sequencing. Sits downstream *verR*
SRB70	CGTTCCTTGATTTGTTATCAGATGTTCC	PCR confirmation and sequencing. Sits in *verB*
SRB71	CGATTTCTTGCCAACGGTCCA	PCR confirmation and sequencing. Sits in *verA*
SRB72	TGGTATCCACATGAATGCTCAAATGC	PCR confirmation and sequencing. Sits in *verA*
SRB73	CAATCAATTCCGTATCCTAAAATGCATGATCC	PCR confirmation and sequencing. Sits in *verC*
SRB74	TCGTGACACCCCTTTGCAC	PCR confirmation and sequencing. Sits in *verC*
SRB75	CCACACAACTTGTTCCGTGC	PCR confirmation and sequencing. Sits in *verC*
SRB76	CCAAGAGCAGTAACACGTTGTTCAG	PCR confirmation and sequencing. Sits in *verK*
SRB77	GCATTTGGCTTAGTTCTTGTTATCGGT	PCR confirmation and sequencing. Sits in *verK*
SRB78	CGAAGCTCGCGATGAATATAGTGAATT	PCR confirmation and sequencing. Sits in *verP*
SRB79	AGCGTACCTGGGCTGTTAAATACG	PCR confirmation and sequencing. Sits in *verP*

SRB85	CGACGGCCAGTGAATTCGAGCCTCGAATGAATC ACAAC	Amplify upstream ACIAD_RS04390 and assemble into pUC18. Sits in ACIAD_RS04380 and pUC18
SRB86	GATCCAAGAAGAGCTCGATCACTTTAATGATTA AATAATCAATTTTC	Amplify upstream ACIAD_RS04390 and assemble into pUC18. Sits downstream ACIAD_ RS16975 and upstream ACIAD_RS04420
SRB87	TAAAGTGATCGAGCTCTTCTTGGATCTCCAAAG TTAACGTTACG	Amplify downstream ACIAD_RS04410 and assemble into pUC18. Sits downstream ACIAD_ RS16975 and upstream ACIAD_RS04420
SRB88	GATTACGCCAAGCTTGCATGCTTCAGCACGGGT GGCAAAC	Amplify downstream ACIAD_RS04410 and assemble into pUC18. Sits in ACIAD_RS04425 and in pUC18
SRB89	GTGCATTATTTCAACATCTTCACTTAAAGCATCG	PCR confirmation and sequencing. Sits in ACIAD_RS04385
SRB90	GGTCGACAAACATCTGATGCTAAAAGAG	PCR confirmation and sequencing. Sits in ACIAD_ RS04380
SRB91	GGAAGTGGTTTATGTTTAATAACAGTTACAGGT	PCR confirmation and sequencing. Sits in ACIAD_ RS16975
SRB92	CTCACCAAAAACAACAAATACATCACACC	PCR confirmation and sequencing. Sits upstream ACIAD_RS04420
SRB93	CAGCAATTACACCAGCAAAGACACTA	PCR confirmation and sequencing. Sits in ACIAD_RS04420
SRB94	AGTGAATTCGAGCTCGGTACGAGCTCTTCTTGG ATCTCC	Amplify downstream SBF fragment and assemble into pUC18. Sits downstream ACIAD_RS04420 and pUC18
SRB95	ATAATGACATATGTATGAAGAAACAATTATAAT CC	Amplify downstream SBF fragment and assemble into pUC18. Sits in ACIAD_RS04420 and ACIAD_RS04385
SRB96	CTTCATACATATGTCATTATTTTTAGAATGCTGC	Amplify downstream SBF fragment and assemble into pUC18. Sits in ACIAD_RS04420 and ACIAD_RS04385
SRB97	TGCATGCCTGCAGGTCGACTGAGCTCGATCACT TTAATGATTAAATAATC	Amplify downstream SBF fragment and assemble into pUC18. Sits downstream ACIAD_ RS16975 and pUC18
SRB247	GGCCAGTGAATTCGAGCTCGGTACCACTCCTTA TCACTTATTGGATCCAA	Amplify ACIAD_RS04520 (*vanK*) allele from ACN2334 and insert into pUC19. Sits upstream ACIAD_RS04520 (*vanK*) and in pUC19
SRB248	ATTACGCCAAGCTTGCATGCAAACCGAACATCA CTCAAAAAATTT	Amplify ACIAD_RS04520 (*vanK*) allele from ACN2334 and insert into pUC19. Sits downstream ACIAD_RS04520 (*vanK*) and in pUC19
SRB249	AAATTTTTTGAGTGATGTTCGGTTTGCATGCAA GCTTGGCG	Amplify ACIAD_RS04520 (*vanK*) allele from ACN2334 and insert into pUC19. Sits upstream ACIAD_RS04520 (*vanK*) and in pUC19
SRB250	TTGGATCCAATAAGTGATAAGGAGTGGTACCGA GCTCGAATTCAC	Amplify ACIAD_RS04520 (*vanK*) allele from ACN2334 and insert into pUC19. Sits upstream ACIADRS04520 (*vanK*) and in pUC19
SRB253	TAGGCAAGACGCGGCCGCCACCGCGGTGGAGC TCTTCTTGGATCT	Delete *verR* from pBAC1601
SRB254	TGGCGGCCGCGTCTTGCCTACCCCTGTTACCTCC TCCCTGATTCC	Delete *verR* from pBAC1601
SRB261	GATCCATTATATCAGTATGCAAATTGAGCAG	PCR confirmation and sequencing. Sits upstream ACIAD_RS04520 (*vanK*)
SRB262	GCCTGTTATTCCGACTTTGCC	PCR confirmation and sequencing. Sits downstream ACIAD_RS04520 (*vanK*)
SRB263	CGTAGTCTTGCTGTTAGTCTCATG	PCR confirmation and sequencing. Sits in ACIAD_RS04520 (*vanK*)
SRB264	GCACTCATGGTCGGTTTATTGC	PCR confirmation and sequencing. Sits in ACIAD_RS04520 (*vanK*)
SRB393	GGATCTCAGTGGTGGTGGTGGTGGTGTAACTGC GGTTCTACGGCATTCAC	Amplify *verR* and assemble in pET-28b. Sits in *verR* and pET-28b
SRB394	AATTTTGTTTAACTTTAAGAAGGAGATATACCA TGGCTCGCCATTTGTGG	Amplify *verR* and assemble in pET-28b. Sits in *verR* and pET-28b

SRB395	GAGCCATGGTATATCTCCTTCTTAAAGTTAAAC AAAATTATTTCTAGAGG	Amplify pET-28b and insert *verR*. Sits in pET-28b and *verR*.
SRB396	GTGAATGCCGTAGAACCGCAGTTACACCACCAC CACCACCACTGAGATCC	Amplify pET-28b and insert *verR*. Sits in pET-28b and *verR*.
SRB600	TATGGTGAAAGGAGGCAGTTATGGTCAC	Delete *verCKP* from pBAC1601
SRB601	TTTCACCATAGCGTCCATTCGAGGG	Delete *verCKP* from pBAC1601

Acknowledgments: Chris Johnson and Gregg Beckham for helpful discussion. Funding from U.S. Department (grant no. 2018-67009-27926).

References

1. Providenti MA, O'Brien JM, Ruff Jr, Cook AM, Lambert IB. 2006. Metabolism of isovanillate, vanillate, and veratrate by *Comamonas testosteroni* strain BR6020. Journal of Bacteriology 188:3862-3869.
2. Mohan K, Phale PS. 2017. Carbon source-dependent inducible metabolism of veratryl alcohol and ferulic acid in *Pseudomonas putida* CSV86. Applied and Environmental Microbiology 83:e03326-16.
3. Morawski B, Segura A, Ornston LN. 2000. Substrate range and genetic analysis of *Acinetobacter* vanillate demethylase. Journal of Bacteriology 182:1383-1389.
4. Morawski B, Segura A, Ornston LN. 2000. Repression of *Acinetobacter* vanillate demethylase synthesis by VanR, a member of the GntR family of transcriptional regulators. FEMS Microbiology Letters 187:65-68.
5. Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S. 2021. The InterPro protein families and domains database: 20 years on. Nucleic Acids Research 49:D344-D354.

6. Correll CC, Batie CJ, Ballou DP, Ludwig ML. 1992. Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]. Science 258:1604-1610.

7. Gassner GT, Ludwig ML, Gatti DL, Correll CC, Ballou DP. 1995. Structure and mechanism of the iron□sulfur flavoprotein phthalate dioxygenase reductase. The FASEB Journal 9:1411-1418.

8. Pavel EG, Martins LJ, Ellis Jr WR, Solomon EI. 1994. Magnetic circular dichroism studies of exogenous ligand and substrate binding to the non-heme ferrous active site in phthalate dioxygenase. Chemistry & Biology 1:173-183.

9. Tsang HT, Batie CJ, Ballou DP, Penner-Hahn JE. 1989. X-ray absorption spectroscopy of the [2-iron-2-sulfur] Rieske cluster in *Pseudomonas cepacia* phthalate dioxygenase. Determination of core dimensions and iron ligation. Biochemistry 28:7233-7240.

10. Pundir S, Martin MJ, O'Donovan C, Consortium U. 2016. UniProt tools. Current protocols in bioinformatics 53:1.29. 1-1.29. 15.

11. Cartwright N, Smith A. 1967. Bacterial attack on phenolic ethers: An enzyme system demethylating vanillic acid. Biochemical Journal 102:826-841.

12. Ribbons D. 1970. Stoicheiometry of O-demethylase activity in *Pseudomonas aeruginosa*. Febs Letters 8:101-104.

13. Venkatesagowda B. 2019. Enzymatic demethylation of lignin for potential biobased polymer applications. Fungal Biology Reviews 33:190-224.

14. Tumen-Velasquez M, Johnson CW, Ahmed A, Dominick G, Fulk EM, Khanna P, Lee SA, Schmidt AL, Linger JG, Eiteman MA. 2018. Accelerating pathway evolution by

increasing the gene dosage of chromosomal segments. Proceedings of the National Academy of Sciences 115:7105-7110.

15. Mallinson SJ, Machovina MM, Silveira RL, Garcia-Borràs M, Gallup N, Johnson CW, Allen MD, Skaf MS, Crowley MF, Neidle EL. 2018. A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. Nature Communications 9:1-12.
16. Grove A. 2017. Regulation of metabolic pathways by MarR family transcription factors. Computational and Structural Biotechnology Journal 15:366-371.
17. Renda BA, Dasgupta A, Leon D, Barrick JE. 2015. Genome instability mediates the loss of key traits by Acinetobacter baylyi ADP1 during laboratory evolution. J Bacteriol 197:872-81.
18. Gralton EM, Campbell AL, Neidle EL. 1997. Directed introduction of DNA cleavage sites to produce a high-resolution genetic and physical map of the Acinetobacter sp. strain ADP1 (BD413UE) chromosome. Microbiology 143:1345-1357.
19. Reams AB, Neidle EL. 2003. Genome plasticity in Acinetobacter: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments. Molecular microbiology 47:1291-1304.
20. Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq, p 165-188, Engineering and analyzing multicellular systems. Springer.
21. Deochand DK, Grove A. 2017. MarR family transcription factors: dynamic variations on a common scaffold. Critical Reviews Biochemistry Molecular Biology 52:595-613.
22. Grove A. 2013. MarR family transcription factors. Current Biology 23:R142-R143.

23. Schasfoort RB. 2017. Handbook of surface plasmon resonance. Royal Society of Chemistry.

24. Maynard JA, Lindquist NC, Sutherland JN, Lesuffleur A, Warrington AE, Rodriguez M, Oh SH. 2009. Surface plasmon resonance for high□throughput ligand screening of membrane□bound proteins. Biotechnology Journal: Healthcare Nutrition Technology 4:1542-1558.

25. Pattnaik P. 2005. Surface plasmon resonance. Applied Biochemistry and Biotechnology 126:79-92.

26. Gestwicki JE, Hsieh HV, Pitner JB. 2001. Using receptor conformational change to detect low molecular weight analytes by surface plasmon resonance. Analytical Chemistry 73:5732-5737.

27. Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification 41:207-234.

28. Ruangprasert A, Craven SH, Neidle EL, Momany C. 2010. Full-length structures of BenM and two variants reveal different oligomerization schemes for LysR-type transcriptional regulators. Journal of Molecular Biology 404:568-586.

29. Tumen-Velasquez MP, Laniohan NS, Momany C, Neidle EL. 2019. Engineering CatM, a LysR-type transcriptional regulator, to respond synergistically to two effectors. Genes 10:421.

30. Pardo I, Jha RK, Bermel RE, Bratti F, Gaddis M, McIntyre E, Michener W, Neidle EL, Dale T, Beckham GT. 2020. Gene amplification, laboratory evolution, and biosensor screening reveal MucK as a terephthalic acid transporter in *Acinetobacter baylyi* ADP1. Metabolic Engineering 62:260-274.

31. Barbe V, Vallenet D, Fonknechten N, Kreimeyer A, Oztas S, Labarre L, Cruveiller S, Robert C, Duprat S, Wincker P. 2004. Unique features revealed by the genome sequence of *Acinetobacter sp*. ADP1, a versatile and naturally transformation competent bacterium. Nucleic Acids Research 32:5766-5779.

32. Biggs BW, Bedore SR, Arvay E, Huang S, Subramanian H, McIntyre EA, Duscent-Maitland CV, Neidle EL, Tyo KEJ. 2020. Development of a genetic toolset for the highly engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. Nucleic Acids Research doi:10.1093/nar/gkaa167.

33. Singh A, Bedore SR, Sharma NK, Lee SA, Eiteman MA, Neidle EL. 2019. Removal of aromatic inhibitors produced from lignocellulosic hydrolysates by *Acinetobacter baylyi* ADP1 with formation of ethanol by *Kluyveromyces marxianus*. Biotechnology For Biofuels 12:91.

34. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press.

35. Neidle E, Hartnett C, Ornston L. 1989. Characterization of *Acinetobacter calcoaceticus* catM, a repressor gene homologous in sequence to transcriptional activator genes. Journal of Bacteriology 171:5410-5421.

36. Seaton SC, Elliott KT, Cuff LE, Laniohan NS, Patel PR, Neidle EL. 2012. Genome‐wide selection for increased copy number in Acinetobacter baylyi ADP1: locus and context‐dependent variation in gene amplification. Molecular microbiology 83:520-535.

37. Stoudenmire JL, Schmidt AL, Tumen-Velasquez MP, Elliott KT, Laniohan NS, Whitley SW, Galloway NR, Nune M, West M, Momany C. 2017. Malonate degradation in

Acinetobacter baylyi ADP1: Operon organization and regulation by MdcR. Microbiology 163:789.

38. Juni E, Janik A. 1969. Transformation of *Acinetobacter calco-aceticus* (Bacterium anitratum). Journal of Bacteriology 98:281-288.

39. Vaneechoutte M, Dijkshoorn L, Tjernberg I, Elaichouni A, de Vos P, Claeys G, Verschraegen G. 1995. Identification of *Acinetobacter* genomic species by amplified ribosomal DNA restriction analysis. Journal of Clinical Microbiology 33:11-15.

40. Norrander J, Kempe T, Messing J. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene 26:101-106.

41. Eraso JM, Kaplan S. 1994. prrA, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*. Journal of Bacteriology 176:32-43.

42. Jones RM, Williams PA. 2003. Mutational analysis of the critical bases involved in activation of the AreR-regulated σ54-dependent promoter in *Acinetobacter sp.* strain ADP1. Applied and Environmental Microbiology 69:5627-5635.

CHAPTER 5:

CONCLUSIONS AND FUTURE DIRECTIONS

Approaches to, and significance of, the expansion of *A. baylyi* metabolism

The studies described in this thesis are unified by the common goal of understanding metabolism and regulation by exploiting the genetic malleability of *Acinetobacter baylyi* ADP1. In many cases, it proved useful to develop new techniques or modifications of methods that are uniquely possible because of the exceptionally high efficiency of natural transformation and homologous recombination in strain ADP1. In **Chapter 2**, I describe the incorporation of foreign DNA into the *A. baylyi* chromosome for the consumption of aromatic compounds through a non-native pathway (a protocatechuate *meta*-cleavage route called the PCA-4,5 pathway). The long-term goal of this work is to increase the number of lignin-derived aromatic compounds that can be consumed by an individual bacterium. To accomplish this goal, I first established a foundation for modular expansion (**Figure 2.1**) and demonstrated that this approach to chromosomal manipulation is feasible. My research approach contributes to improved understanding of how new metabolic functions become regulated, optimized, and integrated in a new host. The results are relevant to fundamental aspects of evolution, horizontal gene transfer, and metabolic regulation.

The successful isolation and characterization of strains that evolved to consume aromatic growth substrates using the PCA-4,5 pathway at 30 °C unveiled a role for an uncharacterized phosphotransferase system (PTS) and peptidyl-prolyl isomerases (PPIase) in *A. baylyi* regulation.

The importance of a PtsP-dependent regulatory circuit was revealed by the frequent occurrence of its genetic inactivation and by the effects of engineered *ptsP* deletions (**Table 2.5, Figures 2.8 and 2.9**). These findings lay the foundation for future studies to determine the components and functions of this novel PTS. As indicated in the diagram below, PtsP of ADP1 is a homolog of the first component of a three-component PTS^{Ntr}, a type of regulatory system that has been studied in other bacteria. PTS^{Ntr} has been found to balance cellular carbon and nitrogen availability by integrating small molecule-signaling pathways and by altering transcriptional, posttranscriptional, and posttranslational controls (1). There is no PtsN homolog in ADP1 indicating that key aspects of PtsP-mediated regulation in this bacterium need to be determined and compared to other PtsP-based circuits (2). Some type of PtsP-dependent regulation appears to be critical for the metabolic integration and function of a foreign pathway for aromatic compound degradation in *A. baylyi.* It is intriguing to speculate that, as in *Legionella pneumophila*, a relay system composed of PtsP and PtsO ultimately affects phosphorylation of a two-component transcriptional regulatory system that participates in global regulation (2). Preliminary data suggest that one such transcriptional system is involved in regulating aromatic compound degradation in ADP1. My results provide a framework for generating and testing specific hypotheses about the role of PtsP.

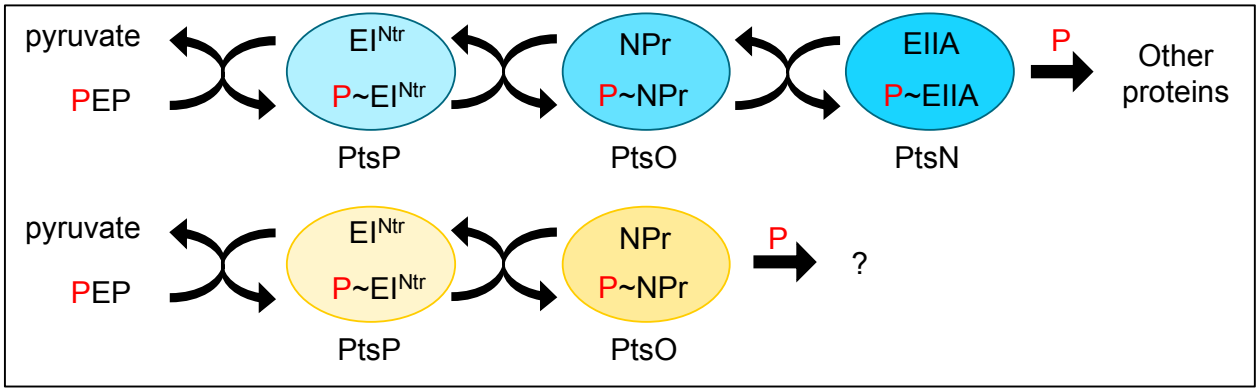


Figure 5.1: Canonical PTS^{Ntr} compared to *A. baylyi* incomplete PTS

Blue ovals represent the canonical PTS^{Ntr} in which a phosphate is relayed from phosphoenolpyruvate (PEP), through each enzyme (PtsP, PtsO, and PtsN) to downstream proteins. The yellow ovals represent a possible functional PTS in *A. baylyi* that lacks a PtsN component. As discussed earlier, a PtsO homolog exists in ADP1, but whether it functions with PtsP has not been studied.

Mutants generated during EASy studies with the PCA *meta*-cleavage pathway revealed an unexpected effect on maximal growth temperature. A single copy of the genes enabled growth using the pathway, but not at temperatures above 25 °C. The genes encoding this pathway came from a marine bacterium that has a lower optimal growth temperature than does ADP1. However, it is unclear whether there is something inherently associated with protein sequence that restricts the maximum temperature of functionality for this pathway. Nevertheless, it is easy to envision that during horizontal gene transfer, the instability of foreign proteins would prove problematic. Thus, the mutants that emerged to improve protein stability during laboratory evolution are likely to reflect common adaptation mechanisms that occur in nature. Growth at increased temperature with the foreign pathway resulted in mutations that appear to increase expression of a well-characterized chaperone, GroE (3). This result is consistent with the well-known effects of GroE on growth at different temperatures (4). The effects of mutations in the *tig, slyD* and *fbp* genes implicate the encoded FKBP PPIase enzymes in critical issues of protein maturation and/or stability affecting growth via the PCA 4,5-pathway. As described in **Chapter 1**, the known functions of these enzymes are consistent with their ability to affect growth at

higher temperatures. However, the specific role of these proteins and their interactions with PtsP cannot yet be deduced. These proteins should be the focus of future studies.

Gene amplification and remarkable chromosomal flexibility

Gene amplification, induced by the EASy method, facilitated the acquisition of new capabilities by *A. baylyi*, such as enabling the use of veratrate and isovanillate as growth substrates and the consumption of compounds via a foreign *meta*-ring cleavage pathway. These results support models that propose a transient role for gene amplification during adaptation and evolution, as previously reviewed (5). The extent to which strains transiently increased the apparent copy number of the PCA 4,5-pathway was impressive. On the *A. baylyi* chromosome, some strains had more than 100 copies (one had 172 copies) of the F87 fragment. Typically, the initial selection for growth in the presence of high levels of kanamycin requires no more than 10 copies of the drug-resistance gene. Growth at 37 °C using the foreign pathway, therefore, initially required and selected massive alteration of the *A. baylyi* chromosome. A region of 172 12-kbp amplicons corresponds to a segment of approximately 2 Mbp. Since the *A. baylyi* ADP1 chromosome is 3.6 Mbp, this additional DNA has a size equal to 50% of the normal chromosome, as shown schematically in **Figure 5.2.** Moreover, since the amplicons are tandemly repeated, all must be on one side of the replicore. This finding raises important questions about chromosomal symmetry and the mechanism(s) by which the genome can be properly replicated.

Genomic flexibility was also evident in the ease with which large deletions could be made. The novel approach to remove all remaining amplicon copies from the genomes of selected mutants enabled analysis of the role played by mutations within and external to the target region. This useful analysis was possible because large deletions could readily be made by

allelic replacement using a linear donor DNA fragment. For example, allelic replacement readily yielded the loss of 130 kbp of DNA, as illustrated in **Figure 5.2.**

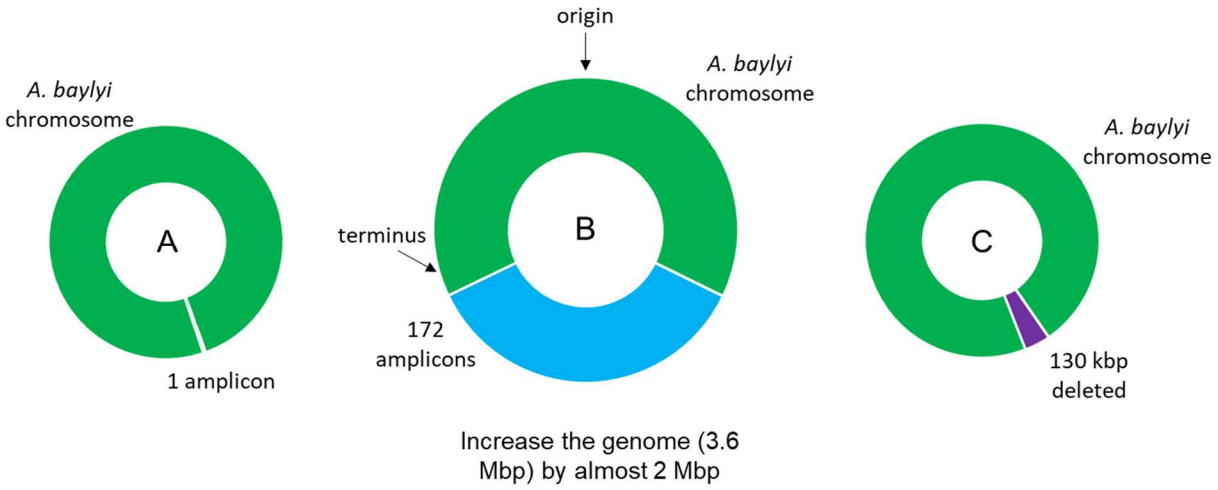


Figure 5.2: Chromosomal changes from gene amplification and chromosomal deletions

Circle A represents an *A. baylyi* chromosome with a single copy of the foreign PCA-4,5 pathway genes. Circle B represents an *A. baylyi* chromosome with 172 copies causing a large increase in the size of the chromosome and significant asymmetry for replication. Circle C represents an *A. baylyi* chromosome in which 130 kbp was deleted in a single step during the amplicon replacement method from Chapter 2.

Development of new genetic and synthetic biology methods for *A. baylyi*

Based on this chromosomal flexibility, I contributed to three publications developing new and modified tools to facilitate the use of *A. baylyi* as a platform organism for synthetic biology and biotechnology applications (6-8). In one of these publications (6), I used a newly developed synthetic promoter library to test multiple hypotheses about aspects of aromatic compound catabolism in *A. baylyi* ADP1 that we did not have tools to test previously. Additionally, in

Chapter 3, I used this promoter library to test the specificity of two homologous transcriptional regulators for the two isomers of aspartate, L-Asp and D-Asp.

Additional method development is underway and can be completed in the future. One such method seeks to discover metabolic pathways and genes by capturing foreign DNA on the *A. baylyi* chromosome without prior characterization. As illustrated in **Chapter 2** and **Chapter 4**, metabolic engineering usually is based on known catabolic pathways, an important scientific challenge is to discover metabolic capabilities and novel pathways that have not yet been elucidated. The full range of microbial capability is far from being understood. During my doctoral studies, I collaborated with Dr. Yo Suzuki at the J. Craig Venter Institute to develop a method of capturing environmental DNA (eDNA), collected without prior culturing, on the *A. baylyi* chromosome (**Figure 5.3**).

Traditional cloning methods capture foreign DNA fragments on plasmids, or similar vectors for expression in model bacteria, such as *E. coli*. There are several problems with this method. There is a limit to the size of DNA fragments that can readily be cloned by these traditional plasmid-based methods. In contrast, our studies with *A. baylyi* suggest that extremely large segments of foreign DNA can be incorporated in the chromosome and replicated. Furthermore, in *A. baylyi* the adjustable regulation of protein expression could be mediated by the cell via gene amplification after the induction of a duplication. There is another possible problem with using a traditional host, such as *E. coli*, to capture and express genes for the catabolism of lignin-derived compounds. Namely, since *E. coli* does not naturally degrade aromatic compounds, it may not have the catabolic versatility needed for the desired proteins and

pathways to function. Thus, developing the ADP1 chromosome to serve essentially as a cloning vector is an attractive idea.

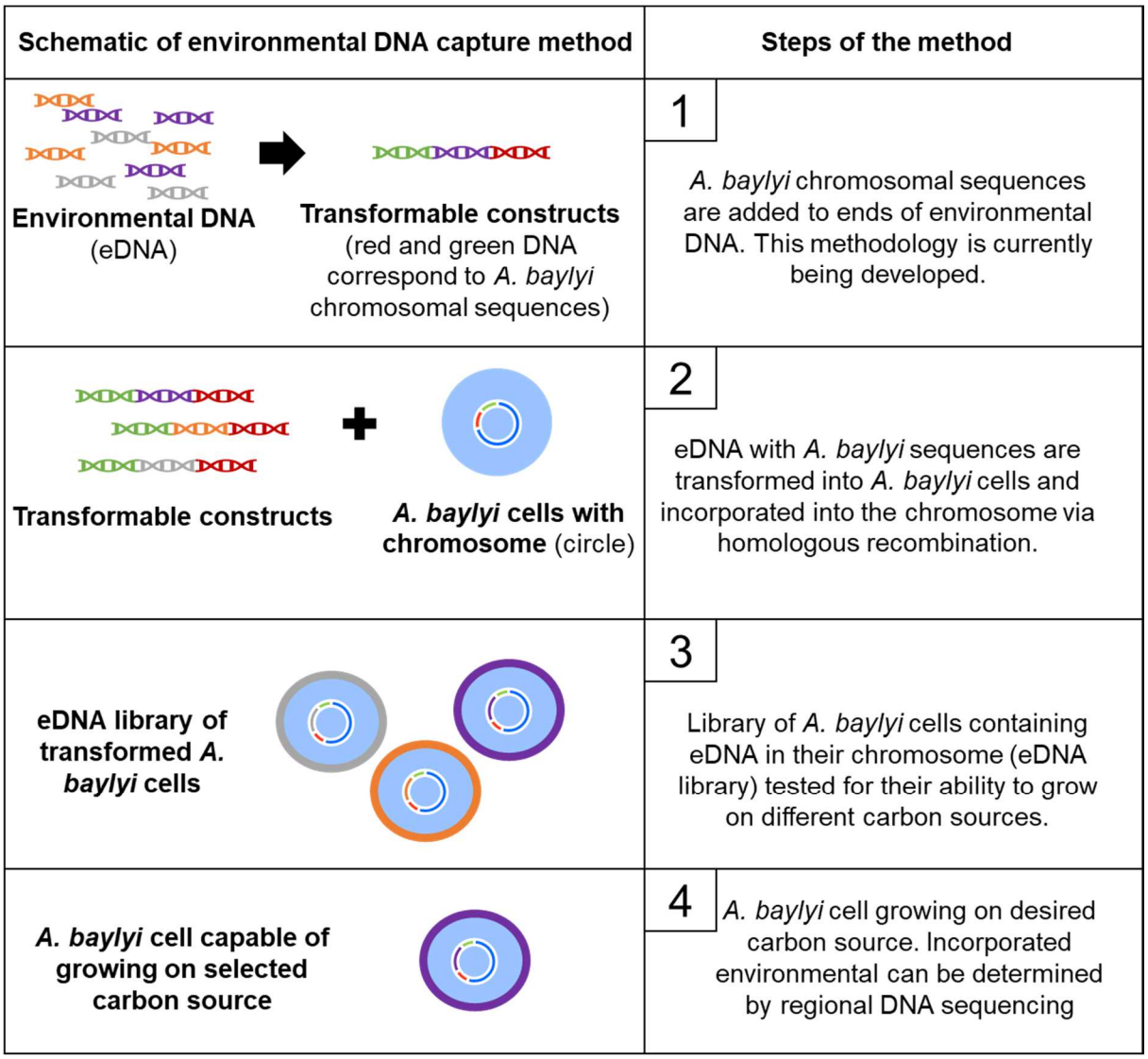


Figure 5.3: Method of capturing environmental DNA on the *A. baylyi* chromosome

Left panels depict schematics of each step described in the corresponding right panel. Environmental DNA represented by orange, purple, and grey double helices, *A. baylyi* sequences are represented by green and red double helices, *A. baylyi* cells are blue circles with a circular chromosome inside, and *A. baylyi* cells that have integrated eDNA have the corresponding color

circle around them. Current work focuses on step 1, adding *A. baylyi* sequences to environmental DNA.

The PCA degradation pathway from *P. atlantica* T6c genomic DNA can be used as proof-of-concept for this method development since I have already shown that integrating the pathway allows *A. baylyi* to grow on PCA or POB at 25 °C. In initial studies, the *P. atlantica* genomic DNA is being used as the eDNA (**Figure 5.3 Step 1**), and selection for integration of the *lig* genes in this sample DNA is the ability to grow on PCA at 25 °C (**Figure 5.3 Step 3**).

This eDNA capture method first requires the tagging of eDNA with *A. baylyi* sequences. The *A. baylyi* DNA sequence allows homologous recombination to mediate allelic replacement that inserts tagged DNA into a known location on the chromosome. Initial attempts to add *A. baylyi* sequences used two methodologies. One method used a Tn5 transposon loaded on DNA containing the *A. baylyi* sequences. The transposon integrates the *A. baylyi* sequences within the eDNA. The second method used tagged primers to amplify eDNA with the phi29 polymerase. The phi29 polymerase has high fidelity and can generate extremely large DNA fragments through multiple displacement amplification (9). The primer tags can then be used for *in vitro* assembly with the *A. baylyi* sequences. Future experiments should continue to refine this methodology.

A different method that I have been developing seeks to facilitate the identification of which mutations are significant for conferring a new phenotype. In laboratory evolution experiments, such as those reported in **Chapters 2 and 4**, whole genome re-sequencing of strains revealed multiple chromosomal mutations, often six or more. Reconstructing all mutations found in a strain would be a massive undertaking that is not feasible. However, if multiple mutations

could be simultaneously transformed into *A. baylyi* cells in a transformation assay like that used to demonstrate the importance of the *clpA* alleles in **Chapter 3**, then we could quickly test what combinations of mutations confer a growth benefit.

The transformation assay I am developing relies on two steps, natural transformation to bring DNA into the cells and allelic replacement to integrate the DNA into the chromosome. To test multiple fragments simultaneously would require that all fragment of interest would efficiently transform an individual cell. For example, to test seven different mutations at once might require a single recipient to take up seven fragments. Future experiments could ensure that all DNA of interest enters the cell in a single uptake step by creating one piece of DNA containing all mutant alleles. An *A. baylyi* strain could be made that expresses a restriction enzyme that does not cut any sequences in its chromosome, such as FseIR (**Figure 5.4**). If the mutant alleles are separated by the FseI recognition sequence, they could be cut from the single fragment once inside the cell and then used for allelic replacement via homologous recombination (**Figure 5.4**). This method could create a high throughput screen for analyzing mutations that arise in *A. baylyi* during laboratory evolution experiments.

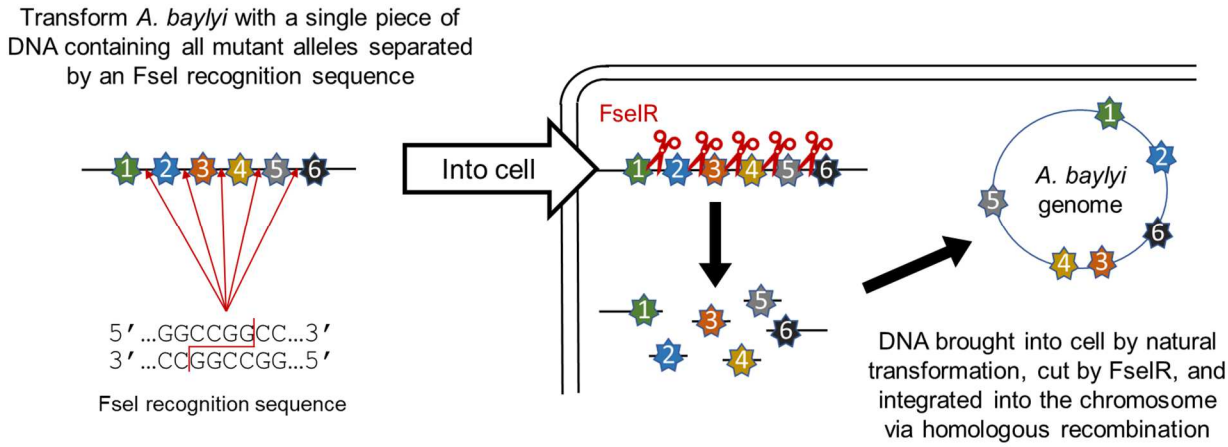


Figure 5.4: Method to analyze multiple mutated alleles by a single natural transformation step

Numbered stars represent mutant chromosomal alleles. A single piece of DNA containing all alleles (line with numbered stars) can be taken into the chromosome by natural transformation and then cut with the FesIR restriction enzyme (represented by red scissors).

Using *A. baylyi* for biotechnology

Collectively, the research described in this thesis addresses fundamental and applied aspects of metabolism and regulation in a genetically malleable soil bacterium. Aromatic compound catabolism is important for the bioeconomy of the future. As depicted in **Figure 5.5**, valorization of lignin by bacteria could convert an energy rich renewable resource such as lignin into industrially important compounds. Other applications of microbial synthesis or degradation of aromatic compounds include alleviating problems caused by the accumulation of plastic wastes, synthesizing new polymers, and/or the bioremediation of aromatic pollutants (10-12). Towards these long-term goals, my studies have increased understanding of metabolism, generated new methods, and helped to develop a microbiology course based on the unique attributes of *A. baylyi* ADP1.

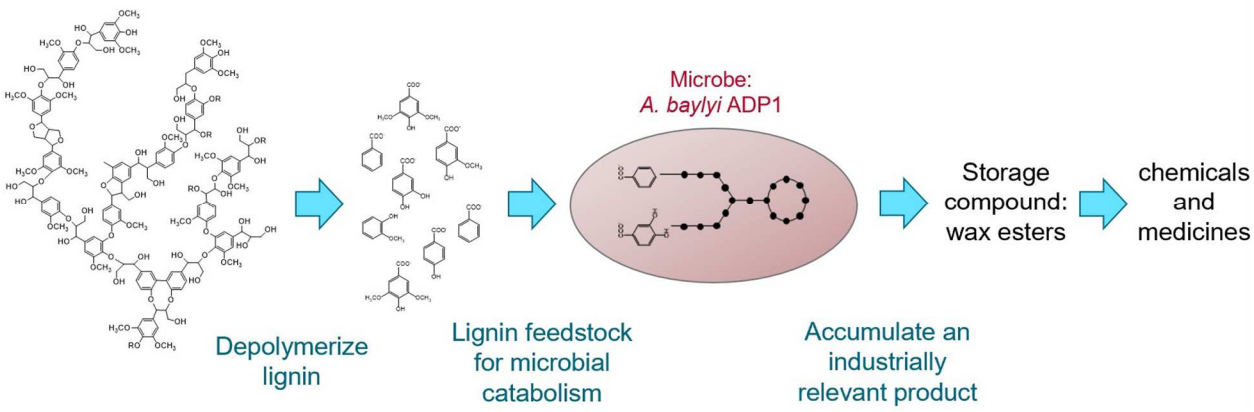


Figure 5.5: Microbial Lignin Valorization

Steps in microbial lignin valorization. Lignin is represented by the interconnected chemical structures which are then depolymerized into the various aromatic compounds. *A. baylyi* ADP1 (red oval) can then catabolize those aromatics and accumulate a product of interest.

References

1. Sánchez-Cañizares C, Prell J, Pini F, Rutten P, Kraxner K, Wynands B, Karunakaran R, Poole PS. 2020. Global control of bacterial nitrogen and carbon metabolism by a PTSNtr-regulated switch. Proceedings of the National Academy of Sciences 117:10234-10245.
2. Speiser Y, Zusman T, Pasechnek A, Segal G. 2017. The *Legionella pneumophila* incomplete phosphotransferase system is required for optimal intracellular growth and maximal expression of PmrA-regulated effectors. Infection and Immunity 85:e00121-17.
3. Ishii N. 2017. GroEL and the GroEL-GroES complex. Macromolecular Protein Complexes:483-504.

4. Fayet O, Ziegelhoffer T, Georgopoulos C. 1989. The groES and groEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. Journal of bacteriology 171:1379-1385.

5. Elliott KT, Cuff LE, Neidle EL. 2013. Copy number change: evolving views on gene amplification. Future microbiology 8:887-899.

6. Biggs BW, Bedore SR, Arvay E, Huang S, Subramanian H, McIntyre EA, Duscent-Maitland CV, Neidle EL, Tyo KEJ. 2020. Development of a genetic toolset for the highly engineerable and metabolically versatile *Acinetobacter baylyi* ADP1. Nucleic Acids Research 48:5169-5182.

7. Singh A, Bedore SR, Sharma NK, Lee SA, Eiteman MA, Neidle EL. 2019. Removal of aromatic inhibitors produced from lignocellulosic hydrolysates by *Acinetobacter baylyi* ADP1 with formation of ethanol by *Kluyveromyces marxianus*. Biotechnology For Biofuels 12:91.

8. Jin Luo EM, Stacy Bedore, Ville Santala, Ellen Neidle, and Suvi Santala. 2021. Characterization of highly ferulate-tolerant *Acinetobacter baylyi* ADP1 isolates by a rapid reverse-engineering method. Applied and Environmental Microbiology. In press.

9. Garmendia C, Bernad A, Esteban JA, Blanco L, Salas M. 1992. The bacteriophage phi 29 DNA polymerase, a proofreading enzyme. Journal of Biological Chemistry 267:2594-2599.

10. Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT. 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proceedings of the National Academy of Sciences 111:12013-12018.

11. Bugg TD, Rahmanpour R. 2015. Enzymatic conversion of lignin into renewable chemicals. Current Opinion in Chemical Biology 29:10-17.

12. Danso D, Chow J, Streit WR. 2019. Plastics: environmental and biotechnological perspectives on microbial degradation. Applied and Environmental Microbiology 85:e01095-19.