UNIVERSITY OF CALIFORNIA
Los Angeles

Evolution of metabolism to reveal metabolic repair and to improve engineered microbial production

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy
in Chemical and Biomolecular Engineering

by

Samuel Patrick Pontrelli

2018

ABSTRACT OF THE DISSERTATION

Evolution of metabolism to reveal metabolic repair
and to improve engineered microbial production

Samuel Patrick Pontrelli
Chemical and Biomolecular Engineering
University of California, Los Angeles, 2018
James C. Liao, Chair

Harnessing the adaptive nature of cell metabolism presents an opportunity to understand the function of biological systems, how they adapt, and how they may respond when challenged. It also stands as a tool that can aid in restoring impaired metabolic function caused by engineering microbial production. Many studies have demonstrated the ability of the cell to overcome metabolite auxotrophies and have elucidated underlying mechanisms. However, these studies have primarily focused on mechanisms that directly replace mutant function. In this work, we first aim to expand this view by evolving and elucidating more complex adaptive mechanisms. As an example, we used a *ΔpanD* strain of *E. coli*, a β-alanine auxotroph, to demonstrate that entire metabolic pathways can evolve to repair auxotrophy. Using directed strain evolution, we showed that *E. coli* successively evolved three distinct metabolic pathways to synthesize β-alanine. The first involved significant rewiring and repurposing of the uracil synthesis and degradation pathways. The second relied on a gain-of-function mutation in ornithine decarboxylase (SpeC) which altered substrate and reaction specificity. The third pathway emerged that relies on synthesis of polyamines.

This work also serves as a demonstration of how metabolism can be evolved to overcome impaired metabolic function that may be incurred through engineering microbes for production.

As an example, we focused on a modified strain of *E. coli* that is capable of producing high titers of butanol in rich media using an anaerobic, growth-coupled, modified *Clostridial* CoA-dependent pathway. For unknown reasons, strain modifications impaired metabolic function. Using directed strain evolution, a strain was acquired that has improved growth, titers and butanol yields. We further identified several mutations that adapted energy and carbon metabolism and optimized expression of pathway enzymes. These works collectively demonstrate the elucidation of adaptive mechanisms of cell metabolism and further, they demonstrate applications in strain engineering.

The dissertation of Samuel Patrick Pontrelli is approved.

Yvonne Chen

Matteo Pellegrini

Yi Tang

James C. Liao, Committee Chair

University of California, Los Angeles

2018

Contents

List of Figures and Tables

Biographical Sketch

EDUCATION

UCLA Chemical and Biomolecular Engineering
PhD Candidate, James Liao Lab 2013 – 2018

Santa Clara University
MS Bioengineering 2012 – 2013
BS Bioengineering 2008 – 2012

PUBLICATIONS

- **Pontrelli S.,** Fricke R.B.C, Sakurai S., Laviña W. A., Putri S.P., Fitz-Gibbon S., Chung M., Chiu T.Y., Pellegrini M., Fukusaki E., Liao J.C. Microbial evolution rebalanced metabolic infrastructure for butanol production in minimal media, *Metabolic Engineering* (Under review)
- Nitta, K., Laviña, W. A., **Pontrelli S.**, Liao, J. C., Putri, S. P., & Fukusaki, E. The glyoxylate shunt as a target for improved 1-butanol production in *Escherichia coli. Journal of Bioscience and Bioengineering* (Under Review)
- Putri, S.P., Nakayama, Y., Shen, C., Noguchi, S., Bamba, T., **Pontrelli, S.,** Liao, J.C., Fukusaki, E., Identifying metabolic elements that contribute to productivity of 1-propanol bioproduction in *Escherichia coli. Metabolomics* (accepted)
- **Pontrelli S.**, Fricke R.B.C, Teoh S. T., Laviña W. A., Putri S.P., Fitz-Gibbon S., Chung M., Pellegrini M., Fukusaki E., Liao J.C. Metabolic Repair through Emergence of New Pathways. *Nature*, (under review)
- **Pontrelli S.,** Chiu T.Y., Lan E.I, Chen F.Y, Chang P.C, Liao J.C., *Escherichia coli* as a host for industrial production *Metabolic Engineering* (review article, under review)
- Nitta, K., Laviña, W. A., **Pontrelli, S.**, Liao, J. C., Putri, S. P., & Fukusaki, E. (2017). Orthogonal partial least squares/projections to latent structures regression-based metabolomics approach for identification of gene targets for improvement of 1-butanol production in Escherichia coli. *Journal of bioscience and bioengineering*, *124*(5), 498-505.
- Ohtake, T., ***Pontrelli, S.**, Laviña, W. A., Liao, J. C., Putri, S. P., & Fukusaki, E. (2017). Metabolomics-driven approach to solving a CoA imbalance for improved 1-butanol production in Escherichia coli. *Metabolic engineering*, *41*, 135-143.
  ***Equal contributing author**
- Wernick, D. G., **Pontrelli, S. P.**, Pollock, A. W., & Liao, J. C. (2016). Sustainable biorefining in wastewater by engineered extreme alkaliphile Bacillus marmarensis. *Scientific reports*, *6*, 20224.
- Liao, J. C., Mi, L., **Pontrelli, S.**, & Luo, S. (2016). Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, *14*(5), 288.

CONFERENCE PRESENTATIONS

- **Pontrelli S.,** Fricke R.B.C, Teoh S. T., Laviña W. A., Putri S.P., Fitz-Gibbon S., Chung M., Pellegrini M., Fukusaki E., Liao J.C. *Emerging metabolic pathways overcome metabolic blocks.* Presentation at the meeting of American Chemical Society, 2017, Washington DC

AWARDS

- Pollination Project Grantee, 2016 for developing Science Education programs in Uganda
- 1st place presentation competition, Shanghai Jiao Tong University Summer Academy of Pharmacology, 2013
- James W. Reites Engineering Award, Santa Clara University, 2012
- Outstanding Bioengineering Senior Award, Santa Clara University, 2012
- 1st place Business Plan Competition, 2012 Santa Clara University

- Received 2012 Roelandt's Fellowship for Master's Thesis research in paper based point-of-care diagnostics
- Received 2011 Jean Donovan Fellowship for completion of community development in Uganda
- 1st place Freshman Math Competition, 2008 Santa Clara University
- Received 2008 L.A. Arch Diocese Youth Community Service Award

---

Chapter 1 : Introduction and background

Portions of this section were adapted partially from Pontrelli S., Chiu T.Y., Lan E.I, Chen F.Y, Chang P.C, Liao J.C., Escherichia coli as a host for industrial production *Metabolic Engineering* (under review)

Overview

Escherichia coli can synthesize all required metabolites from a variety of carbon sources. Specific chemical reactions that comprise biosynthetic pathways may be required for formation of essential metabolites, and therefore, deleterious growth consequences may arise when these reactions are blocked. Metabolism has proven the ability to repair itself from these deleterious consequences(*1*–*5*). By studying this ability, we can gain insight into the function of biological systems, how new functions emerge, and how life may respond when challenged. Further, the ability of metabolism to repair stands as an opportunity to enhance bioproduction when engineered pathways impair metabolic function. This work aims to demonstrate the mechanisms by which *E. coli* can actively utilize its existing genetic components to overcome metabolic damage, and further, how damaged metabolism can be repaired to enhance production of 1-butanol. Two main works, described within chapters 2 and 3, describe these aims. In both works, directed strain evolution plays a critical role, which is discussed in detail below.

In chapter 2, a *ΔpanD* strain of *E. coli,* unable to produce the β-alanine required for synthesis of Coenzyme A, is used as an example to demonstrate mechanisms by which new pathways can form to overcome metabolite auxotrophy. In this, three distinct metabolic pathways are successively evolved to synthesize β-alanine. Here, metabolic pathways emerged using three contributing mechanisms i) rewiring of existing metabolic networks, ii) repurposing promiscuous enzymes and iii) evolution of new enzyme function. Specific contributing genetic mutations and metabolic perturbations are identified to further elucidate how these pathways emerged.

Within chapter 3, a strain of *E. coli* designed to produce 1-butanol in minimal media is used to demonstrate that metabolism can be repaired to enhance bioproduction. Within this strain, all

fermentation pathways are deleted, thus placing the 1-butanol pathway as the sole electron sink. While this strategy worked for production of high titers within rich media, neither growth nor production was observed within minimal media. Here, the burden of the expressed pathway, combined with the altered physiological metabolic flux, impaired metabolic function. With use of directed strain evolution, metabolic function was restored, and resulted in increased growth, titers, and butanol yields. Causal genetic mutations and metabolic perturbations are identified that allow for this improved phenotype.

The following sections provide a background pertaining to metabolic repair. First, strategies for directed strain evolution are discussed and examples are given in which directed strain evolution has previously been used in metabolic engineering. Second, promiscuous enzymes and underground metabolism are discussed, as well as deleterious consequences that they may confer. These physiological components are critical to understanding of this work as they may serve as essential components used to restore metabolic function.

Directed strain evolution

Directed strain evolution involves the continuous culturing of a host with a constantly applied selection pressure, allowing the accumulation of mutations that give rise to a desired phenotype. There are many factors that must be considered when designing a methodology for a directed strain evolution experiment. These include growth medium, timescale, mutagenesis rate, culture and propagation size, and maintenance of growth phases. These considerations are discussed elsewhere(*6*). The rate of evolution can be accelerated by different strategies that enhance mutagenesis. These strategies include chemical mutagenesis(*7*), UV mutagenesis(*8*), compromising DNA repair mechanisms(*9*), and implementation of synthetic variable mutagenesis mechanisms(*10*).

Directed strain evolution stands as a diverse strategy and has previously been employed for generating a variety of industrially relevant phenotypes. *E. coli* and other microbes have been evolved to exhibit enhanced tolerances to conditions such as heat(*11*), pH(*12*), salt concentrations(*13*), and substrate or product toxicity(*8*, *14*)(*15*)(*16*). In one case it was shown that improved tolerance to isobutanol also resulted in tolerance to *n*-butanol and 2-methyl-1-mutanol(*16*). However, the same strain showed no improvement in ethanol tolerance and higher sensitivity to hexane and chloramphenicol. This demonstrates the existence of evolutionary tradeoffs that may come with acquisition of certain phenotypes, and in this case, it must be noted that tolerance mechanisms to the tested compounds relies on different reaction mechanisms. Directed strain evolution has also been used to enhance growth rates(*17*), and to evolve *E. coli* to utilize alternating carbon sources, reducing lag caused diauxic growth phases(*17*).

Substrate consumption or product formation have also been enhanced using directed strain evolution(*17*)(*18*)(*8*, *14*). Increasing product tolerance formed the basis in attempts to improve

production of serine and isobutanol(*16*), as existing titers exceeded toxicity limits for these compounds. In the case of isobutanol, increased tolerance did not alter production titers. However, enhancing tolerance of serine was quite successful in improving production titers and yields(*8*, *14*). Interestingly, two successive studies that both aimed at evolving higher production of serine achieved different results using different strategies for enhancing mutagenesis. In one example, UV mutagenesis was used to enhance tolerance to 25g/l from 1.6g/l(*14*). This improved production titers 20% to 11.3 g/l. A further study used adaptive laboratory evolution to enhance tolerance to 100g/l, and further enhanced titers to 37g/l(*14*).

Many directed strain evolution applications for metabolic engineering require the direct coupling of fitness and production. Apart from those mentioned above, one innovative strategy to enhance production of carotenoids in *Saccharomyces cerevisiae* exploits the antioxidant properties of carotenoids(*19*) which protect against periodic hydrogen peroxide shocking. However, reliance on fitness coupled production greatly limits applications of evolution to production of only a few compounds. Several synthetic biology based strategies have been employed to decouple growth and production that result in improved titers of a target compound. One strategy, termed feedback-regulated evolution (FREP), was developed to employ a molecular sensor to gauge the concentration of a target metabolite that in turn alters mutation rates(*10*). The assembly of synthetic transcription factors that serve as actuators of operons that govern a selection mechanism allow the ability to evolve certain traits that normally have no natural selection mechanism. Unfortunately, this strategy is susceptible to spontaneous mutations that can alter the efficacy of synthetic constructs. “Escapees” are cells that recover normal growth by overcoming induced stress conditions. While strategies have been developed to prevent the formation of escapees(*20*), one recent example employs a synthetic circuit that controls the expression of a

maltose-utilizing enzyme using a biosensor of a target metabolite. Tight coupling of growth and production is yielded in a parent strain that is deficient of enzymes required for maltose utilization.

Besides acquiring a specific phenotype, directed strain evolution also offers insight into cell physiology that can aid in further engineering efforts. In one example, *E. coli* was evolved for enhanced succinate fermentation(*21*). *E. coli*'s native succinate fermentation pathway requires carboxylation of phosphoenolpyruvate (PEP) into oxaloacetate (OAA) using PEP carboxylase, encoded by *ppc*. However, the carboxylation reaction catalyzed by Ppc also releases phosphate from PEP, which would otherwise yield 1 ATP molecule through conversion of PEP into pyruvate. Through evolution, the cell managed to conserve this ATP molecule through two adaptations. First, phosphoenolpyruvate carboxykinase, Pck, became the main enzyme responsible for the carboxylation reaction. Pck, normally used during gluconeogenesis in the decarboxylation direction, here produces ATP in the carboxylation of PEP. Furthermore, a mutation inactivated *ptsI,* part of the phosphotransferase system, which usually functions to phosphorylate glucose while simultaneously converting PEP into pyruvate. By doing so, glucokinase became the main reaction responsible for phosphorylation of glucose, preventing PEP conversion into pyruvate. This step is particularly important for succinate production as converting pyruvate back to PEP, required for flux to succinate, requires an expenditure of 2 ATP molecules.

Several studies that sequenced genomes of evolved strains have noticed common mutations between strains of *E. coli* that have been evolved for enhanced growth within minimal media(*22*). Most common to these are mutations within RNA polymerase subunits (RNAP), *rpoB* and *rpoC*, respectively, which alter regulation of a broad range of cellular processes. Small

deletions in *rpoC* have been well characterized(*23*) that confer systematic transcriptional changes including down-regulation of motility, acid resistance, fimbria, and curlin genes. These adaptive RNAP mutations are believed to enhance growth in minimal media by decreasing open complex longevity. This in turn reduces transcription from promoters with short-lived open complexes such as rRNA, and increases transcription of promoters that require longer engagement of RNAP. Interestingly, this mutation is also accompanied by a decreased growth rate in rich media, presumably caused by decreased transcription of ribosomal units. Observations of mutations within these RNA polymerase subunits has been reported elsewhere(*22*)(*24*)(*14*)(*25*)(*6*)(*11*)(*26*).

Within this work, a strategy was used to accelerate the rate of mutagenesis that relies on *mutD5*. MutD5, a mutant of DNA polymerase III subunit ε, *dnaQ*, was first identified as a mutator gene of *E. coli* that can induce mutations at frequencies that are 50 to 100 times above wildtype levels(*27*). *mutD5* was shown to have two specific point mutations(*28*) that gave it properties of a dominant negative mutant(*29*), allowing its overexpression to confer a mutator phenotype that can later be removed when the gene is extracted. Here, directed strain evolution was carried out by subjecting a culture, harboring a plasmid that contains *MutD5*, to successive serial dilutions with a continuously applied selection pressure. More specific details pertaining to each evolution strategy are described below.

Promiscuous enzymes and spare parts

In certain cases, *E. coli* contains genetic material that may serve as spare parts if an essential enzyme is damaged. For instance, many enzymes have isozymes that can catalyze the same reaction; within *E. coli* two forms of pyruvate kinase exist. Many enzymes are also promiscuous, meaning that they can catalyze the same chemical reaction on different, yet similar substrates. One example is the native *E. coli* 3-hydroxyacid dehydrogenase (YdfG) which can catalyze the NADP+ dependent oxidation of L-serine, L-*allo*-threonine, 3-hydroxypropanoate, among other substrates(*30*). The cell also contains unique biosynthetic pathways that are used to produce the same metabolite. For instance, in *E. coli*, 5-phospho-α-D-ribose 1-diphosphate (PRPP) can be synthesized directly from either ribose 1,5-bisphosphate or ribose 5-phosphate. In these scenarios, each pathway may serve as a spare part in case one pathway is inhibited. Other works have demonstrated the ability of promiscuous enzymes or isozymes to serve as spare parts in case an otherwise essential enzyme is missing(*4*, *5*).

The vast number of unknown chemical reactions that can be catalyzed by known enzymes has been referred to as underground metabolism(*4*, *5*). Many secondary activities have been discovered and published in public databases such as BRENDA and MetaCyc(*31*, *32*). In *E. coli*, more than 260 secondary, or underground reactions, have been discovered that are outside of known metabolic networks(*4*, *5*). It has previously been demonstrated that promiscuous enzymes have the capability to be pieced together to form alternative metabolic pathways. In one case, a library of *E. coli* genomic DNA was transformed into strains of *ΔpdxB E. coli*(*33*). PdxB, 4-phosphoerythronate, catalyzes an essential reaction in synthesis of PLP (pyridoxal 5-phosphate) and therefore a *ΔpdxB* strain is a PLP auxotroph. By transforming the genomic DNA library into this strain, authors tested the ability for promiscuous enzymes to either replace essential mutant

function or to patch unrelated pathways that may form new routes for PLP biosynthesis. As a result, several strains were isolated that used the genomic library to patch unrelated metabolic pathways and thus restore PLP biosynthesis. While this work demonstrates that the cell contains genetic components that have the potential to form new metabolic pathways, it does not demonstrate the mechanisms by which existing genetic components can actively adapt to repair metabolism.

Promiscuous enzymes can be repurposed to complement mutant functions, however, many of these promiscuous reactions are increasingly being discovered that contribute to metabolite damage within the cell(*34*). Metabolite damage refers to side reactions to metabolites that can occur either enzymatically or non-enzymatically to produce wasteful or toxic products. To give an example, methylglyoxal is one of the major metabolites that causes metabolite damage(*35*). Methylglyoxal is formed from the spontaneous degradation of glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. It is also formed as a side product from triose phosphate isomerase during glycolysis(*36*). Excess methylglyoxal can covalently link with lysine, arginine, and cysteine that is either free within the cytosol or bound within a protein(*37*). Methylglyoxal metabolite toxicity therefore presents a major concern for maintenance of optimal metabolic function. Another prevalent example is a side reaction of carbon fixing RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase), which poorly discriminates oxygen and CO_2(*38*). A 2:5 ratio is present in the consumption of oxygen and CO_2, resulting in production of toxic side product glycolate-2-phospate rather than the desired 3-phosphoglycerate product(*39*). When engineering high flux metabolic pathways, metabolite damage becomes an increasingly relevant concern, as increasing flux and pool sizes of metabolites can exacerbate metabolite damage reactions(*40*). Repair enzymes are being discovered that can reverse metabolite damage. For

example, to combat methylglyoxal metabolite damage in *E. coli*, YhbO and YajL can repair glycated proteins including glyceraldehyde-3-phoaphate dehydrogenase and fructose bisphosphate aldolase. Strains deficient in *yhbO* and *yajL* are sensitive to glucose containing media and methylglyoxal(*41*).

Promiscuous enzymes exist that may serve as spare parts for metabolic repair mechanisms. Further, promiscuous reactions may cause unintended metabolite damage. Therefore, the extent and means by which *E. coli* can actively utilize its existing components to repair damage becomes increasingly unclear. Experiments described in this work aim to illustrate the mechanisms that govern metabolic repair.

References

1. Blank, D., L. Wolf, M. Ackermann, O. K. Silander, The predictability of molecular evolution during functional innovation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 3044–9 (2014).

2. Veeravalli, K., D. Boyd, B. L. Iverson, J. Beckwith, G. Georgiou, Laboratory evolution of glutathione biosynthesis reveals natural compensatory pathways. *Nat. Chem. Biol.* **7**, 101–105 (2011).

3. McLoughlin, S. Y., S. D. Copley, A compromise required by gene sharing enables survival: Implications for evolution of new enzyme activities. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13497–13502 (2008).

4. Guzmán, G. I., J. Utrilla, S. Nurk, E. Brunk, J. M. Monk, A. Ebrahim, Model-driven discovery of underground metabolic functions in Escherichia coli. *Proc. Natl. Acad. Sci.*

112, 929–934 (2014).

5. Notebaart, R. A., B. Kintses, A. M. Feist, B. Papp, Underground metabolism: network-level perspective and biotechnological potential. *Curr. Opin. Biotechnol.* **49**, 108–114 (2018).

6. LaCroix, R. A., T. E. Sandberg, E. J. O'Brien, J. Utrilla, A. Ebrahim, G. I. Guzman, R. Szubin, B. O. Palsson, A. M. Feist, Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of Escherichia coli K-12 MG1655 on glucose minimal medium. *Appl. Environ. Microbiol.* **81**, 17–30 (2015).

7. Lee, C. H., D. L. Gilbertson, I. S. Novella, R. Huerta, E. Domingo, J. J. Holland, Negative effects of chemical mutagenesis on the adaptive behavior of vesicular stomatitis virus. *J Virol*. **71**, 3636–3640 (1997).

8. Mundhada, H., K. Schneider, H. B. Christensen, A. T. Nielsen, Engineering of high yield production of L-serine in Escherichia coli. *Biotechnol. Bioeng.* **113**, 807–816 (2016).

9. Antonovsky, N., S. Gleizer, E. Noor, Y. Zohar, E. Herz, U. Barenholz, L. Zelcbuch, S. Amram, A. Wides, N. Tepper, D. Davidi, Y. Bar-On, T. Bareia, D. G. Wernick, I. Shani, S. Malitsky, G. Jona, A. Bar-Even, R. Milo, Sugar Synthesis from CO2 in Escherichia coli. *Cell*. **166**, 115–125 (2016).

10. Chou, H. H., J. D. Keasling, Programming adaptive control to evolve increased metabolite production. *Nat. Commun.* **4**, 2595 (2013).

11. Tenaillon, O., A. Rodríguez-Verdugo, R. L. Gaut, P. McDonald, A. F. Bennett, A. D. Long, B. S. Gaut, The Molecular Diverstiy of Adaptive Convergence. *Science (80-.).* **335**,

457–462 (2012).

12. Hughes, B. S., A. J. Cullum, A. F. Bennett, Evolutionary adaptation to environmental ph in experimental lineages ofEscherichia coli. *Evolution*. **61**, 1725–1734 (2007).

13. Dhar, R., R. Sägesser, C. Weikert, J. Yuan, A. Wagner, Adaptation of Saccharomyces cerevisiae to saline stress through laboratory evolution. *J. Evol. Biol.* **24**, 1135–1153 (2011).

14. Mundhada, H., J. M. Seoane, K. Schneider, A. Koza, H. B. Christensen, T. Klein, P. V. Phaneuf, M. Herrgard, A. M. Feist, A. T. Nielsen, Increased production of L-serine in Escherichia coli through Adaptive Laboratory Evolution. *Metab. Eng.* **39**, 141–150 (2017).

15. Almario, M. P., L. H. Reyes, K. C. Kao, Evolutionary engineering of Saccharomyces cerevisiae for enhanced tolerance to hydrolysates of lignocellulosic biomass. *Biotechnol. Bioeng.* **110**, 2616–2623 (2013).

16. Atsumi, S., T.-Y. Wu, I. M. P. Machado, W.-C. Huang, P.-Y. Chen, M. Pellegrini, J. C. Liao, Evolution, genomic analysis, and reconstruction of isobutanol tolerance in Escherichia coli. *Mol. Syst. Biol.* **6**, 449 (2010).

17. Sandberg, T., C. Lloyd, B. Palsson, A. Feist, Laboratory Evolution to Alternating Substrate Environments Yields Distinct Phenotypic and Genetic Adaptive Strategies. *Appl. Environ. Microbiol.* **83**, 1–15 (2017).

18. Argyros, D. A., S. A. Tripathi, T. F. Barrett, S. R. Rogers, L. F. Feinberg, D. G. Olson, J. M. Foden, B. B. Miller, L. R. Lynd, D. A. Hogsett, N. C. Caiazza, High ethanol Titers

from cellulose by using metabolically engineered thermophilic, anaerobic microbes. *Appl. Environ. Microbiol.* **77**, 8288–8294 (2011).

19. Reyes, L. H., J. M. Gomez, K. C. Kao, Improving carotenoids production in yeast via adaptive laboratory evolution. *Metab. Eng.* **21**, 26–33 (2014).

20. Liu, S.-D., Y.-N. Wu, T.-M. Wang, C. Zhang, X.-H. Xing, *ACS Synth. Biol.*, in press, doi:10.1021/acssynbio.7b00247.

21. Zhang, X., K. Jantama, J. C. Moore, L. R. Jarboe, K. T. Shanmugam, L. O. Ingram, Metabolic evolution of energy-conserving pathways for succinate production in Escherichia coli. *Proc Natl Acad Sci U S A.* **106**, 20180–20185 (2009).

22. Wannier, T. M., A. M. Kunjapur, D. P. Rice, M. J. McDonald, M. M. Desai, G. M. Church, Long-term adaptive evolution of genomically recoded Escherichia coli. *Doi.Org*, 162834 (2017).

23. Conrad, T. M., M. Frazier, A. R. Joyce, B.-K. Cho, E. M. Knight, N. E. Lewis, R. Landick, B. O. Palsson, RNA polymerase mutants found through adaptive evolution reprogram Escherichia coli for optimal growth in minimal media. *Proc. Natl. Acad. Sci.* **107**, 20500–20505 (2010).

24. Long, C. P., J. E. Gonzalez, A. M. Feist, B. O. Palsson, M. R. Antoniewicz, Fast growth phenotype of E. coli K-12 from adaptive laboratory evolution does not require intracellular flux rewiring. *Metab. Eng.* **44**, 100–107 (2017).

25. Sandberg, T., C. Lloyd, B. Palsson, A. Feist, Laboratory Evolution to Alternating Substrate Environments Yields Distinct Phenotypic and Genetic Adaptive Strategies. **83**,

1–15 (2017).

26. Sandberg, T. E., C. P. Long, J. E. Gonzalez, A. M. Feist, M. R. Antoniewicz, B. O. Palsson, Evolution of E. coli on [U-13C]glucose reveals a negligible isotopic influence on metabolism and physiology. *PLoS One*. **11**, 1–14 (2016).

27. Degnen, G. E., E. C. Cox, Conditional mutator gene in Escherichia coli: isolation, mapping, and effector studies. *J. Bacteriol.* **117**, 477–487 (1974).

28. Takano, K., Y. Nakabeppu, H. Maki, T. Horiuchi, M. Sekiguchi, Structure and function of dnaQ and mutD mutators of Escherichia coli. *MGG Mol. Gen. Genet.* **205**, 9–13 (1986).

29. Maruyama, M., T. Horiuchi, H. Maki, M. Sekiguchi, A dominant (mutD5) and a recessive (dnaQ49) mutator of Escherichia coli. *J. Mol. Biol.* **167**, 757–771 (1983).

30. Fujisawa, H., S. Nagata, H. Misono, Characterization of short-chain dehydrogenase/reductase homologues of Escherichia coli (YdfG) and Saccharomyces cerevisiae (YMR226C). *Biochim. Biophys. Acta - Proteins Proteomics*. **1645**, 89–94 (2003).

31. Schomburg, I., A. Chang, O. Hofmann, C. Ebeling, F. Ehrentreich, D. Schomburg, BRENDA: A resource for enzyme data and metabolic information. *Trends Biochem. Sci.* **27**, 54–56 (2002).

32. Keseler, I. M., J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I. T. Paulsen, M. Peralta-Gil, P. D. Karp, EcoCyc: A comprehensive database resource for Escherichia coli. *Nucleic Acids Res.* **33**, 334–337 (2005).

33. Kim, J., J. P. Kershner, Y. Novikov, R. K. Shoemaker, S. D. Copley, Three serendipitous

pathways in E. coli can bypass a block in pyridoxal-5'-phosphate synthesis. *Mol. Syst. Biol.* **6**, 436 (2010).

34. Linster, C. L., E. Van Schaftingen, A. D. Hanson, Metabolite damage and its repair or pre-emption. *Nat. Chem. Biol.* **9**, 72–80 (2013).

35. Richard, J. P., Acid-Base Catalysis of the Elimination and Isomerization Reactions of Triose Phosphates. *J. Am. Chem. Soc.* **106**, 4926–4936 (1984).

36. Richard, J. P., Mechanism for the formation of methylglyoxal from triosephosphates. *Biochem. Soc. Trans.* **21**, 171–174 (1993).

37. Richarme, G., M. Mihoub, J. Dairou, L. Chi Bui, T. Leger, A. Lamouri, Parkinsonism-associated protein DJ-1/park7 is a major protein deglycase that repairs methylglyoxal- and glyoxal-glycated cysteine, arginine, and lysine residues. *J. Biol. Chem.* **290**, 1885–1897 (2015).

38. Erb, T. J., J. Zarzycki, Biochemical and synthetic biology approaches to improve photosynthetic CO2-fixation. *Curr. Opin. Chem. Biol.* **34**, 72–79 (2016).

39. Walker, B. J., A. VanLoocke, C. J. Bernacchi, D. R. Ort, The Costs of Photorespiration to Food Production Now and in the Future. *Annu. Rev. Plant Biol.* **67**, 107–129 (2016).

40. Sun, J., J. G. Jeffryes, C. S. Henry, S. D. Bruner, A. D. Hanson, Metabolite damage and repair in metabolic engineering design. *Metab. Eng.* **44**, 150–159 (2017).

41. Abdallah, J., M. Mihoub, V. Gautier, G. Richarme, The DJ-1 superfamily members YhbO and YajL from Escherichia coli repair proteins from glycation by methylglyoxal and glyoxal. *Biochem. Biophys. Res. Commun.* **470**, 282–286 (2016).

Chapter 2 : Metabolic repair through emergence of new pathways

This section was adapted from Pontrelli S., Fricke R.B.C, Teoh S. T., Laviña W. A., Putri S.P., Fitz-Gibbon S., Chung M., Pellegrini M., Fukusaki E., Liao J.C. Metabolic Repair through Emergence of New Pathways. *Nature* (under review)

Abstract

Microorganisms can derive all metabolites and cofactors essential for growth from a simple carbon source through a highly evolved metabolic system. Damage of pathways may significantly affect cell growth and physiology, but the extent to which metabolic pathways can repair damage is not well understood. Here we use a *ΔpanD* (coding for aspartate 1-decarboxylase) strain of *Escherichia coli* to demonstrate that the metabolic system can repair pathway damage by evolving new metabolic functions and repurposing existing enzymes. The *ΔpanD* strain cannot synthesize β-alanine, which is a required precursor for coenzyme A (CoA). Using directed strain evolution, we showed that *E. coli* repaired this pathway damage using three distinct strategies. The first involved significant rewiring and repurposing of the uracil synthesis and degradation (Rut) pathways. The second evolved through the emergence of a novel pathway involving a gain-of-function mutation within ornithine decarboxylase (SpeC) to alter both reaction and substrate specificity. The mutated enzyme functions as a bifunctional decarboxylase and oxidative deaminase to produce 3-aminopropanal from 2,4-diaminobutyrate, with a kcat/km improvement over 3000 fold. After deletion of both the Rut and SpeC pathways, yet another independent pathway emerged through evolution, demonstrating the vast capacity of metabolic repair.

Main Text

Unlike DNA, metabolic systems are not known to possess a repair mechanism. Some key reactions in metabolism are catalyzed by isozymes that may serve as spare parts(*1*) to avoid catastrophic consequences in pathway damage: when reactions of a metabolic pathway are blocked. Alternatively, promiscuous enzymes may be present to complement a lost function(*2*)(*3*). Other enzymes are highly specialized and catalyze unique reactions, and when damaged may significantly compromise growth and alter physiology. The wide range of possible reactions catalyzed by cryptic and promiscuous enzymes has potential to form pathways that can restore metabolic function when these specialized reactions are damaged (*4*). However, the mechanisms and extent by which the cell can actively utilize its existing components to repair metabolism is poorly understood. Here, we demonstrate that new pathways can be formed with a combination of repurposed enzymes and rewired metabolic networks, and with the emergence of new biochemical function. We use a *ΔpanD* strain of *E. coli*, which is incapable of producing the β-alanine required for synthesis of CoA, as an example to show that β-alanine pathway damage can be repaired by formation of at least three distinct metabolic pathways. Our results demonstrate the intrinsic pliability of biological systems, and that new pathways can form to repair damage at the metabolic level.

Aspartate 1-decarboxylase (PanD), is the only enzyme capable of β-alanine synthesis in *E. coli.* In bacteria, fungi, and plants, β-alanine is a precursor to pantothenate, which is in turn a required metabolite for the synthesis of coenzyme A (CoA) in all organisms(*5*). In animals, β-alanine is synthesized as a precursor to carnosine, which is found at high concentrations within skeletal muscle tissue and the central nervous system and is used for various physiological

purposes(*6*). Without CoA, the cell is incapable of carrying out essential cellular processes that include the TCA cycle, fatty acid biosynthesis, and synthesis of acetyl-CoA which is used as a building block for many essential compounds(*5*). Therefore, unless β-alanine or pantothenate are supplemented, a *ΔpanD* strain cannot grow on minimal media alone. Unlike most decarboxylases that use pyridoxal-5′-phosphate (PLP) as a cofactor, PanD uses a covalently-bound pyruvoyl cofactor(*7*). PanD is first translated as an inactive protoenzyme that is cleaved into an α and β-subunit, triggered by the activator PanZ(*8*). This likely serves as an additional regulatory element to control intracellular levels of pantothenate.

Several other pathways are believed to exist in other organisms to supply β-alanine: degradation of propionate into malonic semialdehyde (MSA) and subsequent transamination(*9*), a reductive uracil degradation pathway using dihydrouracil as an intermediate(*10*), and an oxidative degradation of spermine into β-alanine using 3-aminopropanal as an upstream precursor(*11*). However, within *E. coli*, these pathways are not known to exist.

We first attempted to repair a *ΔpanD* mutant by repeatedly passing cultures that contain limited amounts of β-alanine. This approach did not yield successful results after 20 passages. We then sought to enhance the rate of mutagenesis to generate extensive metabolic innovations within a shorter timeframe. This was accomplished by overexpressing a mutator gene, *mutD5*(*12*), which compromises DNA mismatch repair and proofreading in a dominant negative manner. Within 20 serial dilutions, 12 of 15 independent cultures overcame β-alanine auxotrophy with an average number of 283.5 mutations per genome ($\sigma = 148.2$, Table 2-1). An isolated clonal strain, PS1, was further studied to investigate how the strain repaired the damaged β-alanine synthesis pathway. Interestingly, although many amino acid decarboxylases exist within *E. coli*, none emerged with function sufficient to directly complement PanD. Rather, the

strain re-routed metabolism though pyrimidine synthesis and uracil degradation pathways (Figure 2-1A). This was accomplished solely through the modulation of enzyme expression levels rather than changing enzyme functions. Four essential mutations were identified that allowed for this pathway re-routing. The first two mutations identified are within repressors RutR (L55P) and CsiR (S144P) (Figure 2-1A). These mutations enable derepression of the pyrimidine utilization (Rut) pathway and 4-aminobutyrate transaminase (GabT), respectively (Figure 2-1B). Together these form the core enzymatic constituents of the pathway. The Rut pathway degrades uracil into 3-hydroxypropionic acid (3HP) as a terminal product(*13*). However, GabT is a promiscuous transaminase that is known to convert MSA, the penultimate metabolite of the Rut pathway, into β-alanine(*14*), and therefore is able to reroute the terminal product of the Rut pathway. Deletion of either RutABC or GabT from PS1 abolishes the evolved phenotype, confirming their essential contributions (Figure 2-1C).

The third essential mutation was acquired in the final enzyme of the Rut Pathway(*13*), 3-hydroxy acid dehydrogenase, YdfG (K108E), which diminishes but does not abolish the enzyme function (Figure 2-1D). While this mutation may serve to redirect the terminal product of the Rut pathway to β-alanine, deletion of YdfG from PS1 abolishes the evolved phenotype (Figure 2-1C). Because MSA is a toxic intermediate(*13*), a minor amount of YdfG activity may be essential to prevent toxic buildup. GabT provides a thermodynamically reversible transamination between MSA and β-alanine, and consequently, we observed toxicity caused by excess β-alanine supplementation (Figure 2-1E). We demonstrated the ability of YdfG to relieve this toxicity as overexpression within PS1 restores growth with excess β-alanine (Figure 2-1E).

The fourth significant mutation on Upp (L178P), uracil phosphoribosyltransferase, completely abolishes the enzyme activity (Figure 2-1F). Upp catalyzes the synthesis of UMP

(uridine 5'-monophosphate), the precursor for all pyrimidine nucleotides, from uracil and PRPP (5-phospho-α-D-ribose 1-diphosphate). This reaction recycles the uracil formed as a degradation product from nucleic acids(*15*). Intracellular concentrations of metabolites were compared between PS1 and a *ΔpanD* parent strain and showed decreased concentrations of pyrimidine nucleotides within PS1 (Figure 2-41), suggesting that this mutation enables sufficient outflux of uracil into the Rut pathway. We attempted to reconstruct the pathway within an unevolved *ΔpanD* strain, but with only overexpression of the Rut operon and GabT. In this strain, additional uracil supplementation was required (Figure 2-1G). However, when *upp* was further deleted, growth was observed without additional supplementation (Figure 2-1G). When each of the four identified mutations were individually reverted within PS1 the repaired phenotype was either weakened or completely abolished (Figure 2-1C).

To determine whether this same pathway formed in the remaining 11 repaired *ΔpanD* strains, we deleted the genes *rutABC* from each one. This resulted in an abolished growth phenotype within all but two strains, suggesting the existence of additional mechanisms of metabolic repair. We then obtained strains that specifically overcame auxotrophy independent of uracil and the Rut degradation pathway. A double deletion Δ*panD* Δ*rutABC* mutant was subjected to serial dilutions with limiting β-alanine and expression of mutator *mutD5*. After 85 dilutions, 4 of 12 independent cultures again repaired β-alanine metabolic pathway damage. Genomes of all strains were sequenced to reveal an average of 149 mutations per genome (σ = 36.1,Table 2-1).

All double Δ*panD* Δ*rutABC* strains with repaired phenotypes, in addition to the two remaining single *ΔpanD* repaired strains, acquired a mutation within the same residue of ornithine decarboxylase, SpeC (G655S or G655A). Deletion of SpeC from any of these strains

resulted in complete abolishment of the restored phenotype (Figure 2-2A), demonstrating its essentiality. Reversion of the SpeC point mutation was tested in one strain (PR11) and abolished the growth phenotype (Figure 2-2A).

While SpeC normally functions as an ornithine decarboxylase, 2,4-diaminobutyrate (24DAB) is a structural homolog of ornithine. The analogous decarboxylation product of 24DAB is 1,3-propanediamine, which may serve as an upstream precursor of β-alanine (Figure 2-5). We therefore hypothesized that SpeC acquired a mutation that allowed it to expand its substrate specificity. However, *in vitro* assays with purified enzyme revealed that 24DAB decarboxylase activity was detected from wildtype SpeC (Figure 2-5), but both mutated SpeC variants completely lost this activity.

Surprisingly, in the absence of 13PDA formation, 24DAB consumption was still observed from mutated SpeC variants *in vitro*. Similar PLP dependent amino acid decarboxylases have reported side reactions that result in simultaneous decarboxylation and deamination of the α-amino group to yield the respective aldehyde product(*16*). In this case, product formation from 24DAB yields 3-aminopropanal (3AP), which can further be oxidized to β-alanine, possibly by betaine-aldehyde dehydrogenase, BetB(*17*). A coupled *in vitro* assay of purified BetB and SpeC yielded β-alanine from 24DAB (Figure 2-2B) with SpeC variants. Previous reports of a simultaneous decarboxylation and deamination reaction involve a net reaction that requires the consumption of dissolved oxygen and formation of ammonia and hydrogen peroxide(*16*). We confirmed the presence of this reaction as β-alanine formation was detected in a 1:1 ratio with both ammonia and hydrogen peroxide (Figure 2-2CD).

The native enzyme betaine-aldehyde dehydrogenase, BetB(*17*), has been reported to have secondary activity for converting 3AP to β-alanine. Deletion of BetB abolishes the evolved

phenotype, confirming its contribution (Figure 2-2A). BetI, the corresponding repressor of BetB is mutated within PR11 (BetI I15F), and a measured increase in expression of BetB demonstrates the deleterious consequence of this mutation (Figure 2-2E). Reversion of the BetI mutation also diminishes the growth rate of PR11 (Figure 2-2A). Upstream of these reactions within the pathway, we suspected that aspartate semialdehyde, a central metabolite essential for lysine, threonine, and methionine biosynthesis, may be a direct precursor of an aminotransferase reaction that can produce 24DAB directly (Figure 2-2F). A mutation was found on CsiR (Q99G), the repressor of GabT, within PR11 to suggest its involvement. An increase in GabT expression within PR11 was measured to indicate that CsiR is deleteriously mutated (Figure 2-2E). GabT, which functions as a β-alanine aminotransferase within PS1, was determined to contain the activity of 24DAB aminotransferase needed within PR11 (Figure 2-6). To further confirm the pathway, we proceeded to reconstruct the complete pathway within an unevolved *ΔpanD* strain. GabT, BetB, and SpeC G655S overexpression was sufficient to rescue growth in minimal media without β-alanine supplementation (Figure 2-2G). Here, the identified mutations that contributed to metabolic repair increase expression of pathway constituents (CsiR and BetI) and alter activity (SpeC).

Activity of the SpeC variants using 24DAB, ornithine, and lysine as substrates were tested *in vitro.* While wildtype SpeC has the simultaneous decarboxylation and oxidative deamination activity using ornithine as a substrate (Figure 2-3AC), this activity was almost undetectable with 24DAB. The mutations in SpeC (G655S or G655A) increased Kcat and decreased Km of this reaction for ornithine and 24DAB (Figure 2-3BC). G655A and G655S mutants presented respective 3300 and 640-fold K_{cat}/K_m increases from 24DAB (Figure 2-3C). This gain of function may be the key to the metabolic repair.

SpeC G655A is the only variant capable of using all three substrates for a coupled decarboxylation and deamination reaction (Figure 2-3C, Figure 2-7). On the other hand, wildtype SpeC is the only variant capable of a single decarboxylation reaction on all three substrates (Figure 2-8), while G655A and G655S variants completely lost the ability to perform this reaction on 24DAB and lysine. It appears that metabolic pathway repair takes advantage of the broader reactivity and substrate range made available from existing enzymes that may further evolve to have a highly specified function.

We sought to further test the ability of the cell to repair metabolic pathway damage. To this end, a *ΔpanD ΔrutABC ΔspeC* strain was constructed. With additional overexpression of *mutD5*, this strain was evolved for 45 passages to yield 1 of 12 strains that repaired β-alanine auxotrophy (Mel6) (Figure 2-9). We were able to identify three essential enzymes. These enzymes are arginine decarboxylase (SpeA), S-adenosylmethionine decarboxylase (SpeD), and spermidine synthase (SpeE). Deletion of these genes from mel6 abolished the evolved phenotype, and further, an increase in expression was measured (Figure 2-10). These enzymes normally function to synthesize spermidine within *E. coli*, suggesting that an additional pathway may have formed that degrades polyamines into β-alanine using unknown enzymes. Observing the formation of an additional pathway by which β-alanine auxotrophy can be overcome illustrates the remarkable capability of the metabolic system to repair damage using only endogenous genetic components.

To determine the extent of the repair capability, we selected an additional 34 metabolic auxotrophs, unable to grow in glucose minimal media (Table 2-2), and subjected them to a minimum of 20 serial dilutions with limiting nutritional supplementation, or until the ability to grow without supplementation was acquired. These 34 mutants were chosen to represent

deficiencies in a broad range of cofactor and amino acid biosynthetic pathways. Seven strains acquired the ability to grow without nutritional supplementation (Figure 2-11). These strains all recovered readily, requiring a small number of serial dilutions (mean = 3, Figure 2-11) and number of mutations (mean = 3,Table 2-1), consistent with previous works demonstrating that certain metabolic functions can be easily restored(*18*). For the remaining 27 strains, we overexpressed the *mutD5* gene. Of these strains, four (Figure 2-11) gained the ability to grow without nutritional supplementation with an average of 446.8 mutations per genome (σ = 287.5,Table 2-1). These results indicate that metabolic repair can extend to a diverse group of metabolic pathways.

The cell is innately equipped with a vast set of endogenous genetic components that offer the potential to repair damaged metabolic pathways. The examples here demonstrate the active utilization of these components both by repurposing existing enzymes and by evolving new enzymatic function. Pathways to provide β-alanine were repeatedly evolved, demonstrating the plasticity of metabolism. Further, the evolved pathways are distinct from all currently known routes, indicating how β-alanine can be produced in a diverse number of ways within many different organisms. Metabolic pathway repair poses a challenge for the development of robust engineered phenotypes and metabolic targeted drugs, however, it also serves as an opportunity for the emergence of new biochemical function that can further be exploited in engineering microbes for bioproduction.

Acknowledgements: This research was fully supported by Japan Science and Technology (JST), Strategic International Collaborative Research Program, SICORP and National Science Foundation (NSF) MCB-1139318 for JP-US “Metabolomics for Low Carbon Society.” SFG

acknowledges support from a QCB Collaboratory Postdoctoral Fellowship, and the QCB Collaboratory community directed by Matteo Pellegrini.

Materials and Methods

Strains, reagents, and plasmids.

All chemicals and reagents were purchased from Sigma-Aldrich unless otherwise noted. *Escherichia coli* single knockouts strains were obtained from the Keio collection(*19*), and parent strain BW25113 was obtained from Thermo-scientific.

Primers were purchased through Integrated DNA Technologies (IDTdna.com). RBS sequences were optimized using RBS Calculator v2.0(*20*, *21*). All PCR reactions were first performed using KOD Hot-Start DNA polymerase (EMD Millipore). If these reactions failed, reactions were repeated with KOD Xtreme Hot-Start Polymerase (EMD Millipore). Assembly of fragments were performed using T4 DNA polymerase (NEB) using the following protocol: Gel purified PCR fragments with 20bp overlaps were mixed in equimolar amounts. Approximately 300ng of the fragment mixture was combined with NEB buffer #2 and 0.3µl T4 polymerase to a final volume of 10µl. Reaction was incubated at room temperature for 5 minutes, transformed into *E. coli* XL1Blue (Agilent), and selected on LB agar plates containing appropriate antibiotics. Gene deletions on strains were carried out using λ Red recombinase as previously described(*22*). MutD5 gene was obtained by amplifying DnaQ from XL1Red strains (Agilent). His-tagged proteins are cloned and expressed within pCDFDuet backbones (Addgene).

Growth conditions

Growth curves for PS1, PR11 and Mel6 derived strains were obtained within cultures of glucose M9 minimal media: M9 Minimal Salts (Fisher Scientific), 0.4% glucose, 1mM MgSO4, 0.1mM $CaCl_2$, 1mg/L thiamine. Antibiotics were used at the following concentrations:

Kanamycin sulfate 50 μg/mL, disodium carbenicillin (Gold Biotechnology) 100 μg/mL, chloramphenicol 20 μg/mL, and spectinomycin (Gold Biotechnology) 100 μg/mL.

Optical density was measured using Agilent 8453 Spectrophotometer. During phenotypic analysis of PS1, PR11and Mel6 (and derivatives), colonies were inoculated into 3ml glucose minimal media containing limiting concentrations of β-alanine (1μM) and grown for 24 hours at 37 °C. Cells were used in a 1:1000 inoculation into selection media containing appropriate supplements unless otherwise noted. Growth curves were measured in triplicates.

Evolution of suppressor phenotypes

Strains were evolved to overcome various auxotrophies by gradually reducing the availability of nutritional supplementation. Mutant strains were inoculated into a preculture of LB containing an appropriate antibiotic, grown overnight at 37 °C, and further used to inoculate 1:100 into minimal medium with a respective carbon source and antibiotic. Strains were passed into fresh minimal media at a 1:1000 dilution with limiting supplementation sufficient to maintain growth between OD600 0.4 and 0.6. These cultures were diluted daily. If growth exceeded OD600 0.6, the amount of nutritional supplementation was reduced two-fold for the next dilution. If cultures grew past OD600 1.0, they were passed into minimal media without further supplementation. If this culture grew, they were streaked onto minimal media agar plates without supplementation. Genotypes of growing colonies were verified using PCR and their phenotype was reconfirmed by inoculating directly from the plate into minimal media. Unless otherwise specified, four independent cultures of each strain were subjected to serial dilutions.

Within certain strains, mutation rates were accelerated with use of a mutator plasmid that expresses MutD5. MutD5(*23*) is a dominant negative mutant of DNA polymerase III subunit ε, DnaQ, which catalyzes the 3' to 5' proofreading during DNA replication(*24*).

Strains were either evolved for a minimum of 20 dilutions, or until the ability to grow in minimal media without supplementation was obtained. Occasionally, the addition of a mutator plasmid into a given strain results in the loss of growth within minimal media even with nutritional supplementation. If this occurred, the strain was no longer subjected to serial dilutions.

Curing Plasmids

Plasmids were cured using acridine orange. Overnight cultures were made in LB without antibiotics. The following day, 1% inoculations were made into fresh LB cultures containing 2µL of acridine orange (Thermo Fisher) for every 1ml of LB. These were grown overnight at 42 °C, streaked onto LB plates, and then screened for the correct lack of antibiotic resistance. The lack of plasmid was then confirmed through PCR verification.

Quantification of amino acids and polyamines using HPLC

Amino acids and polyamines were derivatized using OPA and FMOC reagents (Agilent) and analyzed using a Zorbax Eclipse AAA HPLC column (5µM beads, 4.6mm x 150 mm). All protocols were obtained from Agilent.

GC-MS analysis of MCF derivatives

All columns and instruments were purchased from Agilent Technologies. GC-MS data was obtained using a 6890/5973 GC/MS and DB-624UI (GC-MS) column. The oven temperature was initially set at 60 °C for 2 minutes. Then a ramping of 16 °C/min was applied to a gradient reaching 180 °C, followed by a 3 minute hold. Next, a 40 °C/min ramping was applied to a gradient reaching 220 °C, followed by a 3 minute hold. The 40 °C/min ramping was again

applied to a temperature of 240 °C, followed by a 6 minute hold. The flow through the column was held constant at 1.8mL He/min. The injection volume was 2µL and the split ration 20:1. The temperature of the inlet was 180 °C and detector set to 250 °C.

Preparation of cell lysate for enzyme assays

Respective plasmids were transformed into *E. coli* BL-21 DE3 (Invitrogen) and grown overnight. These cultures were used for a 1% inoculation into LB, grown to mid-log phase, and induced using 0.1mM isopropyl-h-D-thiogalactopyranoside (IPTG) overnight at 30 °C. Cells were harvested by centrifugation, resuspended into the respective buffer used for enzyme assays and lysed using Qiagen TissueLyser II.

YdfG assay

Plasmids pALQ82 and pALQ83 were used to amplify wildtype and mutant (K108E) enzymes, respectively, within wildtype strain BL21. Standard reaction mixture was composed of 50mM Tris-HCL buffer (pH 8.5), 0.5mM NADP+ and 5mM 3-hydroxypropionic acid. The reaction was started by adding 10µl of cell lysate to 200µL of reaction mixture, and monitoring OD340 increase over time using Bio-TEK Powerwave XS plate reader.

Upp assay

Plasmids pALQ84 and pALQ85 were used to express wildtype and mutant (L178P), respectively, within wildtype strain BL21. This assay was adapted from previous study(*25*). Standard reaction mixture was composed of 50mM Tris-HCL buffer (pH 8.5), 1mM GTP (Sigma), 2mM PRPP (Sigma), 2mM Uracil (Sigma), and 50mM $MgCl_2$. Cell lysate was normalized to 1.5 mg/mL using Bradford Reagent. Reaction mixture was started by adding 50

μL of lysate to 1mL of reaction buffer. The reaction was stopped by the addition of 80 μL of reaction mixture to 20 μL of ice cold 100% w/v trichloroacetic acid and vortexed. The mixture was kept at -20 °C for 6 hours, centrifuged at maximum speed at 4 °C, and the supernatant was collected. Uracil consumption was measured using high-performance liquid chromatography (HPLC) previously described(*26*).

SpeC assay

Assay for SpeC is adapted from a previous study(*27*). Plasmids pALQ142, pALQ143, and pALQ173 were used to express his-tagged SpeC variants G655S, wildtype, and G655A, respectively. Protein was purified using HisPur Ni-NTA Spin Purification Kit (ThermoFisher). Standard reaction mixture contained 100mM PBS (pH 8), 1mM DTT, 1mM GTP, 100μM PLP and 20mM ornithine, 2,4-diaminobutyrate or lysine. Reactions were initiated by addition of 30μl of enzyme to 1mL of enzyme mixture to reach a final concentration around 0.05mg/ml. Concentrations were later normalized using Bradford reagent. Reactions were allowed to proceed overnight and were stopped using Whatman Mini-UniPrep Nylon Syringeless Filters. Quantification of consumption of 2,4-diaminobutyrate was measured by HPLC using ZORBAX Eclipse AAA column. Samples were derivatized using OPA and FMOC reagents (Agilent)(*28*). Product formation was later confirmed after MCF derivatization and analysis using GCMS.

GabT assay for 2,4-diaminobutyrate transaminase

His-tagged GabT was overexpressed from pALQ144 and purified using HisPur Ni-NTA Spin Purification Kit (ThermoFisher). The standard reaction mixture contained 100mM Tris-HCL buffer (pH 8), 100μM PLP, 20mM 2,4-diaminobutyrate, 20mM α-ketoglutarate and 1mM DTT. The reaction was allowed to proceed overnight at room temperature and was stopped using Whatman Mini-UniPrep Nylon

Syringeless Filters. Formation of glutamate was measured using ZORBAX Eclipse AAA column on HPLC and was further derivitized using MCF derivatization and confirmed on GCMS.

Coupled SpeC and BetB assay

Decarboxylase and oxidative deaminase activity of SpeC was measured in a coupled reaction with BetB. The reaction mixture contained the following components: 25µg/ml purified SpeC, 50µg/ml purified BetB, 1mM GTP, 1mM NAD+, 100µM PLP, 20mM 24DAB or ornithine, and 1mM DTT in 100mM pH 8.0 PBS. Increase in OD340 was measured over time at 37°C. When identifying product formation, the reaction was filtered using Whatman Mini-UniPrep Nylon Syringeless Filters, derivatized using MCF derivatization, then analyzed using GCMS.

Hydrogen peroxide assay

Hydrogen peroxide formation was measured simultaneously with β-alanine formation in a reaction coupling SpeC, BetB, and Peroxidase from horseradish (Sigma). The peroxidase reacts with H_2O_2, phenol, and 4-aminoantipyrine to produce a colorimetric readout at OD505(*29*). 3-aminopropanal is oxidized into β-alanine using BetB with a simultaneous conversion of NAD+ into NADH, producing a readout at 340nm. The reaction mixture contained 25µg/ml SpeC G655S, 50µg/ml BetB, 1mM GTP, 1mM DTT, 100µM PLP, 1mM NAD+, 20mM 24DAB, 4mM phenol, 6mM 4-aminoantipyrine, and approximately 100U/ml peroxidase, in PBS pH 8.0 100mM. The reaction was allowed to proceed at 37°C.

Ammonia assay

Ammonia formation was measured by coupling SpeC with glutamate dehydrogenase (Sigma). Other reactions involving SpeC G655S use GTP as an activator, however, we observed that GTP inhibits the activity of glutamate dehydrogenase. Therefore, we used CTP as an activator as it has previously been demonstrated to serve this purpose(*30*). The reaction mixture contained 25µg/ml SpeC G655S, 4U/ml glutamate dehydrogenase, 5mM 24DAB, 1mM CTP, 1mM DTT, 100µM PLP, 20mM αKG, and 1mM NADH. The reaction was allowed to proceed at 37°C and OD340 was measured over time.

MCF derivitizaion

Samples used for GCMS analysis were derivatized using methylchloroformate (MCF) derivatization adapted from a previously reported protocol(*31*). 100µL of sample was added to 100µL of 2M NaOH. 167µL of methanol was added, followed by 34µL of pyridine, 20µL of MCF, and vortexed for 30 seconds. Immediately, another 20µL of MCF was added, followed by another 30 seconds of vortexing. 400µL of chloroform was added, followed by an additional 10 seconds of vortexing. Next, 400µL of 50mM sodium bicarbonate was added, followed by 10 seconds of vortexing. A glass Pasteur pipette was used to remove the top phase, and 100mg of anhydrous sodium sulfate was added to the chloroform solution to bind the remaining water. A glass Pasteur pipette was used to transfer the entire dried chloroform solution to a glass vial for injection into the GCMS. Samples that were not immediately analyzed were stored at -80 °C.

Genomic sequencing

Genomic DNA was extracted from relevant samples using Qiagen DNeasy Blood & Tissue Kit. All samples were diluted to 0.2 ng/µL using a Qubit measurements. This was used as input sample for Nextera XT sample preparation (Illumina). Final libraries were eluted in Qiagen EB

buffer with 0.1% Tween 20. The individual sample libraries were normalized to equimolar concentrations and diluted to 10nM solution. Samples were sequenced using 100bp SE reads on an Illumina HiSeq2000 sequencing system with single end reads to a minimum of 10x coverage, average = 42x coverage.

Sequencing data analysis

Adapter sequences were removed from reads using Trim Galore! (http://www.bioinformatics.babraham.ac.uk/) with quality trimming turned off. Trimmed reads were mapped using BWA-MEM v.0.7.12-r1039(*32*) to the Escherichia coli str. K-12 substr. MG1655 genome (NCBI Accession NC_000913). Variant discovery and filtering was done with GATK v 3.7-0-gcfedb67(*33*) using HaplotypeCaller in GVCF mode with ploidy 1 and ploidy 3 as needed, followed by GenotypeGVCFs, and finally VariantFiltration setting a minimum QD of 2. SnpEff(*34*) was used to determine the context of the variants and predict the functional impact. Additional custom scripts were used to identify variants of interest.

Reverse transcriptase quantitative PCR

Genomic DNA was purified using Qiagen DNeasy Blood and Tissue Kit for initial determination of primer efficiency using Luna Universal qPCR Master Mix. The DNA was initially diluted to approximately 1 ng/uL and three additional subsequent serial dilutions were used to determine primer efficiency. Primers were designed using Primer3(*35*) software to have annealing temperatures at 60 °C and product sized between 75 and 150bp within their respective target gene. Primers are listed below in Table 2-6, and are verified to have efficiency between 90 and 105%. Gene frr, ribosome-recycling factor, was used as a reference gene(*36*).

RNA was purified from cultures of glucose minimal media at OD600 0.4-0.6. These cultures were inoculated at 1% with precultures originating from glucose minimal media with 1μM β-alanine. Approximately $2x10^8$ cells were added to two volumes of Qiagen RNAprotect Bacteria Reagent, vortexed and centrifuged. 200μL of 15 mg/mL of lysozyme with 30μL of Proteinase K (NEB) was added, and the mixture was shaken at room temperature for 45 minutes. RNeasy RNA purification protocol was followed with additional DNase Digestion (Qiagen) step to remove genomic DNA. RT-qPCR was performed using Luna Universal One-Step RT-qPCR Kit and results were analyzed using Bio-Rad CFX Manager 2.0.

Genomic point mutations

Point mutations were introduced using λ-red recombinase system(*22*). Linear fragments were designed similar to those used for gene deletions, which amplify a kanamycin cassette from pKD13 that have 50bp overlaps intended for site specific homologous recombination. However, to introduce a point mutation, one of the overlaps was extended 400-1000bp and contains a mutation that can be introduce in the genome when recombined. Here, the overlap contained a wildtype sequence so that when the cassette is recombined several hundred basepairs upstream or downstream of a mutated gene, a point mutation will be repaired.

Briefly, primers were designed that amplify a kanamycin cassette from pKD13 (Table 2-7). A 400-1000bp overlap was also amplified from a wildtype BW25113 genome (Table 2-8). These were attached using SOE PCR. The target strain, harboring pKD46, was grown overnight at 30 °C with ampicillin. The next morning, it was inoculated 1% into fresh LB containing ampicillin and 1mM L-arabinose and grown until OD600 0.6. This was then washed 3 times in ice cold 10% glycerol. 400ng of linear fragment were added and the reaction was electroporated. The strain was rescued at 37 °C for one hour and then plated onto kanamycin plates. Successful

colonies that formed were sequenced around the area of integration using Sanger sequencing to verify successful point mutation reversion.

Metabolomic Analysis

Cultures grown until OD600 = ~1 were harvested for metabolome analysis by fast filtration of 10 mL culture broth through PTFE membrane filters (pore size 0.45 µm, diameter 47 mm; Millipore, MA, USA). The filter-bound cells were transferred to 2 mL tubes and flash-frozen in liquid nitrogen to quench metabolism before storage at -80°C. Metabolite extraction was performed by adding 1.8 mL extraction solvent (methanol/water/chloroform in 5:2:2 ratio, additionally spiked with 20 µg/mL ribitol and 30 µg/mL camphosulfonic acid as internal standards) and incubating at -30°C for 1 h. For each sample, 1.2 mL solvent containing extracted metabolites was then mixed with 600 µL ultrapure water, vortexed briefly then centrifuged at 9390 × g, 4°C for 3 min to separate the aqueous and organic phases. The aqueous phase containing hydrophilic metabolites was filtered through syringe-mounted PTFE filter units (pore size 0.20 µm; Millipore, MA, USA) then 350 µL and 700µL were taken for GC/MS and LC/MS analysis respectively. Residual organic solvent was removed from the samples by centrifugal concentration, then the samples were freeze-dried overnight and stored at -80°C until analysis.

For GC/MS analysis, extracted metabolites were first derivatized by oximation (addition of 100 µL 20 mg/mL methoxyamine hydrochloride in pyridine, 1200 rpm for 90 min at 30°C) and silylation (addition of 50 µL, N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA), 1200 rpm for 30 min reaction at 37°C). The derivatized samples were analyzed on a GCMS-QP2010 Ultra (Shimadzu, Kyoto, Japan) with InertCap 5MS/NP column (0.25 mm ID x 30 m, df = 0.25 µm; GL Sciences, Tokyo, Japan). An alkane standard mix (C8 to C40) was injected prior to sample analysis for calculating retention indices. Peak detection, baseline correction and

retention time alignment were performed using MetAlign [1] followed by automated peak identification with AIoutput2 ver.1.29 [2]. For each sample, peak intensities were normalized to the internal standard (ribitol).

For LC/MS analysis, extracted metabolites were resuspended in 35µL ultrapure water and analyzed on a Shimadzu Nexera UHPLC system coupled to LCMS 8030 Plus (Kyoto, Japan) using a Mastro C18 reversed phase HPLC column (150 mm × 2.1 mm, particle size 3 µm; Shimadzu, Kyoto, Japan) operated in multiple reaction monitoring (MRM) mode. The mobile phases were 10 mM tributylamine and 15 mM acetic acid in water (A) and methanol (B). Peak identification and quantitation of peak areas were done using the LabSolutions software (Shimadzu, Kyoto, Japan).

Figures and tables

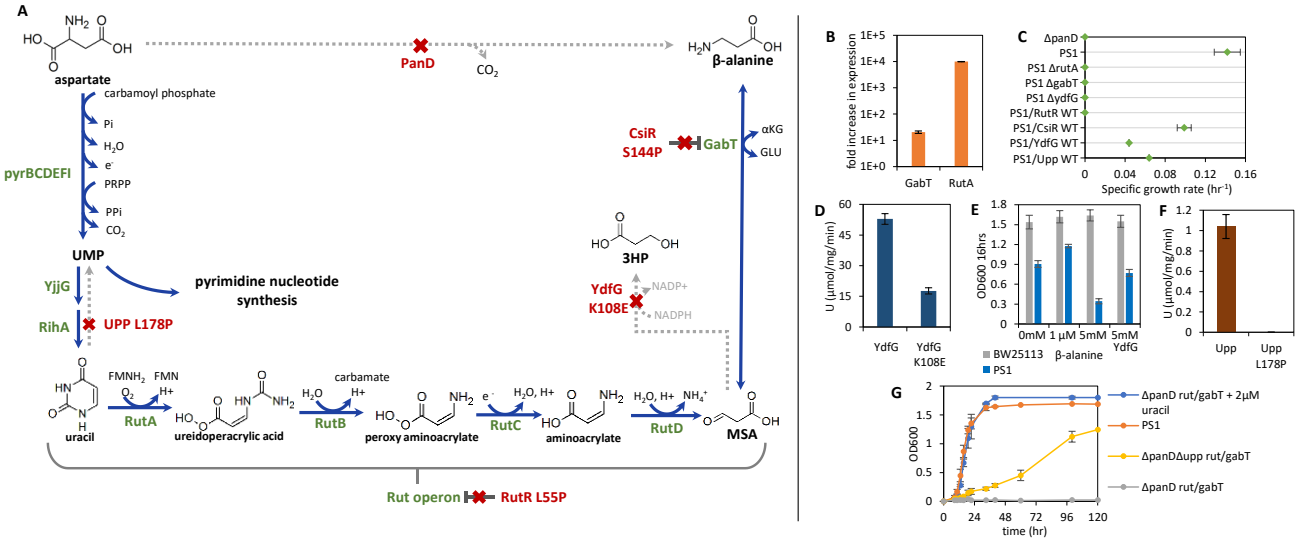


Figure 2-1: Metabolic repair through rewiring of existing metabolic components

A) A pathway emerged to repair β-alanine pathway damage that repurposed the uracil degradation pathway (Rut operon) and GabT. Mutations that contribute to this phenotype are noted in red. B) RT-qPCR measurements of GabT and RutA in PS1 compared to wildtype parent strain BW25113. Three independent sets of cultures were analyzed, each with three technical repeats. C) Growth phenotypes of PS1 with deletion of pathway enzymes or reversion of point mutations. D) *In vitro* specific activity of mutant YdfG compared to wildtype when expressed within cell lysate. Measurements were taken from three independent replicate cultures. E) Growth of PS1 in minimal media with varying amounts of β-alanine supplementation show that excess β-alanine is toxic to PS1. Overexpression of YdfG is able to relieve this toxicity. F) *In vitro* specific activity of mutant Upp expressed within cell lysate shows complete loss of activity. Measurements were taken from three independent replicate cultures. G) Growth in minimal media of pathway reconstruction. Overexpression of the Rut operon and GabT, noted rut/gabT, within a *ΔpanD* mutant can only rescue growth with uracil supplementation. Further deletion of

upp relinquishes the need for uracil supplementation. All growth curves were measured using three cultures originating from three separate colonies. Error bars represent standard deviation

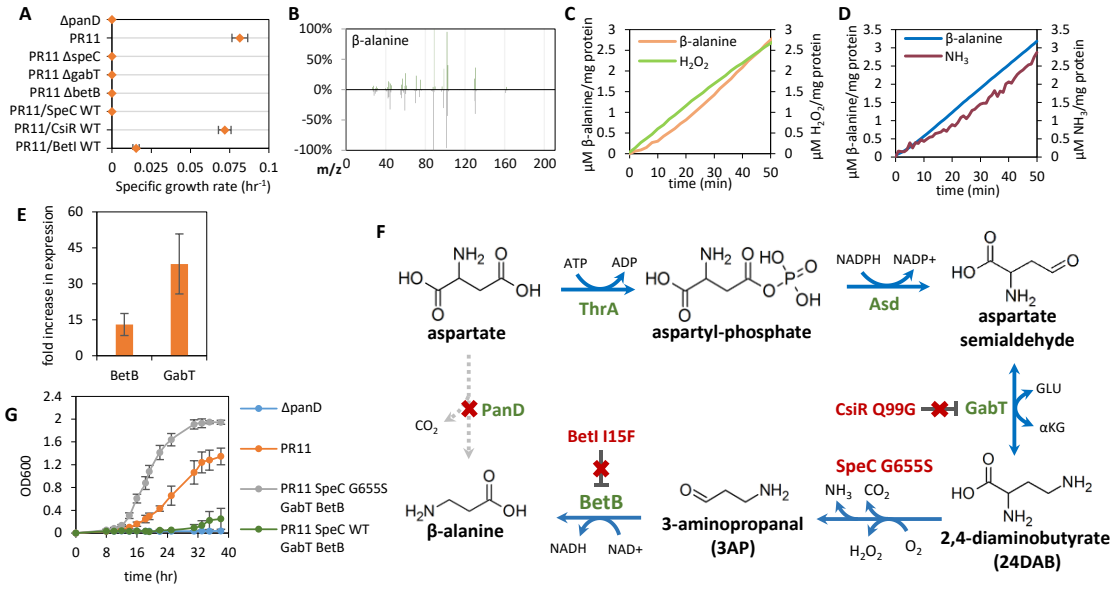
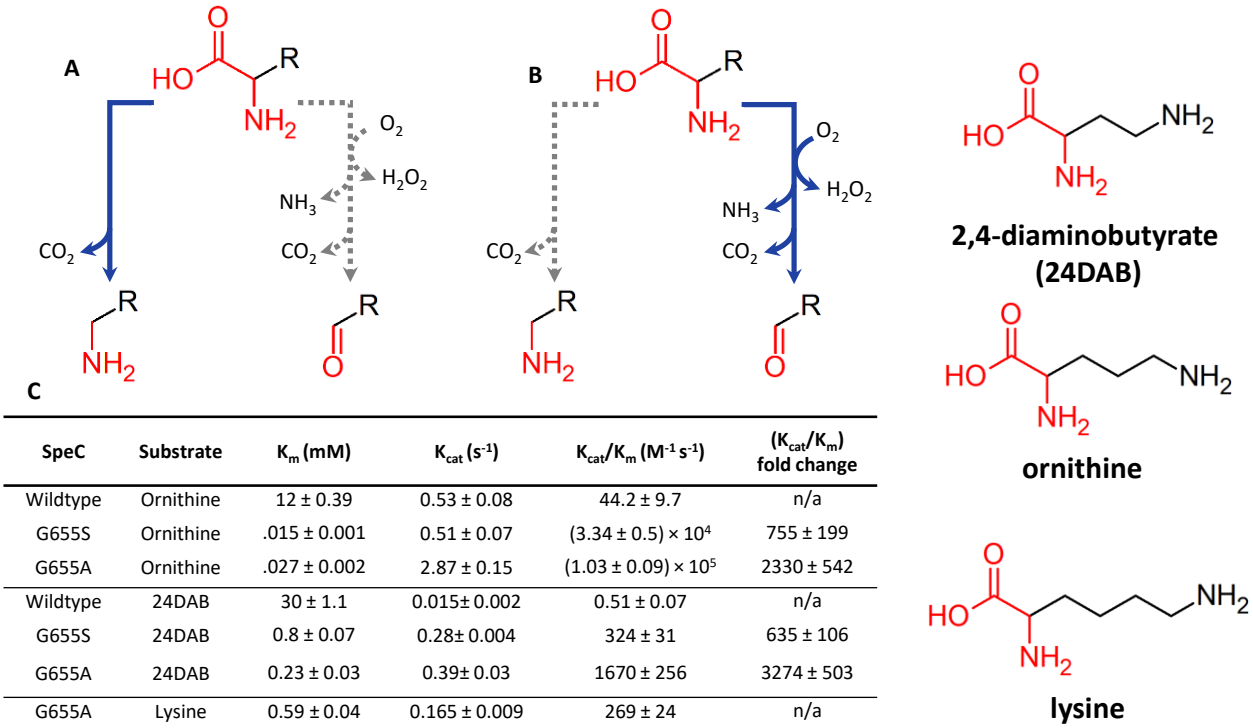


Figure 2-2: Metabolic repair through evolution of new enzyme function.

A) Growth in minimal media of evolved strain PR11 with deletion of pathway enzymes or point mutation reversions. B) Mass spectrum of β-alanine produced from 24DAB (2,4-diaminobutyrate) *in vitro* using a coupled assay with purified SpeC and BetB. C/D) β-alanine, NH_3, and H_2O_2 measured from a coupled SpeC and BetB assay show equimolar formation of all products. E) Relative gene expression of BetB and GabT in PR11 compared to wildtype parent BW25113. F) Illustration of the pathway that formed within PR11 to reroute damaged β-alanine biosynthesis. This pathway utilizes a gain of function mutation within ornithine decarboxylase (SpeC) that allows a bifunctional decarboxylation and deamination. Relevant mutations are noted in red. G) Reconstruction of this pathway to rescue growth of a *ΔpanD* strain in minimal media is accomplished with overexpression of SpeC G655S, GabT, and BetB.



SpeC	Substrate	K_m (mM)	K_{cat} (s^{-1})	K_{cat}/K_m ($M^{-1} s^{-1}$)	(K_{cat}/K_m) fold change
Wildtype	Ornithine	12 ± 0.39	0.53 ± 0.08	44.2 ± 9.7	n/a
G655S	Ornithine	.015 ± 0.001	0.51 ± 0.07	$(3.34 \pm 0.5) \times 10^4$	755 ± 199
G655A	Ornithine	.027 ± 0.002	2.87 ± 0.15	$(1.03 \pm 0.09) \times 10^5$	2330 ± 542
Wildtype	24DAB	30 ± 1.1	0.015± 0.002	0.51 ± 0.07	n/a
G655S	24DAB	0.8 ± 0.07	0.28± 0.004	324 ± 31	635 ± 106
G655A	24DAB	0.23 ± 0.03	0.39± 0.03	1670 ± 256	3274 ± 503
G655A	Lysine	0.59 ± 0.04	0.165 ± 0.009	269 ± 24	n/a

Figure 2-3: Altered activity of evolved ornithine decarboxylase

Two reaction types are observed from ornithine decarboxylase (SpeC) variants on three amino acid substrates. One is a single decarboxylation reaction, while the other is a bifunctional decarboxylation and deamination reaction. The relevant functional groups are highlighted in red. A) Wildtype SpeC can perform the single decarboxylation reaction on any of the three substrates, but only has minimal activity for the bifunctional reaction. B) SpeC G655S and G655A have high activity for the bifunctional reaction, but lost the ability to perform the single decarboxylation on both 24DAB and lysine. C) Enzyme kinetics of the bifunctional reaction using each SpeC variant.

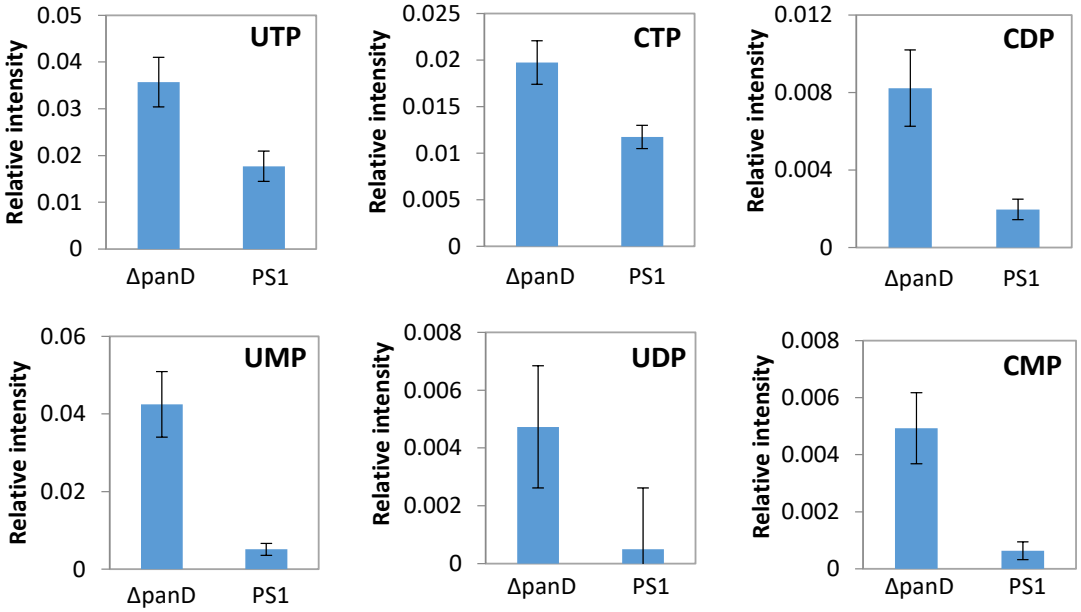


Figure 2-4: Intracellular metabolite measurements of pyrimidines

Intracellular metabolite measurements comparing PS1 to unevolved parent strain BW25113*ΔpanD*. Polar metabolites were extracted and analyzed by LC/MS using reversed phase HPLC coupled by negative mode electrospray ionization (ESI) to a triple-quadrupole mass spectrometer. Relative concentrations of pyrimidine nucleotides are significantly diminished within the evolved strain PS1.

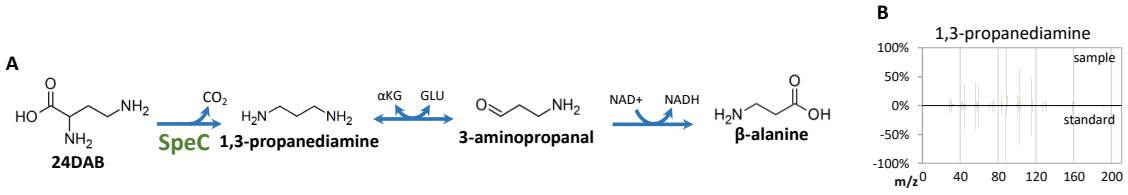


Figure 2-5: Hypothetical pathway using 13PDA as an intermediate

A) A hypothetical pathway that can convert 24DAB (2,4-diaminobutrate) into β-alanine. SpeC was assayed for 24DAB decarboxylase activity to produce 1,3-propanediamine. Wildtype SpeC can catalyze this reaction, but mutant SpeC variants cannot. B) Mass spectrum of 1,3-propanediamine formed as a product of this reaction using wildtype SpeC *in vitro*.

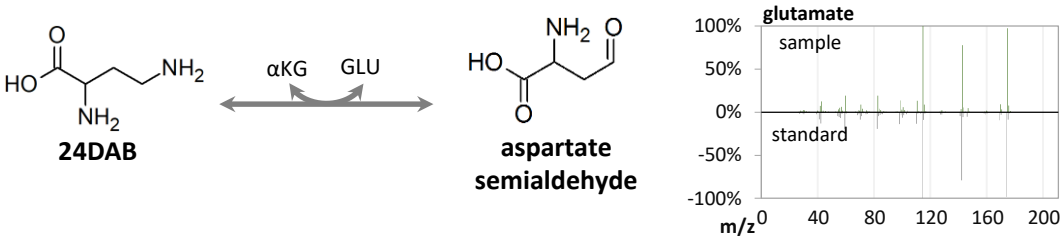


Figure 2-6: 24DAB transaminase reaction

Mass spectrum of glutamate formed as a product from 24DAB transaminase using purified GabT *in vitro.*

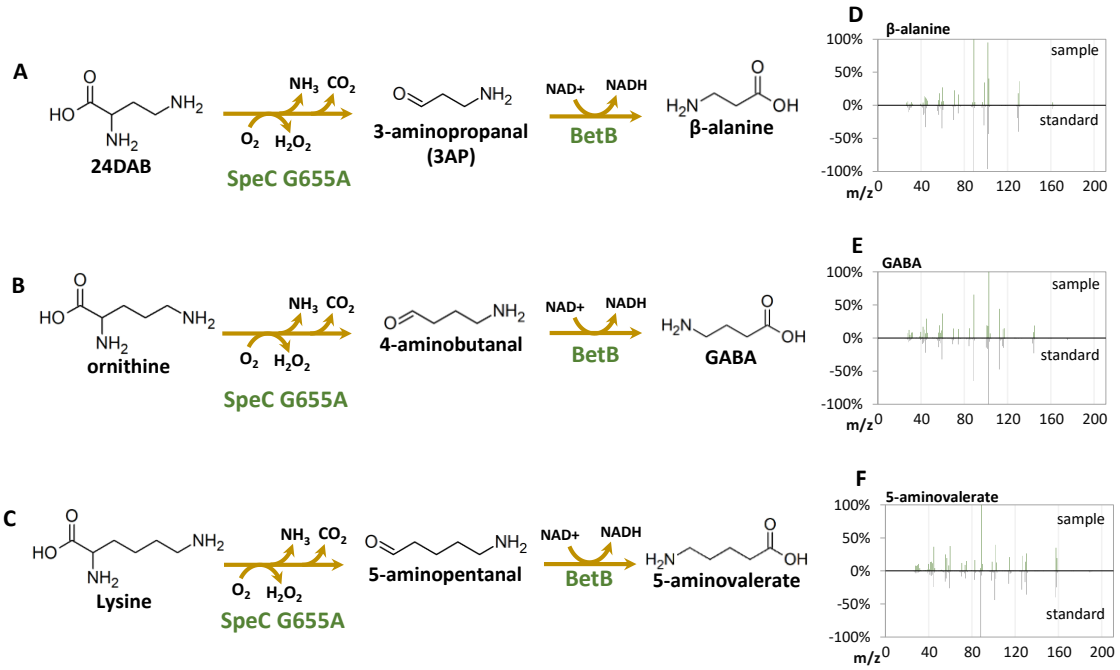


Figure 2-7: Promiscuous decarboxylation deamination reactions from SpeC G655A

In vitro assays using purified SpeC G655A were performed to detect a bifunctional decarboxylation and deamination reaction that is coupled with BetB to form the corresponding acid products. Substrates A) 2,4-diaminobutyrate (24DAB), B) Ornithine and C) Lysine were used. Mass spectra of corresponding product formation D) β-alanine E) 4-aminobutyrate (GABA) and F) 5-aminovalerate.

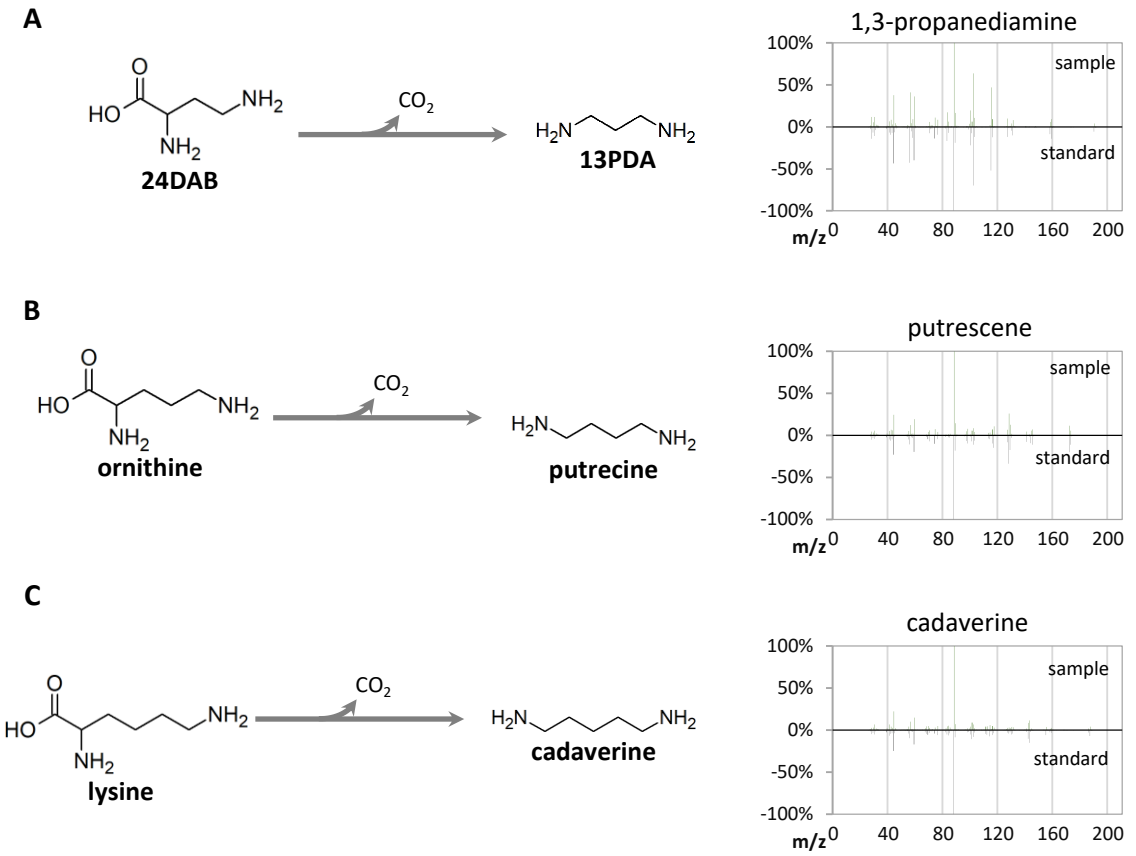


Figure 2-8: Promiscuous decarboxylation reactions from wildtype SpeC

Mass spectra of products formed from in vitro enzymatic reactions catalyzed by wildtype SpeC. A) 24DAB decarboxylase B) ornithine decarboxylase C) lysine decarboxylase.

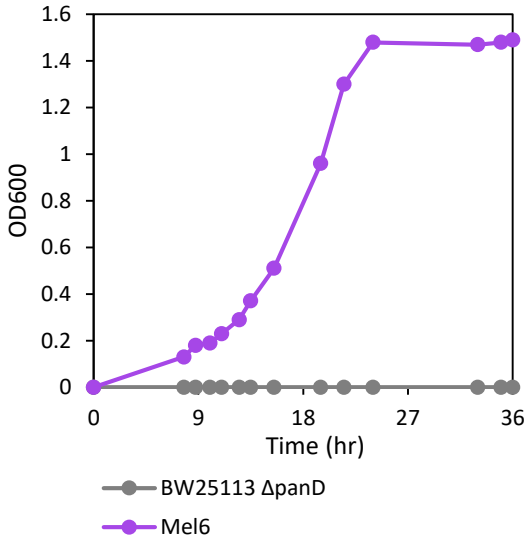


Figure 2-9: Growth curve of Mel6

Growth of evolved triple deletion mutant *ΔpanD ΔrutABC ΔspeC* strain (Mel6) in minimal media. Precultures were grown overnight in minimal media with 0.1μM β-alanine. This was inoculated into fresh minimal media 1:1000.

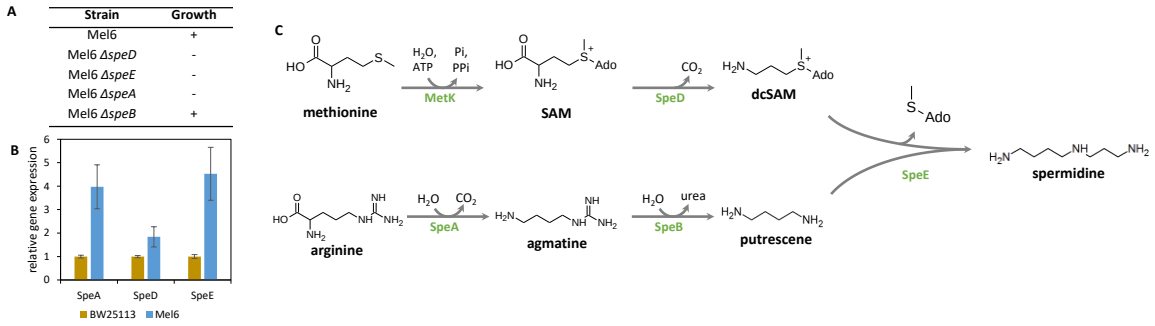


Figure 2-10: Polyamine biosynthesis and knockout data from Mel6

A) Growth in minimal media of Mel6 with various knockouts. With deletion of SpeD, SpeE, or SpeA no growth in minimal media is observed. B) RT-qPCR measurements of gene expression comparing Mel6 to wildtype parent BW25113. SpeA, SpeD, and SpeE have increased expression. C) Illustration of SpeD, SpeE, and SpeA catalyzed reactions used in polyamine biosynthesis.

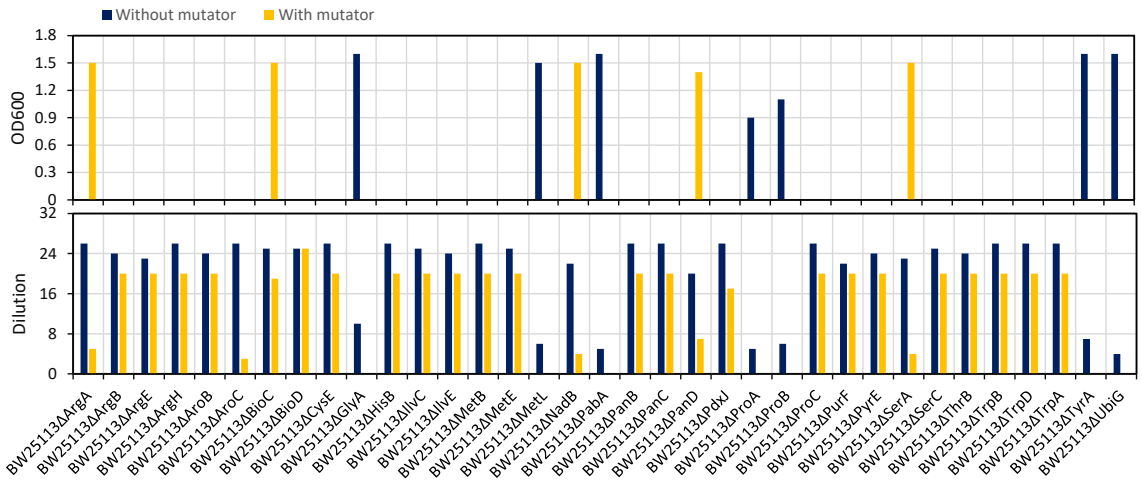


Figure 2-11: Large scale evolution using MutD5

Strains with various damaged pathways that confer metabolic auxotrophies were grown in glucose minimal media with sufficient nutritional supplementation to sustain growth between OD600 0.4-0.6. Cultures were diluted 1:1000 into fresh minimal media. Every dilution, the amount of nutritional supplementation was adjusted to restrict growth of the following passage to remain around OD600 0.4. In the event that the culture grew past OD600 of 1.0, it was diluted into minimal media with no supplementation. If growth was observed past OD600 of 0.6, it was streaked onto solid media. Resulting colonies were tested for a retained growth phenotype in a new liquid minimal media culture. Final genotypes were verified using genomic sequencing. Including *ΔpanD* strains, 35 total auxotrophs were tested. (bottom) The number of serial dilutions each set of strains was subjected to with or without a mutator plasmid. (top) Growth in minimal media with no supplementation after the noted number of serial dilutions. Strains that have repaired phenotypes exhibit growth.

Table 2-1: All mutation are listed in an auxiliary file

All genomic mutations identified using whole genome sequencing on strains that repaired damaged metabolic pathways.

File submitted as an auxiliary document.

BW25113*ΔargA*	Ornithine	LB	pALQ14	Yes
BW25113*ΔargB*	Arginine	LB	pALQ14	No
BW25113*ΔargE*	Ornithine	LB	pALQ14	No
BW25113*ΔargH*	Ornithine	LB	pALQ14	No
BW25113*ΔaroB*	Chorismate	LB	pALQ14	No
BW25113*ΔaroC*	Chorismate	LB	pALQ14	No
BW25113*ΔbioC*	Biotin	LB	pALQ14	Yes
BW25113*ΔbioD*	Biotin	LB	pALQ14	No
BW25113*ΔcysE*	Cysteine	LB	pALQ14	No
BW25113*ΔglyA*	Glycine	LB	n/a	Yes
BW25113*ΔhisB*	Histidine	LB	pALQ14	No
BW25113*ΔilvC*	Valine, Leucine, Pantothenate, Isoleucine	LB	pALQ14	No
BW25113*ΔilvE*	Valine	LB	pALQ14	No
BW25113*ΔmetB*	Methionine	LB	pALQ14	No
BW25113*ΔmetE*	Methionine	LB	pALQ14	No
BW25113*ΔmetL*	Homoserine	LB	n/a	Yes
BW25113*ΔnadB*	NAD	LB	pALQ14	Yes
BW25113*ΔpabA*	4-aminobenzoate	LB	n/a	Yes
BW25113*ΔpanB*	Pantothenate	LB	pALQ14	No
BW25113*ΔpanC*	Pantothenate	LB	pALQ14	No
BW25113*ΔpanD*	Pantothenate	β-alanine	pALQ14	No
BW25113*ΔpdxJ*	Pyridoxal 5'-phosphate	LB	pALQ14	No
BW25113*ΔproA*	Proline	LB	n/a	Yes
BW25113*ΔproB*	Proline	LB	n/a	Yes
BW25113*ΔproC*	Proline	LB	pALQ14	No
BW25113*ΔpurF*	Purine nucleotides	LB	pALQ14	No
BW25113*ΔpyrE*	UMP	LB	pALQ14	No
BW25113*ΔserA*	Serine	LB	pALQ14	No
BW25113*ΔserC*	Serine	LB	pALQ14	No

BW25113*ΔthrB*	Threonine	LB	pALQ14	No
BW25113*ΔtrpB*	Tryptophan	LB	pALQ14	No
BW25113*ΔtrpD*	Tryptophan	LB	pALQ14	No
BW25113*ΔtrpA*	Tryptophan	LB	pALQ14	No
BW25113*ΔtyrA*	Tyrosine, Phenylalanine	LB	pALQ14	Yes
BW25113*ΔubiG*	Ubiquinone	LB	n/a	Yes

Table 2-2: Auxotrophic strains subjected to evolution

Auxotrophic strains that were subjected to serial dilutions in minimal media with limited nutritional supplementation. Listed are the type of nutritional supplementation provided, as well as the mutator plasmid used to enhance the rate of mutagenesis. Wildtype strains are BW25113

Plasmid Construction of N-terminal His-Tag expression vectors

Plasmid	Template	Primers	
pALQ143 SpeC Insert	*E. coli* BW25113	F	CCAGGATCCGATGAAATCAATGAATATTGCCGCCA
		R	AGCGGTGGCATTACTTCAACACATAACCGTACAACC
pALQ142 Backbone	pCDFDuet	F	GTTGAAGTAATGCCACCGCTGAGCAATAAC
		R	TCATTGATTTCATCGGATCCTGGCTGTGGTGA
pALQ142 SpeC G655S	PR11	F	CCAGGATCCGATGAAATCAATGAATATTGCCGCCA
		R	AGCGGTGGCATTACTTCAACACATAACCGTACAACC
pALQ143 Backbone	pCDFDuet	F	GTTGAAGTAATGCCACCGCTGAGCAATAAC
		R	TCATTGATTTCATCGGATCCTGGCTGTGGTGA
pALQ144 GabT	*E. coli* BW25113	F	CCAGGATCCGATGAACAGCAATAAAGAGTTAATGCAGC
		R	AGCGGTGGCACTACTGCTTCGCCTCATCAAAAC
pALQ144 Backbone	pCDFDuet	F	GAAGCAGTAGTGCCACCGCTGAGCAATAA
		R	TTGCTGTTCATCGGATCCTGGCTGTGGTGA
pALQ153 BetB	*E. coli* BW25113	F	CCAGGATCCGATGTCCCGAATGGCAGAACA
		R	TCAGCGGTGGCATTAGAATATGGACTGGAATTTAGCCATC
pALQ153 Backbone	pCDFDuet	F	GTCCATATTCTAATGCCACCGCTGAGCAATAA
		R	TTCGGGACATCGGATCCTGGCTGTGGTG

Table 2-3: Primers used to construct plasmids that express N-terminal His-Tag proteins

Plasmid construction of mutator plasmids

All fragments were amplified for insertion between P_LLacO_1 directly upstream a 20bp RBS and T1 terminator.

P_LLacO_1	AATTGTGAGCGGATAACAATTGACATTGTGAGCGGATAACAAGATACTGAGC ACATCAGCAGGACGCACTGACC
RBS	GAATTCATTAAAGAGGAGAAAG
T1 Terminator	AGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTA TCTGTTGT

Insert	Template	Primers	
MutD5	XL1 Red (Agilent)	F	AGAGGAGAAAGATGAGCACTGCAATTACACGCC
		R	TTGATGCCTTTATGCTCGCCAGAGGCA

Table 2-4: Primers used in the construction of plasmids that harbor mutator genes

Strains

BW25113	rrnBT14 DlacZWJ16 hsdR514 DaraBADAH33 DrhaBADLD78
PS1	BW25113 *ΔpanD* Evolved to grow in minimal media
PR11	BW25113 *ΔpanDΔrutABC* Evolved to grow in minimal media
Mel6	BW25113 *ΔpanDΔrutABCΔspeC* Evolved to grow in minimal media

Plasmid	**Promotor**	**Genes**	**Origin/ Antibiotic**
pALQ9	P_LLacO_1	*mutD5* (XL1 Red, Agilent)	ColE1/Kan^r
pALQ68	P_LLacO_1	*gabT (E. coli)*	p15A/Spec^r
pALQ78	P_LLacO_1	*RutABCDEF (E. coli)*	ColE1/Amp^r
pALQ82	P_LLacO_1	*ydfG (E. coli)*	p15A/Amp^r
pALQ83	P_LLacO_1	*ydfG* D263V *(E. coli)*	p15A/Amp^r
pALQ84	P_LLacO_1	*upp (E. coli)*	colA/Kan^r
pALQ85	P_LLacO_1	*upp* L178P *(E. coli)*	colA/Kan^r
pALQ142	pT7	*speC* His *(E. coli)*	CDF/Spec^r
pALQ143	pT7	*speC* G655S His *(E. coli)*	CDF/Spec^r
pALQ144	pT7	*gabT* His (*E. coli*)	CDF/Spec^r
pALQ146	P_LLacO_1	*betB gabT (E. coli)*	ColA/Kan^r
pALQ153	pT7	*betB* His (*E. coli*)	CDF/Spec^r
pALQ173	pT7	*speC* G655A His *(E. coli)*	CDF/Spec^r

Table 2-5: Strain and plasmid designations used in this work

Genes	Direction	Primer Sequence
RutA	F	CGGCGCAGATCACTTTCATG
	R	GCACGGGGAAAAGCGATTTT
BetB	F	GGGCAATCTGGATCGGGTAG
	R	CGTTGCGTGAAACGTCCTTT
GabT	F	TGAAGACGGCGATCACAACA
	R	CGCAGCACGTTGTAATACGG
Frr	F	GTTGCAGTACGTAACGTGCG
	R	CAGCGCCGCTTCAATTTTCT
SpeA	F	AGTTGCAGTACCTGAGTCGC
	R	AGGGTTCGGGTAAATGGCAG
SpeE	F	TCCGGGTAGGTATGTACGCA
	R	AGTGAAGAACCGGTTGACCC
SpeD	F	GATAGGATCGGTGCAGTCGG
	R	GCGCTTTAAGCTGGTGATCG

Table 2-6: Primer sequences used for RT-qPCR.

All product sizes are between 70-150bp and have a measured efficiency between 90-150%.

Primers used for amplification of pKD13 kanamycin cassette

Gene	Target strain	Primers	
RutR	PS1	F	AGTGGACTAAACGGTCAAAACAGTTGCACATAAAA CATGCATTCCGGGGATCCGTCGACC
		R	CTTCTGCACTCTCATCGCGCTGTAGGCTGGAGCTGC TTCG
CsiR	PS1	F	AGGCGATGGCTGGCAATTAAATTCCGGGGATCCGT CGACC
		R	AATTTATCCGGGGCAAGTGTTGCGTATTCCGGAAG AGTAGTGTAGGCTGGAGCTGCTTCG
YdfG	PS1	F	CTGCCGGGTTATTGCTTGTCATTCCGGGGATCCGTC GACC
		R	TCTTGCTTGTGAGTGAGTTAACTGCATGAGTCTACC ACTTTGTAGGCTGGAGCTGCTTCG
Upp	PS1	F	CTCTGTATTATGTGTTATAGGCGCTTTACTCAAAAA AAAGATTCCGGGGATCCGTCGACC
		R	CAAAATCTTTGGTACGAAATAAAGAATAAAAATTG TAGGCTGGAGCTGCTTCG
SpeC	PR11	F	CAAACGGTCATAATAAGAAAATCAAACAAATTCCG GGGATCCGTCGACC
		R	GCACTAGCGTTGATAAAAGGGCCGATGACCACAAG AGTTCTGTAGGCTGGAGCTGCTTCG
CsiR	PR11	F	AGGCGATGGCTGGCAATTAAATTCCGGGGATCCGT CGACC
		R	AATTTATCCGGGGCAAGTGTTGCGTATTCCGGAAG AGTAGTGTAGGCTGGAGCTGCTTCG
BetI	PR11	F	AATAGTAACAATAACAGTGGGGATACTGATTCCGG GGATCCGTCGACC
		R	TGATTTTGTCCTTTTCCCTGCTGTGTGAAAGGTCTGT CATTGTAGGCTGGAGCTGCTTC

Table 2-7: Primers used to amplify a kanamycin cassette from pKD13 using PCR

The formed amplicon contains overlaps that are homologous to specific sites within the genome.

Gene	Target strain	Primers	
RutR	PS1	F	GCGCGATGAGAGTGCAGAAG
		R	CTGCCAGCATCTCCATACAGAA
CsiR	PS1	F	TTACGGGAAGCTCTTTCGCAA
		R	TTAATTGCCAGCCATCGCCT
YdfG	PS1	F	GTTAAAAGACGAACTGGGAGATAATCTG
		R	GACAAGCAATAACCCGGCAG
Upp	PS1	F	ATTTTTATTCTTTATTTCGTACCAAAGATTTTG
		R	AGCCGGTACCGTACTTCCA
SpeC	PR11	F	AGGAGATGGCGTGCATCGGG
		R	TTGTTTGATTTTCTTATTATGACCGTTTG
CsiR	PR11	F	TTACGGGAAGCTCTTTCGCAA
		R	TTAATTGCCAGCCATCGCCT
BetI	PR11	F	AGCAGACCATTTTTGTCCCTGA
		R	CAGTATCCCCACTGTTATTGTTACTATT

Table 2-8: Primers used to amplify genomic regions used for point mutations

Primers used to amplify 400-1000bp regions of BW25113 genomic DNA. These were then used as templates to SOE with linear pKD13 amplicons to create linear fragments used for point mutation reversions.

References

1. Reed, J. L., T. D. Vo, C. H. Schilling, B. O. Palsson, An expanded genome-scale model of Escherichia coli K-12 (i JR904 GSM / GPR). *genome Biol.* **4**, 1–12 (2003).

2. Guzmán, G. I., J. Utrilla, S. Nurk, E. Brunk, J. M. Monk, A. Ebrahim, Model-driven discovery of underground metabolic functions in Escherichia coli. *Proc. Natl. Acad. Sci.* **112**, 929–934 (2014).

3. Tawfik, O. K. and D. S., Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annu. Rev. Biochem.* **79**, 471–505 (2010).

4. Kim, J., J. P. Kershner, Y. Novikov, R. K. Shoemaker, S. D. Copley, Three serendipitous pathways in E. coli can bypass a block in pyridoxal-5'-phosphate synthesis. *Mol. Syst. Biol.* **6**, 436 (2010).

5. Webb, M. E., G. Smith, C. Abell, Biosynthesis of pantothenate. *Nat. Prod. Rep.* **21**, 695–721 (2004).

6. Boldyrev, A. A., Carnosine: new concept for the function of an old molecule. *Biochem. Biokhimiĭa*. **77**, 313–26 (2012).

7. Joanne, W. M., and G. M. Brown, Purification and Properties of L-Aspartate-α-decarboxylase , That Catalyzes the Formation of β-Alanine in Escherichia. *J. Biol. Chem.* **254**, 8074–8082 (1979).

8. Nozaki, S., M. E. Webb, H. Niki, An activator for pyruvoyl-dependent l-aspartate α-decarboxylase is conserved in a small group of the γ-proteobacteria including Escherichia

coli. *Microbiologyopen*. **1**, 298–310 (2012).

9. Rathinasabapathi, B., Propionate, a source of β-alanine, is an inhibitor of β-alanine methylation in Limonium latifolium, Plumbaginaceae. *J. Plant Physiol.* **159**, 671–674 (2002).

10. Fritzson, P., The catabolism of C14-labeled uracil, dihydrouracil, and β-ureidopropionic acid in rat liver slices. *J. Biol. Chem.* **226**, 223–228 (1956).

11. White, W. H., P. L. Gunyuzlu, J. H. Toyn, Saccharomyces cerevisiae Is Capable of de Novo Pantothenic Acid Biosynthesis Involving a Novel Pathway of beta-Alanine Production from Spermine. *J. Biol. Chem.* **276**, 10794–10800 (2001).

12. Maruyama, M., T. Horiuchi, H. Maki, M. Sekiguchi, A dominant (mutD5) and a recessive (dnaQ49) mutator of Escherichia coli. *J. Mol. Biol.* **167**, 757–771 (1983).

13. Kim, K. S., J. G. Pelton, W. B. Inwood, U. Andersen, S. Kustu, D. E. Wemmer, The Rut pathway for pyrimidine degradation: Novel chemistry and toxicity problems. *J. Bacteriol.* **192**, 4089–4102 (2010).

14. Borodina, I., K. R. Kildegaard, N. B. Jensen, T. H. Blicher, J. Maury, S. Sherstyk, K. Schneider, P. Lamosa, M. J. Herrgård, I. Rosenstand, F. Öberg, J. Forster, J. Nielsen, Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. *Metab. Eng.* **27**, 57–64 (2015).

15. Andersen, P. S., J. M. Smith, B. Mygind, Characterization of the upp gene encoding uracil phosphoribosyltransferase of Eschevichia coli K12. *FEBS J.* **204**, 51–56 (1992).

16. Bertoldi, M., V. Carbone, C. B. Voltattorni, S. Neurologiche, C. Biologica, C. Nazionale, C. Internazionale, V. Pansini, Ornithine and glutamate decarboxylases catalyse an oxidative deamination of their α-methyl substrates. *Biochem. J.* **512**, 509–512 (1999).

17. Incharoensakdi, A., N. Matsuda, T. Hibino, Y. Meng, H. Ishikawa, A. Hara, T. Funaguma, T. Takabe, T. Takabe, Overproduction of spinach betaine aldehyde dehydrogenase in Escherichia coli. *Eur. J. Biochem.* **7023**, 7015–7023 (2000).

18. Blank, D., L. Wolf, M. Ackermann, O. K. Silander, The predictability of molecular evolution during functional innovation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 3044–9 (2014).

19. Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. a Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Mol. Syst. Biol.*, in press, doi:10.1038/msb4100050.

20. Espah Borujeni, A., A. S. Channarasappa, H. M. Salis, Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* **42**, 2646–2659 (2014).

21. Salis, H. M., E. A. Mirsky, C. A. Voigt, Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–50 (2009).

22. Datsenko, K. a, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–6645 (2000).

23. Degnen, G. E., E. C. Cox, Conditional mutator gene in Escherichia coli: isolation, mapping, and effector studies. *J. Bacteriol.* **117**, 477–487 (1974).

24. Livingston, D., Deoxyribonucleaic Acid Polymerase III of Escherichia coli. *J. Biol. Chem.* **250**, 489–497 (1975).

25. Jensen, K. F., B. Mygind, Different oligomeric states are involved in the allosteric behavior of uracil phosphoribosyltransferase from Escherichia coli. *Eur. J. Biochem.* **645**, 637–645 (1996).

26. Wernick, D. G., S. P. Pontrelli, A. W. Pollock, J. C. Liao, Sustainable biorefining in wastewater by engineered extreme alkaliphile Bacillus marmarensis. *Sci. Rep.* **6**, 20224 (2016).

27. Canellakis, E. S., a a Paterakis, S. C. Huang, C. a Panagiotidis, D. a Kyriakidis, Identification, cloning, and nucleotide sequencing of the ornithine decarboxylase antizyme gene of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7129–7133 (1993).

28. Choi, K. Y., D. G. Wernick, C. a. Tat, J. C. Liao, Consolidated conversion of protein waste into biofuels and ammonia using Bacillus subtilis. *Metab. Eng.* **23**, 53–61 (2014).

29. Kaminaga, Y., J. Schnepp, G. Peel, C. M. Kish, G. Ben-Nissan, D. Weiss, I. Orlova, O. Lavie, D. Rhodes, K. Wood, D. M. Porterfield, A. J. L. Cooper, J. V. Schloss, E. Pichersky, A. Vainstein, N. Dudareva, Plant phenylacetaldehyde synthase is a bifunctional homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation. *J. Biol. Chem.* **281**, 23357–23366 (2006).

30. Applebaum, D. M., J. C. Dunlap, D. R. Morris, Comparison of the Biosynthetic and Biodegradative Ornithine Decarboxylases of Escherichia coli. *Biochemistry*. **16**, 1580–1584 (1977).

31. Smart, K. F., R. B. M. Aggio, J. R. Van Houtte, S. G. Villas-bôas, Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography – mass spectrometry. *Nat. Protoc.* **5** (2010), doi:10.1038/nprot.2010.108.

32. Li, H., R. Durbin, Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. **26**, 589–595 (2010).

33. McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Proc. Int. Conf. Intellect. Capital, Knowl. Manag. Organ. Learn.* **20**, 254–260 (2009).

34. Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin).* **6**, 80–92 (2012).

35. Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, S. G. Rozen, Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, 1–12 (2012).

36. Liu, M., T. Durfee, J. E. Cabrera, K. Zhao, D. J. Jin, F. R. Blattner, Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli. *J. Biol. Chem.* **280**, 15921–15927 (2005).

Chapter 3 : Directed strain evolution restructures metabolism for 1-butanol production in minimal media

This section was adapted from Pontrelli S., Fricke R.B.C, Sakurai S., Laviña W. A., Putri S.P., Fitz-Gibbon S., Chung M., Chiu T.Y., Pellegrini M., Fukusaki E., Liao J.C. Microbial evolution rebalanced metabolic infrastructure for butanol production in minimal media (manuscript in preparation)

Abstract

Engineering a microbial strain for production sometimes entails metabolic modifications that impair essential physiological processes for growth or production. Restoring these functions may require amending a variety of non-obvious physiological networks, and thus, rational design strategies may not be practical. Here we demonstrate that growth and production may be restored by evolution that repairs impaired metabolic function. Previously, high titers of butanol production were achieved by *Escherichia coli* using a growth-coupled, modified *Clostridial* CoA-dependent pathway after all native fermentative pathways were deleted. However, production was only observed in rich media. Native metabolic function of the host was unable to support growth and production in minimal media. We use directed strain evolution to repair this phenotype and observed improved growth, titers and butanol yields. We further used genomics, metabolomics and proteomics to identify several underlying mutations and metabolic perturbations that allow metabolism to repair: mutations in the ArcAB two-component system and integration host factor (IHF) tune expression of enzymes within central carbon metabolism and alter energy metabolism to result in increased butanol yields. A mutation in *pcnB* resulted in decreased relative plasmid copy numbers and pathway enzymes to balance resource utilization. Increases in glycolysis and biosynthetic enzymes, as well as decreases in degradation pathways, were also observed. These results demonstrate that metabolic impairment caused by strain engineering can be repaired using directed strain evolution, and further, they illustrate the diverse strategies that may underlie such repair.

Introduction

Engineering microbes for production sometimes require extensive modifications, which may perturb essential physiological processes and impair growth. Coupling between growth and production has proven to be an effective way to address this problem. One example of this is a previously constructed butanol producing strain of *E. coli*. Within this system, all native fermentation pathways were deleted: alcohol dehydrogenase (*adhE*), lactate dehydrogenase (*ldhA*), and fumarate reductase (*frdBC*). As a result, no electron sink is available under anaerobic conditions and growth is not possible. A modified *Clostridial* CoA-dependent pathway was introduced in which NAD+ can be regenerated, 1-butanol is established as the sole electron sink, and growth and 1-butanol production are coupled under anaerobic conditions (Figure 3-1a). Production using this system as a starting platform has realized titers of 18.3 g/l in anaerobic batch fermentation(*1*). However, growth and production in minimal media are severely restricted.

In minimal media, a major shift in cell resources occurs as all metabolites and cofactors must be synthesized *de novo.* As such, the physiological modifications used to enhance butanol production likely impair metabolic function in conjunction with the redistribution of cell resources. The underlying physiological causes of a severely weakened phenotype are unknown. In this case, rational engineering strategies aimed at restoring impaired metabolic function is not practical. Nutritional supplements or other operational strategies can be employed to augment the production phenotype in lean media(*2*)(*3*)(*4*)(*5*). However, these strategies are costly and do not shed light on the underlying causes of the problem.

Previously, directed strain evolution has been employed to adapt cells to altered environmental conditions. With a continuous selection pressure, directed strain evolution offers an efficient

strategy to address system-wide, non-obvious physiological factors that can lead to metabolic innovation by the cells. Some examples include improving product or substrate toxicity(*6*)(*7*), thermotolerance(*8*), elongation of growth phase(*9*), growth using alternative substrates(*10*)(*11*)(*12*) (reduce to one cite), enhancing function of native pathways(*13*), and replacing native glycolysis with a synthetic non-oxidative glycolysis pathway(*14*). Directed strain evolution has also been used to repair damaged metabolic function that confers metabolite auxotrophy(*15–17*).

Here, we demonstrate the use of directed strain evolution to restore metabolic impairment of the 1-butanol producing strain (JCL166) in minimal media. Our strategy employs a mutator gene *MutD5*(*18*), which enhances the rate of accumulation of genomic mutations over time. While *MutD5* results in a large number of mutations, it also allows innovative adaptations to occur within a shorter timeframe. We utilize genomic sequencing, proteomics and metabolomics to identify causal mutations and metabolic perturbations. As a result, we have identified several mutations that aid in balancing electron supply for altered product formation, decreased expression of pathway enzymes, and increased expression of biosynthetic processes.

Results

Evolving growth in minimal media

All strains used in this study for production of butanol are deficient in native *E. coli* fermentation pathways. JCL166 (*ΔadhEΔldhAΔfrdBC*) with the expression of the modified *Clostridial* butanol pathway (JCL166B)(Figure 3-1a), has previously been demonstrated to produce 6.5g/l of 1-butanol after 72 hours(*19*). The high production titers were only observed in rich media, and severely reduced growth and production is observed in minimal media. Although other strain modifications previously improved titers further, certain modifications interfered with the efficacy of our evolution strategy, as discussed below. Therefore, JCL166B was chosen as a starting point for our efforts to enable production of butanol in minimal media. Owing to nutritional limitations in minimal media, we initially hypothesized that lack of growth was caused by insufficient expression of pathway enzymes. To test this, specific activity of pathway enzymes in cell lysate was compared between JCL166B in rich and minimal media (Figure 3-1b). These results demonstrate that enzyme activity is similar regardless of media composition, and suggest that other factors must be responsible for lack of growth and production.

We therefore attempted to evolve JCL166B using directed strain evolution to gradually reduce the amount of nutritional supplementation required for growth. Our evolution strategy involves the use of mutator gene *mutD5* (Figure 3-1c). MutD5, a dominant negative mutant of DnaQ(*20*), encodes for DNA pol III subunit ε which catalyzes 3’ to 5’ exonuclease activity in proofreading during DNA replication(*21*). With overexpression of MutD5, mutations accumulate with every division cycle and are increased. This effectively increased the rate of evolution. Previously reported modifications to improve butanol titers involved deletion of *pta*, encoding for phosphate

acetyltransferase. However, it was previously demonstrated that deletion of *pta* improves DNA replication fidelity and directly negates the increased mutagenesis rate conferred by overexpression of MutD5(*22*). Further, overexpression of formate dehydrogenase (FDH) from *Candida boidinii* was employed to oxidize formate produced in the anaerobic conversion of pyruvate into acetyl-CoA. This increased NADH supply for enhancing butanol formation. Because overexpression of the previously published plasmid containing FDH would introduce plasmid stability issues alongside overexpression of the additional MutD5 plasmid, and because the *pta* deletion would interfere with the efficacy of MutD5, we chose JCL166B as the base strain and these further modifications were employed following evolution.

Strains of JCL166B were constructed to express the butanol pathway (Table 3-1), along with mutator gene *mutD5*. These strains were inoculated into rich media in anaerobic conditions. Cultures were then subjected to successive serial dilutions with decreasing nutritional supplementation that limits growth to OD_{600} 0.4-0.6 (Figure 3-1c). Growth without nutritional supplementation, with only minimal media remaining, was observed after 5 successive serial dilutions (Figure 3-2a). However, the growth phenotype was weak. Therefore, these strains were further evolved for an additional 25 serial dilutions.

After 30 serial dilutions, a significant improvement of both growth and production was observed (Figure 3-2a). As expected, production of 1-butanol increased over the course of evolution in correlation with growth, and reached a plateau of 600mg/l in 24 hours by the 16^{th} dilution. To verify that growth remains dependent on the butanol pathway, all plasmids were cured from the evolved strain to reveal a complete abolishment of the phenotype (Figure 3-2b).

Mutator plasmid pALQ32 was cured from a culture subjected to 16 serial dilutions, and clonal strains were individually tested for their evolved phenotype (Figure 3-2c). Growth and production

varied between these colonies, however, butanol production remained correlated with growth. Further, byproduct formation of acetate and formate also correlated with growth. On the contrary, little or no ethanol, lactate, or succinate were produced from any culture to verify that the native fermentation pathways were absent and do not contribute to the evolved phenotype.

Of the 24 colonies that have variable growth phenotypes, butanol yields are also positively correlated with growth (Figure 3-2f). For the other byproducts, yields of formate and acetate remained relatively constant as growth improved, while the yield of butyrate decreased. While *E. coli* has no annotated butyryl-CoA transferase that can produce butyrate, a number of CoA transferases that catalyze this reaction on different substrates likely have unknown substrate promiscuity. Such enzymes exist within other organisms. The formation of butyrate and butanol each have different energy requirements. Commencing with the condensation of 2 acetyl-CoA into acetoacetyl-CoA, 4 additional NADH are required for production of 1 butanol molecule, while only 2 NADH are required for production of 1 butyrate molecule. This shift in byproduct formation suggests the presence of mutations that increase the supply of electrons in the catabolism of glucose into acetyl-CoA.

The best producing colony, BP1, was isolated for further study and improvement. Previous studies have employed two rational modifications to further enhance butanol production using this pathway(*19*). Formate is produced as a byproduct of butanol in anaerobic conditions from native pyruvate formate-lyase, PflB. Formate dehydrogenase (FDH) oxidizes this excess formate into carbon dioxide while producing a molecule of NADH. We overexpressed NAD+ dependent FDH from *Candida boidinii* to enhance the oxidation of formate(*23*). Further, *pta* was deleted, encoding for phosphate acetyltransferase, which catalyzes the first reaction in the degradation of acetyl-CoA into acetate. This effectively enhances acetyl-CoA pool sizes for further downstream

conversion into 1-butanol. Together, these modifications raised titers from 0.960mg/l to 2g/l in 96 hours (Figure 3-2d). Curiously, comparing production of butanol in rich media between JCL166B and BP1, titers were markedly decreased from BP1 (Figure 3-2e). This is consistent with other evolutionary studies, which identified mutations that offer advantages for growth in minimal media that decrease growth in rich media(*24*).

Genomic sequencing reveals mutations in anaerobic redox metabolism

To investigate how energy metabolism was being altered by evolution, as well as other mutations that contribute to the augmented phenotype, BP1 was sequenced to reveal 113 mutations that include 5 insertions and 2 deletions (Table 3-2Table 3-3). 89 coding regions were mutated in total that include 56 non-synonymous mutations. Two stop codons were found in uncharacterized genes *ycaK* and *yfdT*. Of four frameshift mutations, one was found on uncharacterized protein *yfjH*. Two of the observed frameshift mutations were observed on enzymes involved with anaerobic redox metabolism: *fdnH*, encoding for a subunit of formate dehydrogenase N, and *hyaF*, encoding for a subunit of hydrogenase-1.

Anaerobic respiration in *E. coli* can proceed in the presence of several external electron acceptors: nitrate, dimethyl sulfoxide (DMSO), trimethylamine-*N*-oxide (TMAO), and fumarate. All respiratory chains in *E. coli* are comprised of dehydrogenases and are linked to terminal reductases or oxidases by a quinone pool. Because a variety of each of these components exists in *E. coli*, a large diversity in the composition of respiratory chains is observed, which are responsive to different environmental factors. Formate dehydrogenase N and hydrogenase-1 respectively oxidize formate or H_2 into CO_2 or H^+, while donating an electron to a membrane

bound quinone. Anaerobic fermentation in wildtype strains can use fumarate as a terminal acceptor with fumarate reductase. However, this enzyme is deleted in JCL166B. Here, no other electron acceptor is available. Therefore, the presence of deleterious mutations in formate dehydrogenase N and hydrogenase-1 likely confer beneficial effects by preventing excess electrons from being dumped into the respiratory chain.

The fourth frameshift mutation was found in *barA*, encoding for a sensory histidine kinase of the BarA/UvrY two-component signal transduction system. This two component system plays a regulatory role in concert with the carbon storage regulator (csr), composed of the two non-coding RNAs *csrB* and *csrC*(*25*). These non-coding RNAs in turn react with multiple copies of the RNA binding protein CsrA(*26*), which has been shown to regulate either directly, or indirectly, a large number of enzymes involved with carbon metabolism, biofilm formation, motility, biofilm formation, and protein uptake(*27*). We speculate that this mutation relates to the adapted phenotype in BP1 as it has been shown that formate acts as a direct activator of the BarA/UvrY transduction system(*27*). Native anaerobic fermentation is here disrupted both through deletions in JCL166B and frameshift mutations in *fdhH* and *hyaF*. These alterations likely cause altered formate concentrations and may confer unintended damage to the BarA/UvrY global regulatory system, which may be mitigated by the frameshift mutation in *barA*.

Optimized carbon and energy metabolism through mutations in redox regulators

Two mutated genes in BP1, *ifhB* (T98C, M33T) and *arcB* (C315A, N105K), are both involved in regulation of a large number of genes used for growth in anaerobic conditions(*28*)(*29*)(*30*). We

speculated that these genes play a role in adapting anaerobic redox metabolism of BP1 for growth in minimal media.

Global regulation of enzymes involved in anaerobic redox metabolism is largely controlled by the ArcAB two-component system, which senses oxygen limitations and elicits a system wide response. ArcB is a constitutively expressed, membrane bound sensory histidine kinase that phosphorylates DNA-binding transcriptional regulator ArcA during anaerobic or microaerobic conditions(*31*). This two component system is known to be responsible for regulation of a wide range of enzymes and metabolic processes that include the TCA cycle, glyoxylate shunt, and fatty acid degradation(*32*). It has been shown that the genetic disruption of *arcB* causes systematic deviations in response to the availability of oxygen, and it is believed that responses generated by the ArcAB system are initiated solely by the ArcB sensor protein(*33*). The point mutation observed on ArcB in BP1 was reverted to the wildtype sequence (BP1 WT*arcB*) and a reduced growth rate was observed, demonstrating the importance of this mutation (Figure 3-3a). To determine whether this mutation conferred a total loss of function of ArcB, we further deleted it from BP1 (BP1 *ΔarcB*). The resulting growth phenotype of this strain is decreased in a similar manner to growth with the reverted wildtype ArcB sequence (Figure 3-3a). These results indicate that the mutation on ArcB confers a beneficial physiological effect without destroying enzyme function.

The other mutated gene, *ihfB*, encodes for a polypeptide that forms the integration host factor (IHF) of *E. coli* along with polypeptide IhfA. IHF binds to DNA at site-specific locations to induce sharp bends into the DNA backbone, and plays a role in DNA replication, recombination and transcription regulation(*34*)(*30*). Many genes have been identified that have altered regulation when IHF is deleted in *E. coli*(*30*). In anaerobic conditions, IHF has been shown to

alter regulation of many genes that contribute to energy metabolism. Notably, the expression of *pflB*, encoding for pyruvate formate-lyase, has been shown to be dependent on IHF in anaerobic conditions(*29*). To determine whether the mutation in *ihfB* confers a beneficial growth phenotype to BP1, it was reverted to the wildtype sequence in BP1 (BP1 WT*ihfB*). Further, *ihfB* was deleted from BP1(BP1 *ΔihfB*). The resulting growth phenotypes are significantly decreased compared to BP1 (Figure 3-3c), demonstrating that the mutation contributes to the evolved phenotype without destroying the function of IhfB.

To test how these mutations may be affecting butanol production in BP1, we set to determine the effect of these mutations on the expression of genes in central carbon metabolism using RT-qPCR. Expression of operons in the TCA cycle and glycolysis were compared between JCL166B and BP1 after growth in minimal media. As shown in Figure 3-3a, 12 of 28, of these operons exhibited significant expression differences after evolution (log2 ratio ≥ 1 or ≤ -1). To determine whether genes with significant expression changes are influenced by the *arcB* (C315A) mutation, gene expression was compared between BP1, BP1 *ΔarcB*, and BP1 WT*arcB* (Figure 3-3c). Similarly, to determine if these expression changes are caused by the *ifhB* (T98C) mutation, gene expression was compared between BP1, BP1 *ΔihfB*, and BP1 WT*ihfB* (Figure 3-3d). As observed in Figure 3-3cd, all significantly perturbed genes comparing BP1 to BP1 WT*arcB* or BP1 WT*ihfB* have increased expression, suggesting that these mutations aid in increasing activity of central carbon metabolism. To verify this increase in activity, we measured specific activity of glycolytic enzymes within cell lysate in both JCL166B and BP1 and determined that reactions within glycolysis are increased (Figure 3-7cd).

Interestingly, expression of the *aceEF* operon, encoding for two subunits of pyruvate dehydrogenase (PDH) was dependent on the presence of mutated sequences of both *ihfB* and

arcB (Figure 3-3cd). Gene expression measurements of *aceEF* were repeated in BP1 and all *arcB* and *ihfB* derivatives for comparison to JCL166B (Figure 3-4ab) to confirm that both the *arcB* and *ihfB* mutations are required to alter expression of *aceE* from wildtype levels. The products of AceEF encode for two subunits of PDH in addition to the third subunit, lipoamide dehydrogenase, Lpd. Under anaerobic conditions, the primary conversion of pyruvate to acetyl CoA is catalyzed by PflB, pyruvate formate-lyase. PDH produces acetyl-CoA and NADH as products, while PflB produces acetyl-CoA and formate. Usage of PflB under anaerobic conditions is preferred by the cell because formate can act as an electron acceptor, negating the need to produce excess fermentation products for use as an electron sink. Besides PDH and PflB, two other genes are capable of converting pyruvate into acetate or acetyl-CoA: pyruvate oxidase (PoxB) and 2-ketobutyrate formate-lyase (TdcE). These genes were all measured comparing JCL166B with BP1 and only *aceE* has increased expression (Figure 3-3b), suggesting that it may be conferring beneficial physiological effects that extend beyond increased conversion of pyruvate into acetyl-CoA.

The use of PDH serves advantageous for improving butanol production because additional NADH can be produced to balance the stoichiometric conversion of glucose into butanol. In the conversion of pyruvate into acetyl-CoA, if only PflB is used and only formate is produced, then there will be an electron deficiency for butanol production and theoretically only butyrate will be formed:

1 glucose + 2 CoA ➔ 2 acetyl-CoA + 2 formate + 2 NADH

2 acetyl-CoA + 2 NADH ➔ 1 buryrate + 2 CoA

The net reaction becomes: 1 glucose + ➔ 2 formate + butyrate

However, if these additional formate are oxidized to produce NADH, then butanol formation should increase:

$$1 \text{ glucose} + 2 \text{ CoA} \rightarrow 2 \text{ acetyl-CoA} + 4 \text{ NADH} + 2 \text{ CO}_2$$

$$2\text{-acetyl-CoA} + 4 \text{ NADH} \rightarrow 1 \text{ butanol} + 2 \text{ CoA}$$

$$\text{The net reaction becomes: } 1 \text{ glucose} \rightarrow 1 \text{ butanol} + 2 \text{ CO}_2$$

Any increased NADH supplied by PDH upregulation can therefore aid in conversion of butyrate into butanol. It is important to note that PDH has been shown to have decreased activity under anaerobic conditions as it is inhibited by high NADH/NAD+ ratios. Further, formate is still being produced as a major biproduct. This suggests that the activity of PDH likely only supplements pyruvate formate-lyase in conversion of pyruvate to acetyl-CoA.

To validate that mutations within *arcB* and *ihfB* contribute to the altered byproduct formation observed within BP1, product formation was compared between JCL166B, BP1, BP1 *ΔarcB* and BP1 *ΔihfB* (Figure 3-4cde). These results confirm a 2-fold increase of butanol yield in BP1 compared to JCL166B, with a 1.7-fold decrease in butyrate yield. Within BP1 WT*arcB* and BP1 WT*ihfB*, a decrease in butanol yield is observed, suggesting a lack of sufficient reducing power. Increased byproduct formation was observed in acetate rather than butyrate, which requires the same amount of reducing equivalents.

JCL166B, without expression of FDH, was here and previously demonstrated to produce butanol in rich media(*19*)(Figure 3-2e). These results demonstrate that FDH is neither essential for growth nor butanol production. While our results suggest that further oxidation of formate into NADH aids in a robust growth phenotype, additional overexpression of FDH within JCL166B

(JCL166F) is not sufficient to rescue growth in minimal media (Figure 3-3a). These results suggest that other physiological factors contribute to metabolic impairment.

Redistribution of carbon resources through altered pathway expression

We further aimed to isolate additional physiological factors that contribute to the robust growth of BP1. We compared activity of butanol pathway enzymes before and after evolution. Specific activity of enzymes within cell lysate was measured after 24 hours of anaerobic growth. AtoB, Hbd and Crt show a decrease in specific activity (Figure 3-5a). We suspected that this observation is due to an alteration in plasmid copy number. ColE1 plasmids, as well as ColE1 derivatives that include ColA, are a class of Pol I-dependent plasmids whose replication relies solely on host-encoded proteins(*35*). Genomic sequencing showed a mutation in *pcnB* (Plasmid Copy Number B) (G632A, R312H), which encodes for poly(A) polymerase I (Pol I). It has been reported that even though *pcnB* deletion strains show no growth defects, they are defective for plasmid maintenance with ColE1 origins(*36*). As a result, ColE1 related plasmids show instability in *pcnB* deletion mutants and exhibit decreased copy numbers. We therefore suspected that the mutation in *pcnB* (G632A) may alter or diminish the copy number of the pathway plasmids.

The relative plasmid copy number was measured using qPCR and compared between JCL166B and evolved strain BP1. As anticipated, after evolution there was roughly a 4-fold decrease in plasmid copy number of pEL11 and a 5-fold decrease in pIM8 (Figure 3-5b). This observation can explain the diminished activity of butanol enzymes. To verify that the mutation in PcnB is the cause of this diminished plasmid copy number, we reverted the point mutation in PcnB (BP1

WT*pcnB*) and measured the relative plasmid copy number as well. This subsequently restored levels of plasmid to match those of unevolved parent JCL166B (fFigure 3-5b). BP1 WT*pcnB* also exhibited a reduced growth rate compared to BP1 in minimal media (Figure 3-3a).

The presence of plasmids that overexpress butanol pathway enzymes forces a large amount of *de novo* synthesized amino acids and nucleotides to be used to maintain high protein expression levels and plasmid copy numbers. In general, it has been observed that enzymes involved with protein synthesis, as well as synthesis of nucleotides, are decreased within cells growing in rich media in response to the surplus of sufficient building blocks(*37*). These results suggest that by modulating the concentration of intracellular plasmids, this burden is deceased to allow for carbon flux to be redirected toward synthesis of compounds essential for growth

Metabolic perturbations optimize carbon usage

To evaluate systematic perturbations that contribute to butanol production in minimal media, we used both metabolomic and proteomic analysis to compare BP1 with JCL166B. Ion pair LC/MS/MS was used for obtaining the metabolome profiles for JCL166B and BP1. A total of 44 metabolites with significantly perturbed values ($P \leq 0.05$) were annotated (Figure 3-6a). This data was subjected to principal component analysis and showed that JCL166B and BP1 separated by PC1 (68.9%), while the PC2 (10.2%) separated between the replicates within each strain.

Of the significantly reduced metabolites are nucleosides uridine, thymidine, cytidine, and guanosine. In addition to this are other nucleotide monophosphates and diphosphates such as CDP, CMP, ADP, GMP, AMP and UMP. Relative concentrations of nucleotides of BP1

compared to JCL166B are listed in Figure 3-6B. As shown, nucleotides with decreasing phosphorylation have deceased relative concentrations, corresponding to a relative increase in adenosyl energy charge. The physiological effects of energy charge are based on the ratio of these nucleotides(*38*), and an increase in energy charge is an indicator that overall metabolism supports greater biosynthetic processes.

Lactate and succinate showed increased measurements within BP1. While native fermentation pathways are deleted, several other enzymes can produce these metabolites. Specifically, succinate can be produced from basal activity of the glyoxylate shunt, and lactate can be produced using membrane bound, FAD dependent lactate dehydrogenase (Dld). Contributions of lactate and succinate fermentation to the growth phenotype of BP1 was shown to be negligible as no growth was observed after pathway plasmids were cured (Figure 3-2b).

Using the tandem mass tagging (TMT) quantitative proteomics, we have identified 1874 proteins. Of which, 1847 proteins are quantifiable. Measured proteins were categorized by biological process, molecular function, and usage within specific metabolic pathways by cross referencing them within the EcoCyc database. Protein functionality groups are listed in Figure 3-4A, and the number of proteins that had significantly increased reads (> 1.5 fold increased) compared between JCL166B and BP1 were noted. Consistent with expectations and the observed increase in energy charge, levels of proteins involved in biosynthetic processes were increased in BP1. These include amino acid biosynthetic processes and fatty acid biosynthesis (Figure 3-7ab). On the contrary, proteins involved in catabolic processes have relatively higher levels in JCL166B including β-oxidation and various amino acid degradation pathways (Figure 3-7a). Also consistent with previous analysis, *aceE* and *aceF* showed respective increases of 1.61 fold, 1.57 fold.

Metabolomic measurements show that concentrations of CoA were decreased, suggesting that excess CoA is likely being utilized within the CoA dependent butanol pathway. This is supported by measurements of acetyl-CoA and butanoyl-CoA, the only two butanol pathway intermediates detected, which both showed significantly increased concentrations in BP1. A previous report that improved titers of butanol in rich media did so by increasing the expression of AdhE2, which released free CoA from the pathway to improve pathway recycling. It was also hypothesized that by supplying additional CoA, flux could be increased through the pathway. Within BP1, aspartate 1-decarboxylase (PanD) shows a 2.1 fold increase in protein level. PanD has been shown to be the rate limiting enzyme in synthesis of pantothenate, a CoA precursor. Increases in pantothenate were measured in BP1, suggesting an adaptive mechanism to increase the supply of CoA for function of the butanol pathway.

Discussion

While deletion of native fermentation pathways results in high-titer production of butanol in rich media, this major systematic modification impaired metabolic function required for growth coupled production in minimal media. To restore metabolic function, we employed an evolution based strategy with an enhanced rate of mutagenesis conferred by mutator MutD5. Following evolution and further engineering, butanol production of 2g/l was achieved. Interestingly, evolution also resulted in increased butanol yields and decreased butanol titers, which suggest alterations to energy metabolism.

Genome sequencing was performed to identify mutations in BP1 that may contribute to altered energy metabolism or other factors that contribute to growth. Mutations in *ihfB* and *arcB* were identified that work in combination to upregulate expression of pyruvate dehydrogenase (PDH). An increase in PDH activity is able to increase supply of NADH as an alternative to formate production with PflB. Reversion of either of these mutations causes significant decreases in butanol yields, supporting the notion that PDH affects growth and byproduct formation.

It is not apparent as to why increased NADH supply improves the growth phenotype. Nevertheless, because additional overexpression of FDH (JCL166F) is further unable to grow in minimal media, we can conclude that other factors likely contribute to the robust growth phenotype of BP1.

A mutation in *pcnB* was observed that decreased copy numbers of plasmids containing butanol pathway enzymes. This mutation effectively reduced the activity of these enzymes. Overexpression of these enzymes is likely burdensome to the cell as all amino acids must be synthesized *de novo.* Therefore, this mutation allows a preservation of amino acids and other resources that can be redirected towards cell division and homeostasis.

These results demonstrate the utility of directed strain evolution for addressing non-obvious factors that impair metabolic function. These results also demonstrate that directed strain evolution has potential to be used for reducing byproduct formation and enhancing production yields. Further, by elucidating the underlying strategies used by the cell, we are able to observe important considerations that can aid in further rational engineering approaches.

Materials and Methods

Reagents

T4 DNA polymerase was purchased from New England Biolabs (Ipswich, MA). KOD Xtreme DNA polymerase and Bugbuster were from EMD Millipore (Billerica MA). Kapa HiFi was from Kapa Biosciences (Wilmington, MA). Oligonucleotides were purchased from IDT (San Diego, CA). All chemicals were purchase from Fisher Scientific (Pittsburg, PA) or Sigma-Aldrich (St. Louis, MO).

Bacterial Strains

A list of strains and plasmids are listed in table 1. XL1-Blue and XL1-Red were purchased from Agilent Technologies (Santa Clara, CA). Strain JCL166 was acquired from previous studies(*39*).

Plasmid and strain construction.

A list of plasmids are listed in table 1. RBS sequences were designed using the Salis RBS calculator(*40*). All fragments were amplified with Kapa HiFi or KOD Xtreme and subjected to further digestion with DPN1 to remove methylated DNA template. Fragments were designed with 20bp overlaps. Fragments were assembled using T4 DNA polymerase using the following protocol: Fragments are pooled in equimolar concentrations with a final concentration between 30-80µl. 8.7µl of this mixture was mixed with 1µl of NEB buffer #2, and 0.3 µl of T4 DNA polymerase. The reaction was incubated at room temperature for 10 minutes and the entire reaction mixture was transformed into XL1-Blue. Plasmid sequences were further verified using Sanger sequencing with Laragen (Culver City, CA). MutD5 was acquired by amplifying the

DnaQ gene from XL1-Red. Gene deletions were performed using λ-Red recombination as previously described(*41*).

Media and Growth Conditions.

Strains were routinely cultured in LB broth. Commercial 5x M9 salts were used as the basis for minimal media with final concentrations of 0.4g/l of glucose and 2mM $MgSO_4$, 0.1mM$CaCl_2$, 1ml/L thymine.

Minimal media growth conditions are as follows: Cultures were placed into Vacutainer serum tubes or for larger cultures, Wheaton Glass Serum Bottles. The bottles were topped and crimped, then a needle (21G by 1 ½ in.; BD) paired with a Millipore polyethersulfone (PES) 0.22 μm filter was inserted through the rubber stopper of the serum bottle. The bottles were placed in an anaerobic transfer chamber where the oxygen in the headspace and medium was purged by repeated cycles of vacuuming and filling with nitrogen and hydrogen. The needles were removed from the stoppers inside the anaerobic chamber.

Evolution

Parent strains were transformed with pALQ32. Single colonies were precultured in LB and grown overnight at 37C. 1:100 inoculations were made into minimal media in anaerobic conditions containing appropriate antibiotics and 0.1mM of isopropyl-β-D-thiogalactopyranoside (IPTG). Cultures were provided with minimal amounts of supplementation so that the culture could grow to an OD_{600} of 0.4-0.6 in 24 hours. This culture was used to inoculate fresh media with limiting supplementation at 1:100. This process continued until a desired growth phenotype was observed.

Quantification of mutation rate

The mutation rate of pALQ32 was determined by observing the increase in rifampicin resistance. JCL16 with and without pALQ32 were grown in LB overnight with 0.1mM IPTG and plated the next day onto LB agar plates containing 100µl/ml rifampicin. Corresponding fold increases in CFU formation were used to determine the increased rate of mutagenesis when grown with pALQ32.

Butanol production

Individual colonies were used to inoculate 3ml of LB overnight preculture containing appropriate antibiotics. This culture was used to inoculate 3ml of minimal media in vacutainer tubes and immediately made anaerobic. Media contained antibiotics (carbenicillin 100 µg/ml, kanamycin 50µg/ml, and chloramphenicol 50µg/ml), 4g/l glucose and 0.1mM IPTG. The culture was grown at 37C. Every 24 hours, the pH was adjusted to pH 7 using 10M NaOH, and 4g/l of glucose was added.

Quantification of metabolites

Alcohols were quantified using GC (gas chromatography) with flame ionization detection. Agilent 6890N with 6850 autosampler. Culture samples were centrifuged at 6000rpm for 5 minutes. Supernatant was injected directly into the GC using 1-propanol as an internal standard. Detailed procedures are described previously(*42*).

Glucose and organic acids were quantified using HPLC with Aminex HPX-87H column (Bio-Rad) with 30mM H_2SO_4. Detailed procedures are described previously(*43*).

Enzyme assays

Cell extracts were prepared in aerobic conditions. Cells were collected by centrifugation. The pellets were lysed using Bugbuster Protein Extraction Reagent (Millipore). The lysate was centrifuged at 13,200 rpm, 10 minutes, 4ºC and the supernatant was collected for enzyme assays.

All enzyme assays were performed spectrophotometrically using the Biotek Powerwave XS microplate reader at 30ºC aerobically. Reaction mixture volumes were 150 µL. Protein concentrations were determined using a Bradford assay. All assays were initiated by the addition of the cell extract.

Enolase assay. The enolase activity was measured by the increase in absorption at 240 nm, corresponding to the formation of phosphoenol pyruvate (PEP). The reaction mixture contained 87 mM Tris-HCl, pH 8.0, 100 mM KCl, 1 mM Mg SO4, and 1 mM 2-phosphoglycerate.

Aldolase assay. The aldolase activity was measured by an increase in absorption at 340 nm. The reaction mixture contained 100 mM Tris-HCl, pH 8.0, 0.5 mM NAD+, triose phosphate isomerase (1 unit/reaction), glyceraldehyde-3-phosphate dehydrogenase (2 units/reaction), and 5 mM fructose-1,6-bisphosphate.

Glucokinase assay. The glucokinase activity was measured by an increase in absorption at 340 nm. The reaction mixture contained 100 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 1 mM ATP, 0.4 mM NADP+, glucose-6-phosphate dehydrogenase (0.2 units/reaction), and 20 mM glucose.

Phosphoglucose isomerase assay. The phosphoglucose isomerase activity was measured by an increase in absorption at 340 nm. The reaction mixture contained 100 mM Tris-HCl, pH 8.0, 3 mM MgCl2, 0.2 mM NADP+, glucose-6-phosphate dehydrogenase (0.2 units/reaction), and 0.3 mM fructose-6-phosphate.

Phosphofructokinase assay. The phosphofructokinase activity was measured by a decrease in absorption at 340 nm. The reaction mixture contained 100 mM Tris-HCl, pH 8.0, 0.4 mM NADH, 1 mM ATP, 10 mM MgCl2, 2 mM NH4Cl, aldolase (0.2 units/reaction), triose phosphate isomerase (1 unit/reaction), glycerophosphate dehydrogenase (2 units/reaction) and 5 mM fructose-6-phosphate.

Phosphoglyceromutase assay. The phosphoglyceromutase activity was measured by a decrease in absorption at 340 nm. The reaction mixture contained 100 mM Tris-HCl, pH 7.0, 20 mM KCl, 5 mM MgSO4, 2 mM ADP, 0.4 mM NADH, enolase (2 units/reaction), pyruvate kinase (4 units/reaction), lactate dehydrogenase (4 units/reaction), and 2 mM 3-phosphoglycerate.

Glyceraldehyde-3-phosphate dehydrogenase assay. The glyceraldehyde-3-phosphate dehydrogenase activity was measured by an increase in absorption at 340 nm. The reaction mixture contained 100 mM potassium phosphate buffer, pH 7.6, 1 mM NAD+ and 1 mM glyceraldehyde-3-phosphate.

Phosphoglycerate kinase assay. The phosphoglycerate kinase activity was measured by a decrease in absorption at 340 nm. The reaction mixture contained 100 mM Tris-HCl, pH 7.5, 2 mM MgSO4, 1 mM ATP, 0.5 mM NADH, 1 mM 3-phosphoglycerate, and glyceraldehyde-3-phosphate dehydrogenase (5 units/reaction).

AtoB assay. The AtoB activity was measured by the decrease in absorption at 303 nm, corresponding to the absorption bond enolate complex formed by acetoacetyl-CoA and Mg2+. The reaction mixture contained 100 mM Tris-HCl, pH 8.0, 10 mM MgSO4, 0.2 mM acetoacetyl-CoA, and 0.2 mM CoA.

Crt assay. The Crt activity was measured by the decrease in absorption at 263 nm, corresponding to the disruption of the α-β unsaturation of crotonyl-CoA. The reaction mixture contained 100 mM Tris-HCl, pH 7.5 and 0.1 mM crotonyl-CoA.

Ter assay. The Ter activity was measured by a decrease in absorption at 340 nm. The reaction mixture contained 100 mM potassium phosphate buffer, pH 6.2, 0.2 mM NADH, and 0.2 mM crotonyl-CoA.

Hbd assay. The Hbd activity was measured by a decrease in absorption at 340 nm. The reaction mixture contained 100 mM 3-(N-morpholino) propanesulfonic acid (MOPS), pH 7.0, 0.2 mM NADH, and 0.2 mM acetoacetyl-CoA.

DNA Sequencing

Genomic DNA was purified using a Qiagen DNAeasy Blood and Tissue Kit. DNA concentrations were normalized to 1ng/µl using a Qubit fluorescent reader. This was used as input DNA for Illumina Nextera XT DNA Sample Preparation. Resulting libraries were pooled a 5nM. Libraries were sequenced on Illumina 3000 with 100bp single end reads to a minimum of 10x coverage, and an average of 42x coverage. The adapter sequences were removed from reads using Trim Galore! (http://www.bioinformatics.babraham.ac.uk/) with quality trimming turned off. Trimmed reads were mapped using BWA-MEM v.0.7.12-r1039 (Li 2013) to the Escherichia coli str. K-12 substr. MG1655 genome (NCBI Accession NC_000913). Variant discovery and filtering was done with GATK v 3.7-0-gcfedb67(*44*) using HaplotypeCaller in GVCF mode with ploidy 1, followed by GenotypeGVCFs, and finally VariantFiltration setting a minimum QD of 2. SnpEff[45] was used to determine the context of the variants and predict the functional impact. Additional custom scripts were used to identify important variants.

Metabolite Sampling and Extraction

After cultivating strains for 24 hours, three OD_{600} units of cells were collected by fast filtration, using 47 mm diameter nylon membrane with pore size 0.45 µm (Millipore, MA, USA), for intracellular metabolome analysis. Cells were quenched by liquid nitrogen immediately and then stored at -80 °C until extraction. Extraction was performed by addition of 1.8 mL extraction

solvent (methanol/water/chloroform = 5:2:2 v/v/v) with internal standard (+)-10 camphorsulfonic acid (25 µg/L) to the filter in 2 mL sampling tube. The sampling tubes were incubated at -30 °C for 1 hour, after which 1050 µL of solution was collected and added to 525 µL of ultrapure water. After vortexing, the solution was centrifuged at 4 °C, 10,000 rpm for 5 min to separate polar and non-polar phase. 700 µL of the upper polar phase was then transferred to a new tube via syringe filtration (0.2 µm PTFE hydrophilic membrane, Millipore, MA, USA). Methanol in the samples was removed by centrifugal concentration for about 2 hours after which the samples were lyophilized by freeze drying overnight. Samples were stored at -80 °C till analysis. Lyophilized samples were dissolved in 50 µL of ultrapure water for ion-pair liquid chromatography mass spectrometry (IP-LC/QqQ-MS) analysis.

Intracellular sample analysis using IP-LC/QqQ-MS:

IP-LC/QqQ-MS analysis was carried out using Shimadzu Nexera UHPLC system coupled with LCMS 8030 Plus (Shimadzu Corp., Kyoto, Japan). The column used was CERI (Chemicals Evaluation and Research Institute, Tokyo, Japan) L-column 2 ODS (150 mm × 2.1 mm, particle size 3 µm). The mobile phase (A) was 10 mM tributylamine and 15 mM acetic acid in ultrapure water whereas mobile phase (B) was methanol. The column oven temperature was 45 °C. The flow rate was 0.2 mL/min and the concentration of B was increased from 0% to 15% from 1.0 to 1.5 min, 15% to 50% from 3.0 to 8.0 min and 50% to 100% from 8.0 to 10.0 min, and held until 11.5 min. From 11.5 min, the concentration was decreased to 0% and held until 17 min. Negative ion mode was used for mass analysis. Injection volume was 3 µL and probe position was +1.5 mm. Desolvation line temperature and heat block temperature was 250 °C and 400 °C, respectively, nebulizer gas flow was 2 L/min and drying gas flow was 15 L/min. Analysis was performed with multiple reaction monitoring (MRM). The raw data obtained was converted to

analysis base file (abf) format by using freely available file format converter (Reifycs Inc., Tokyo, Japan). MRMPROBS (Tsugawa, et al., 2013) was used for automatic peak picking and for calculation of peak area. The peaks were normalized by internal standard (+)-10 camphorsulfonic acid.

Multivariate analysis

SIMCA-P+ version 13 (Umetrics, Umeå, Sweden) was used for principal component analysis (PCA). Metabolome data was mean centered to unit variance and transformation was not performed. Heatmap of the annotated metabolites was obtained using Multiexperiment Viewer Version 4.9, freely available for download at http://mev.tm4.org/. Hierarchical cluster analysis was obtained based on Pearson correlation with gene leaf optimization and complete linkage clustering.

Proteomics

RT-qPCR

Bacterial samples were harvested and immediately mixed with Qiagen RNA Protect bacterial reagent as per the manufacturers instructions. RNA was purified using RNeasy purification kit with additional DNase digestion step as per the manufacturer's instructions. Purified RNA was eluted in 10mM EDTA. Luna Universal One-Step RT-qPCR Kit (NEB) was used to measure all RNA samples. Primers are listed below in table 2, and are verified to have efficiency between 90 and 105%. Gene *frr*, ribosome-recycling factor, was used as a reference gene[46]. Results were analyzed using Bio-Rad CFX Manager 2.0.

Figures

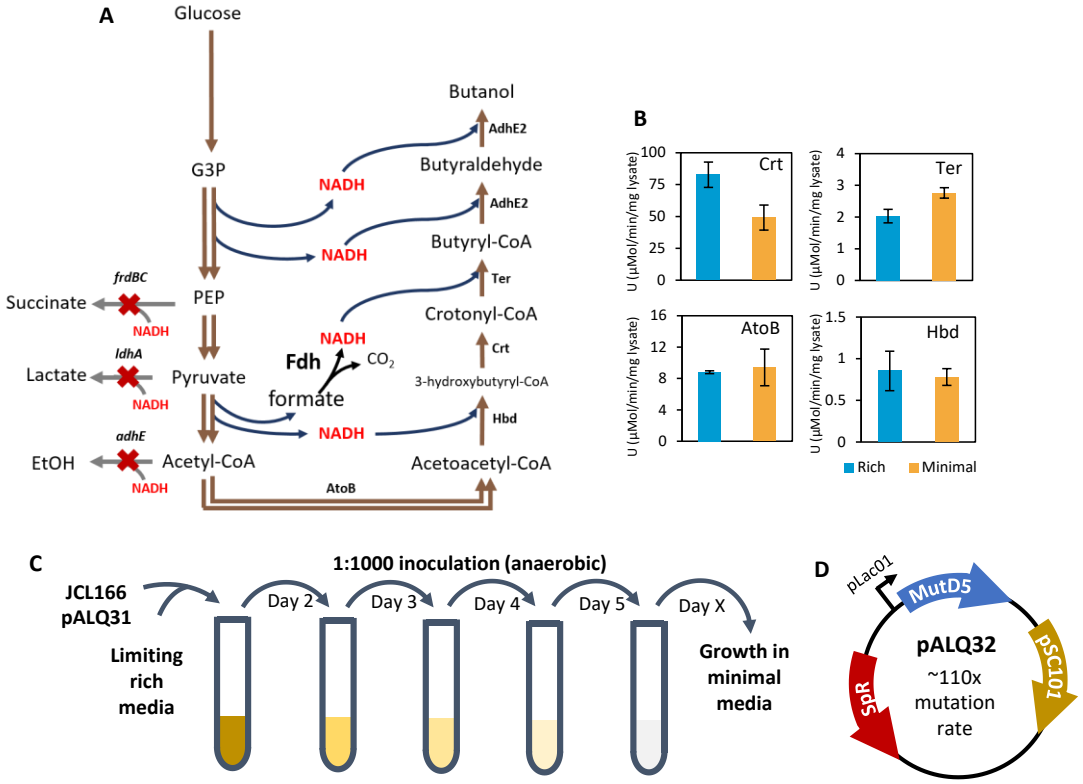


Figure 3-1: Overview of butanol pathway and evolution strategy

a) Illustration of the *Clostridial* CoA-dependent butanol pathway. Native fermentation pathways are deleted, forcing the cell to utilize the butanol pathway as a sole electron sink under anaerobic conditions.

b) Specific activity of enzymes in the butanol pathway are measured and compared between JCL166B grown in rich media (blue) and minimal media (yellow)

c) The evolution strategy involves anaerobic successive serial dilutions with decreasing nutritional supplementation until growth is observed within minimal media.

d) Illustration of pALQ32, which contains mutator gene MutD5.

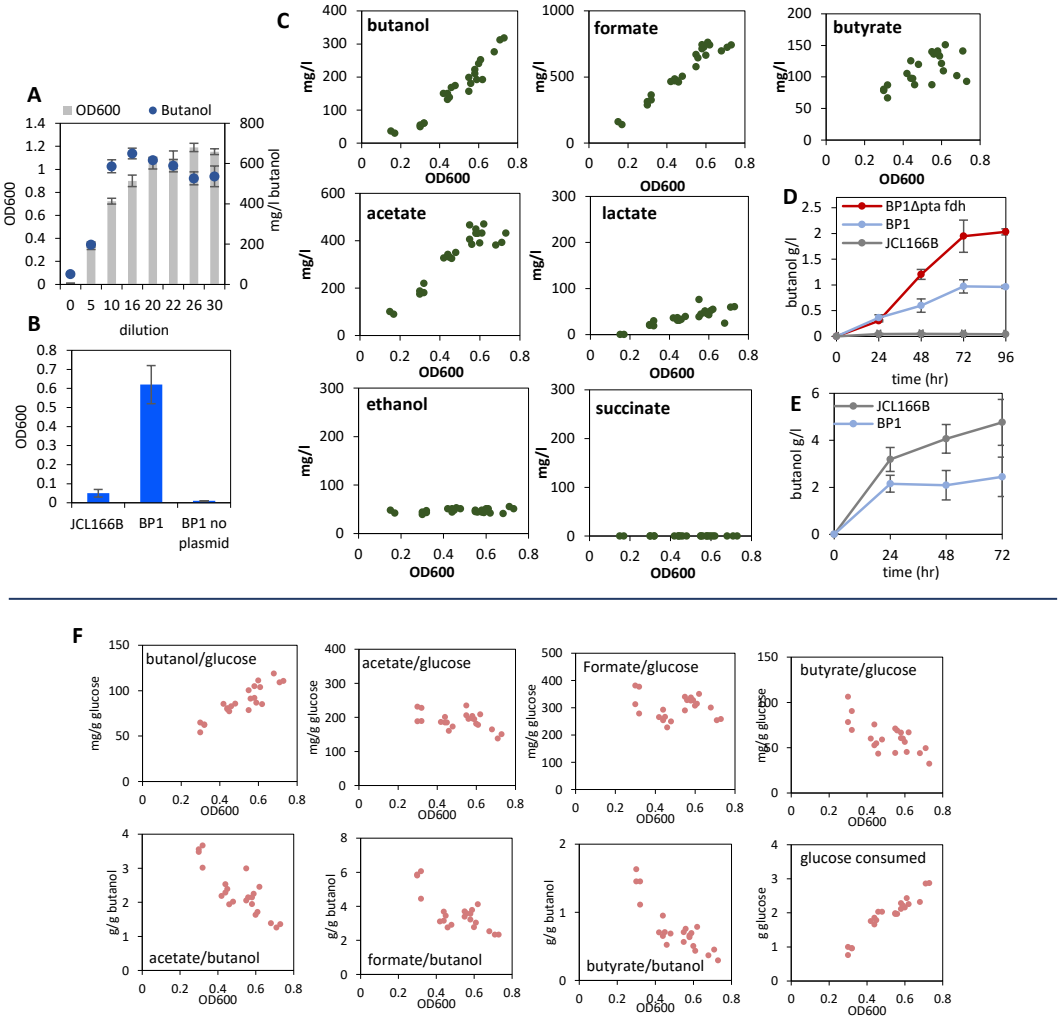


Figure 3-2: Growth and production phenotypes after evolution

a) Growth and butanol production over the course of evolution. After 5 serial dilutions JCL166B obtained the ability to grow on minimal media. After additional dilutions, growth and butanol production both improved.

b) Curing the plasmids that harbor butanol enzymes abolishes the growth phenotype in minimal media.

c) MutD5 was cured from an evolved strain. Individual clonal strains were tested for growth and production. Increased growth phenotypes were correlated with increased byproduct formation and butanol production.

d) After deletion of *pta* and overexpression of Fdh (pCS138) in clonal strain BP1*Δpt*a fdh, 2g/l of butanol was produced after 96 hours.

e) Production of BP1 in rich media is lessened compared to production before evolution by JCL166B.

f) Product yields, product ratios, and glucose consumption from clonal evolved strains plotted against increasing OD600.

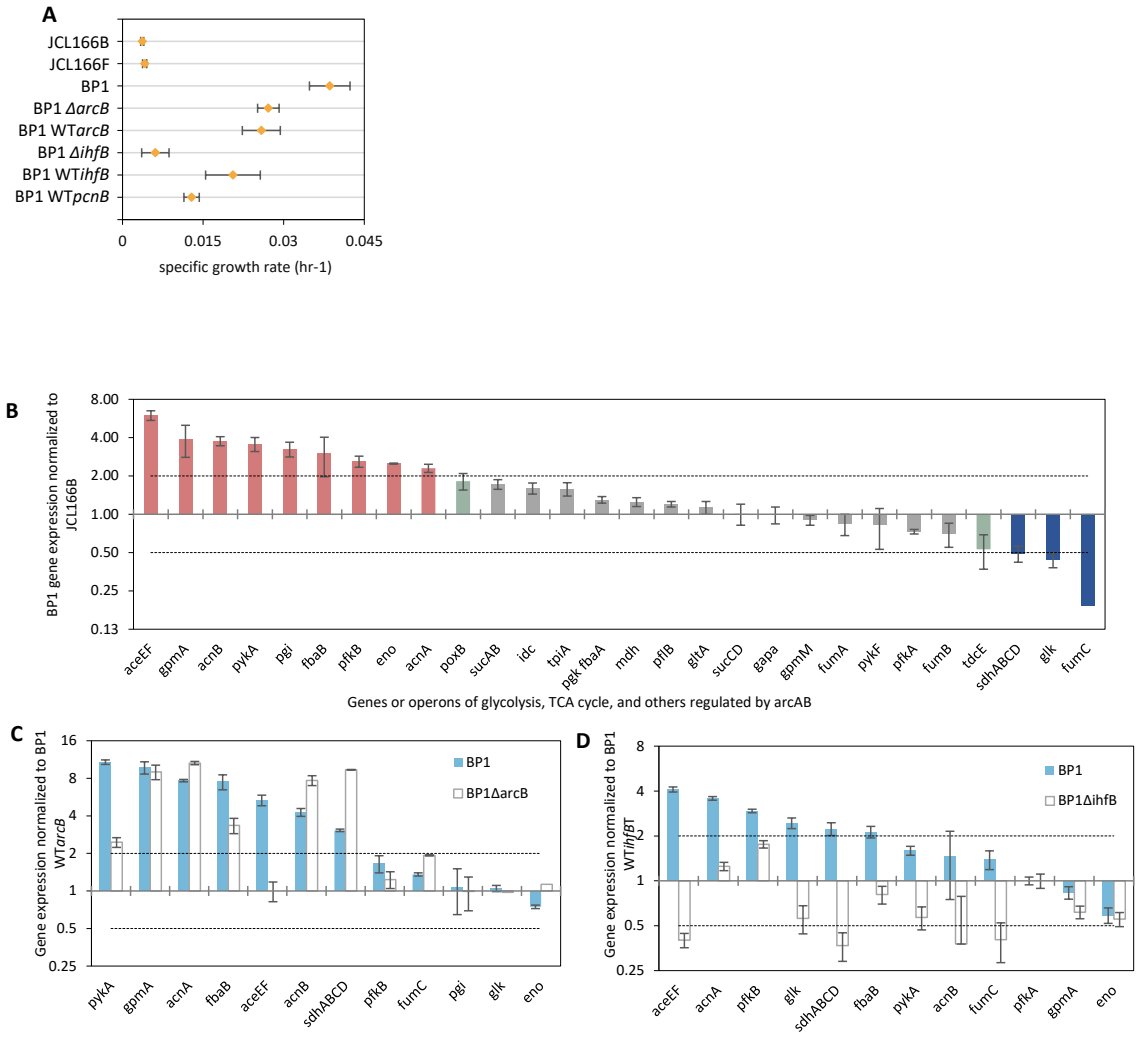


Figure 3-3: Specific growth rates and gene expression

a) Specific growth rate of JCL166B, BP1 and all derivatives. BP1 with deletions of *arcB* (BP1 *ΔarcB*) or *ifhB* (BP1 *ΔihfB*). BP1 with wildtype sequence of *arcB* (BP1 WT*arcB*), *ihfB* (BP1 WT*ihfB*), or *pcnB* (BP1 WT*pcnB*).

b) Relative mRNA abundance (measured by RT-qPCR) of genes within central carbon metabolism comparing BP1 to JCL166B. Red bars represent enzymes with significantly increased expression, while blue bars represent those with decreased expression. The gray bars represent those insignificantly changed. Cutoff is set with a $\log_2$ ratio of 1.

c) Enzymes with significantly perturbed expression in BP1 compared to JCL166B are tested to see if the expression change is due to the point mutation observed in either *arcB* or *ihfB*. C) Gene expression of BP1 (blue) and BP1*ΔarcB* (white) are normalized to BP1 WT*arcB.* D) Gene expression of BP1 (blue) and BP1*ΔihfB* (white) are normalized to BP1 WTihf*B.*

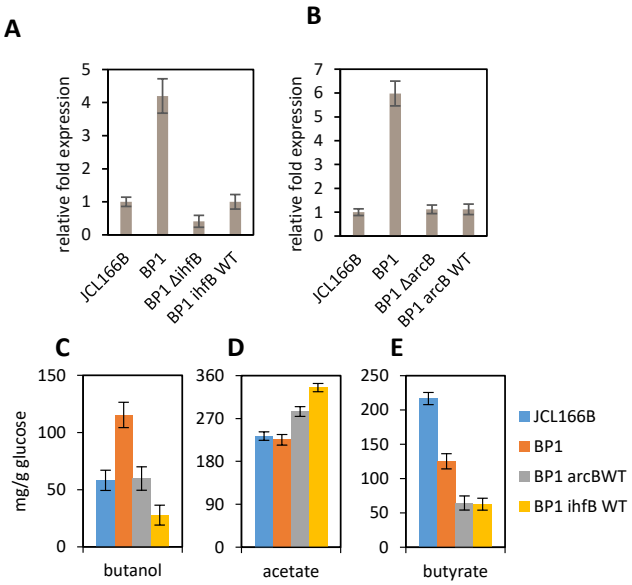


Figure 3-4: Byproduct formation and PDH expression after BP1 modifications

A/B) – Relative expression of *aceE* between JCL166B and BP1 derivatives. Results show that increased *aceE* expression is dependent on mutations in both IhfB and ArcB.

C/D/E) **–** Byproduct yields after 24 hours production in minimal media

A

	JCL166B	BP1
	(µMol/min/mg protein)	
AtoB	9.41 ± 0.39	2.76 ± 0.39
Hbd	0.78 ± 0.10	0.15 ± 0.01
Crt	49.2 ± 9.8	33.5 ± 0.43
Ter	2.76 ± 0.16	3.46 ± 0.56

B

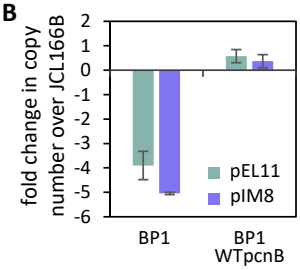


Figure 3-5: Altered plasmid copy number and pathway expression in BP1

a) Specific activity of butanol pathway enzymes measured within cell lysate of BP1 and JCL166B

b) Relative plasmid copy number of pEL11 and pIM8 within BP1 compared to JCL166B. Plasmid copy numbers are decreased within BP1. BP1 WT*pcnB* (reversion of *pcnB* to wildtype sequence) restores plasmid levels to match JCL166B.

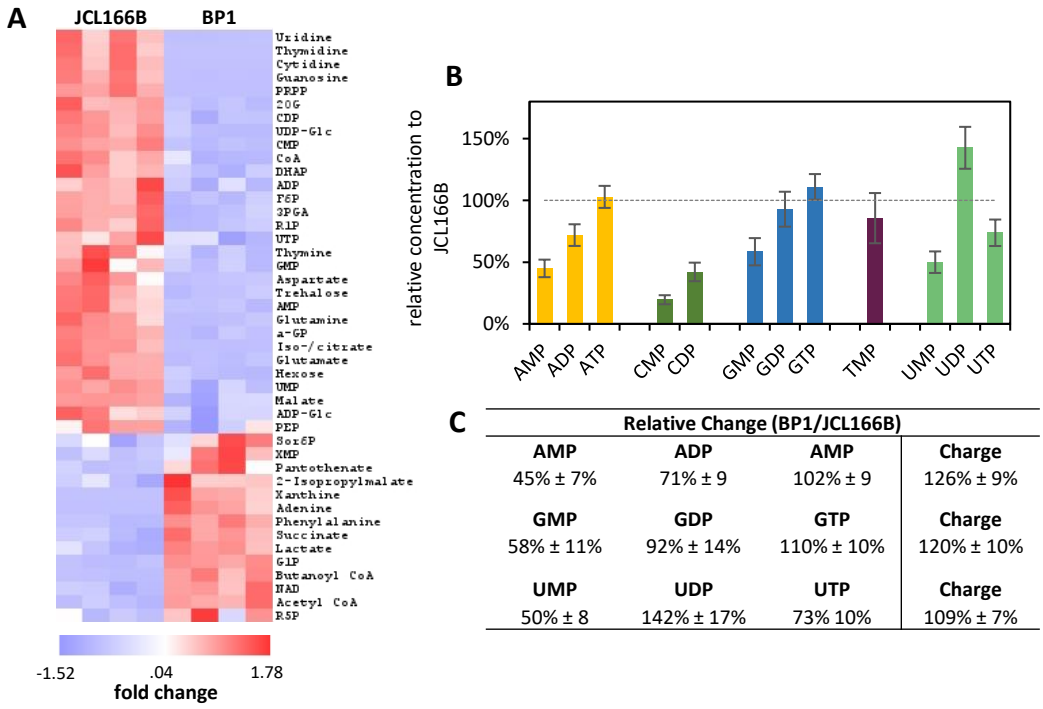


Figure 3-6: Metabolomic analysis of BP1

a) Relative intracellular metabolites comparing BP1 to JCL166B. All 44 metabolites that are significantly perturbed ($P \leq 0.05$) are listed.

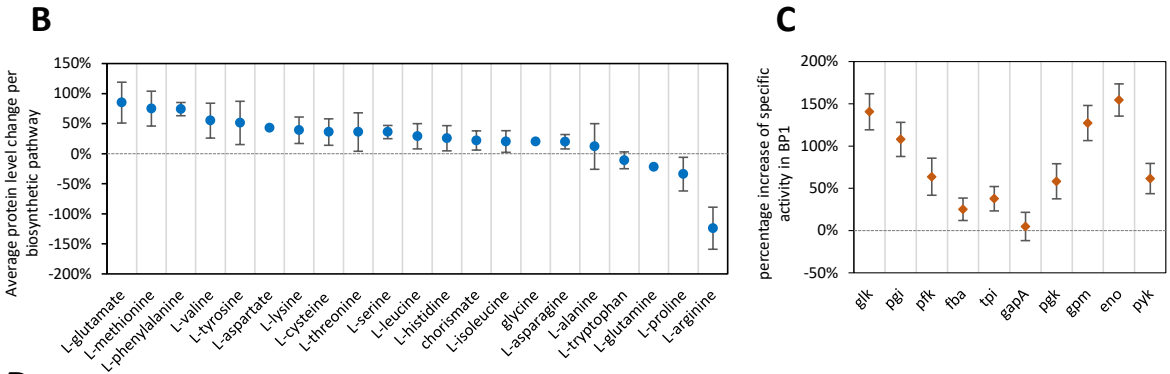
b) Relative intensities of nucleotides comparing BP1 to JCL166B.

c) Change in nucleotide energy charge after evolution in BP1 shows 26% increasing adenosyl energy charge.

A

Functionality	No of proteins* Higher in JCL166B	Higher in BP1	Total proteins
Whole genome	129	152	1741
All biosynthetic processes	12	47	390
Amino acid biosynthesis	5	28	89
fatty acid biosynthesis	0	2	14
Nucleotide metabolism	1	2	68
Catabolic processes	37	11	223
β-oxidation	6	1	10
DNA binding	8	18	231
DNA replication	1	3	44
Oxidoreductase	23	23	228
Transporter	7	9	100

*Number of proteins with greater than 1.5 fold increase



D

	Units (µMol/min/mg lysate)									
	Glk	**Pgi**	**Pfk**	**Fba**	**Tpi**	**GapA**	**Pgk**	**Gpm**	**Eno**	**Pyk**
JCL166B	0.133 ± 0.02	5.45 ± 1.1	2.51 ± 0.43	1.19 ± 0.41	2.44 ± 0.31	3.71 ± 1	18 ± 4.1	0.11 ± 0.03	0.11 ± 0.02	0.13 ± 0.01
BP1	0.32 ± 0.034	11.33 ± 2	4.11 ± 0.65	1.49 ± 0.35	3.26 ± 0.75	3.98 ± 0.53	28 ± 1.36	0.25 ± 0.01	0.21 ± 0.04	0.21 ± 0.04

Figure 3-7: Proteomic analysis of BP1 and glycolysis enzyme activity

a) Proteins with increased expression in JCL166B and BP1 are grouped by functionality.

b) Average protein level changes of enzymes within various amino acid biosynthetic pathways in BP1 compared to JCL166B.

c) Relative specific activity of glycolytic enzymes in BP1 compared to JCL166B shows an increase in activity after evolution.

d) Specific activity of glycolytic enzymes within cell lysate of JCL166B and BP1

Table 3-1: Plasmids and strains used in this work

Strains		**Reference**
JCL166	BW25113/F' [traD36 proAB$^+$ lacI^{q}ZΔM15 (Tet^r)] *ΔldhA ΔadhE ΔfrdBC*	[19]
JCL166B	JCL166/pEL11/pIM8	This work
JCL166F	JCL166/pEL11/pIM8 /pCS138	This work
BP1	JCL166B evolved to function in minimal media	This work
BP1*Δpta* fdh	BP1*Δpta*/pCS138	This work
BP1 *ΔarcB*	BP1*ΔarcB*	This work
BP1 WT*arcB*	BP1 with wildtype *arcB* sequence	This work
BP1 *ΔihfB*	BP1*ΔihfB*	This work
BP1 WT*ihfB*	BP1 with wildtype *ihfB* sequence	This work
BP1 WT*pcnB*	BP1 with wildtype *pcnB* sequence	This work
Plasmids		
pALQ32	$P_L lacO_1$::*mutD5* pSC101 *ori* Spec^r	
pEL11	$P_L lacO_1$::*atoB*$_{EC}$-*adhE2*$_{CA}$-*crt*$_{CA}$-*hbd*$_{CA}$ ColE1 *ori* Amp^r	[19]
pIM8	$P_L lacO_1$::ter$_{TD}$ Cola *ori* Kan^r	[19]
pCS138	$P_L lacO_1$:: fdh$_{CB}$ pSC101 *ori* Cm^r	[19]

Table 3-2: Primers used for RT-PCR

fumC	ACCGCCGAGTAATTCACTGG
	CGAATTCCCGCTGGCTATCT
idc	CTCTGGTGCACAAAGGCAAC
	TCGATCAGTTCACCGCCAAA
gltA	GTCACTGTGCATTTCGCTCC
	TCCGTCTTCCATGTTCACCG
acnA	TTGGTACTGACTCGCACACC
	TCTGCTTCGATCCCACCAAC
acnB	AAATGGCTGGCGAATCCTGA
	GAGCACACAGGATTGGCTCT
mdh	CAGGTTTTTCACGATGCCGG
	AAGGCGCAGATGTCGTTCTT
fumA	CCCAGACCAAGATTTTGCGC
	AAGGGAATGAGCACGGTCAG
fumB	GATGCTTGATGCTCTGCTGC
	TGATCATGCTGGCGAAAGGT
sucB	TTTGAAAAACGCCACGGCAT
	TTCCGGGTAACGTTTCAGGG
sucC	GCGATGGGTACGATGGACAT
	ACGCTTCGGTTACACGTTCT
sdhA	CCGACCAAAGTTACCGGTCA
	TTAGCGCCGTGTACCGATAC
aceE	CTCCGGTTCTATCCTGCGTC
	CAGCATGTTCCAGCGTTCAC
pflB	CTGAGTCGGGATAGCGTCAC
	CCTGGCTGTTGACCTGGTAG
poxB	CGTTGCAGGGCTTCATCAAC
	GACACAAACTTTGCCCGCAT
frr	TGGATGGCATTGTCGTGGAA
	ATCGGACGCCATAATCGCTT
pgi	AGCTCTGCGTCCGTACAAAA
	CAAGAACAGCGTGGTTTCCG
gapA	GCTCGTAAACACATCACCGC
	CGATGTCCTGGCCAGCATAT
pgk	CTTGTGCCGCGATAAAGGTG
	ATGGTGGCTATCGTTGGTGG
fbaA	AGAGCGGGTTAGCAGCAAAT
	GCAGACGTTGTACAACACCG
tpiA	GGCTGAGCAAACAGTTCTGC
	AAATCTGCAACTCCGGCTCA
pfkA	AAAAACGTGGTATCGACGCG
	GGAAGCCCATTTCGGTCAGA

pfkB	TCCCCGTCGCTACTGTAGAA
	ATACTGCTCACCGCTTGCTT
gpmM	GCACGCCAGTCGGGATATTA
	ATCATCGCTGCACACGGTAA
gpmA	TCTGCAACCGGAAATGAGCT
	GATGGTGTCGTATTTGCCGC
eno	CTTTCACAGTTTTCGCGCCA
	ACATCGCTGAACTGAACGGT
pykA	GTACGCTGAACCTGACTGCT
	GCAGATTAACCGCTTCGCTG
pykF	ATCCGTGCACGTAAAGTCGT
	CGTCAGTACCGTCGAGGATG
tdcE	GAATCGACTGGGTAGGGACG
	TCGACGGTGAATATCCGCAG

Table 3-3: Mutations in BP1

Position	Reference	Alternate	Mutation Type	Gene	Residue change
4514230	C	T	non-synonymous	fecB	Glu27Lys
4467419	A	G	intergenic		
4442349	A	AG	intergenic		
4389447	A	G	non-synonymous	psd	Ile305Thr
4304171	A	G	non-synonymous	ytcA	Leu78Pro
4250594	A	G	intergenic		
4234120	C	T	synonymous	pgi	Asn121Asn
4233494	A	G	intergenic		
4214221	T	C	intergenic		
4154289	T	C	non-synonymous	argE	Ile187Val
4154275	G	A	synonymous	argE	Gly191Gly
4073641	C	T	intergenic		
4030338	T	C	non-synonymous	fadB	Lys212Glu
3861296	A	G	non-synonymous	glvG	Val231Ala
3840009	T	C	intergenic		
3787593	A	G	synonymous	envC	Glu252Glu
3720309	AT	A	intergenic		
3698666	A	G	synonymous	bcsG	Val151Val
3681691	T	C	non-synonymous	yhjJ	Ile84Val
3584679	G	A	intergenic		
3413071	T	C	non-synonymous	envR	Lys132Arg
3352711	G	T	non-synonymous	arcB	Asn105Lys
3330165	C	T	synonymous	dacB	Tyr401Tyr
3249169	T	C	synonymous	yqjC	Arg67Arg
3212398	T	C	non-synonymous	dnaG	Met431Thr
3205051	T	A	non-synonymous	plsY	Trp120Arg
3197632	A	G	synonymous	glnE	Gly670Gly
3192194	C	T	intergenic		
3189871	G	A	intergenic		
3155282	C	T	intergenic		
3073957	C	A	intergenic		
3031570	C	T	synonymous	uacT	Ser68Ser
3018689	T	C	non-synonymous	ygfK	Ile877Thr
2916214	A	AG	frameshift	barA	Leu388fs
2884230	T	C	intergenic		
2784674	A	G	non-synonymous	ypjC	Val113Ala
2768451	G	A	non-synonymous	yfjP	Gly248Ser
2757982	A	AT	frameshift	yfjH	Leu207fs
2757730	T	C	non-synonymous	yfjH	Thr291Ala

2712176	T	C	non-synonymous	trmN	Asn197Asp
2709467	A	G	synonymous	rpoE	Asp182Asp
2698979	T	A	non-synonymous	yfhH	Leu74Gln
2690054	T	C	non-synonymous	glrK	Thr349Ala
2598177	A	G	non-synonymous	bamC	Val230Ala
2572607	A	T	non-synonymous	eutD	Met300Lys
2543967	A	G	synonymous	ucpA	Arg219Arg
2508833	C	T	synonymous	glk	Leu198Leu
2476124	G	A	stop gained	yfdT	Trp84*
2368763	T	C	non-synonymous	arnA	Val242Ala
2360707	G	A	synonymous	rhmD	Ala236Ala
2284682	A	C	non-synonymous	yejM	Ile103Leu
2137719	T	C	intergenic		
2130550	T	C	non-synonymous	wcaD	Asn110Ser
2108903	T	C	non-synonymous	wzxB	Ser228Gly
1878882	A	G	synonymous	yoaJ	Lys3Lys
1852869	A	G	non-synonymous	ydjE	Ser371Pro
1814698	C	T	non-synonymous	katE	Arg278Cys
1803680	T	C	non-synonymous	arpB	Leu196Ser
1613671	T	C	non-synonymous	sad	Asn345Asp
1611933	C	T	intergenic		
1583403	A	G	synonymous	ydeO	Tyr95Tyr
1577667	G	T	non-synonymous	yddA	Ala559Glu
1554486	A	T	non-synonymous	maeA	Leu395Gln
1551003	CG	C	frameshift	fdnH	Ala183fs
1545608	T	C	non-synonymous	yddK	Tyr36Cys
1531486	G	T	synonymous	yncI	Ala300Ala
1524223	T	C	non-synonymous	yncE	Val306Ala
1492909	A	G	non-synonymous	trg	Tyr147Cys
1459344	A	G	synonymous	paaH	Glu97Glu
1444291	A	G	non-synonymous	ydbH	Asp414Gly
1440176	T	A,C	synonymous	ydbK	Ala203Ala
1415744	C	T	non-synonymous	recE	Ser548Asn
1398582	T	C	non-synonymous	uspE	Asn14Ser
1385439	G	A	non-synonymous	ycjX	Met441Ile
1369783	A	G	non-synonymous	pspE	Tyr32Cys
1307059	T	A	non-synonymous	yciU	Asp31Val
1295208	G	A	non-synonymous	insZ	Leu39Phe
1288160	G	A	synonymous	purU	Ala155Ala
1231716	T	A	synonymous	umuC	Ser177Ser
1192294	A	T	non-synonymous	hflD	Ile113Asn

1189403	T	C	non-synonymous	phoQ	Asp125Gly
1118261	A	G	synonymous	yceI	Asn72Asn
1036863	G	GC	frameshift	hyaF	Leu40fs
1005001	A	G	synonymous	pyrD	Gly78Gly
994065	A	G	synonymous	ssuC	Tyr256Tyr
974213	A	G	non-synonymous	smtA	Glu226Gly
963925	T	C	non-synonymous	ihfB	Met33Thr
959997	T	C	synonymous	aroA	Leu396Leu
949881	C	T	stop gained	ycaK	Gln72*
893474	G	A	intergenic		
864043	T	C	synonymous	fsaA	Arg134Arg
724138	A	G	non-synonymous	kdpD	Phe93Leu
659226	A	T	intergenic		
655916	G	A	non-synonymous	dcuC	Ala18Val
627875	C	A	non-synonymous	entB	Ala61Asp
602206	T	C	non-synonymous	pheP	Val83Ala
522807	C	T	synonymous	ybbP	Leu798Leu
506465	A	C	non-synonymous	ushA	Thr518Pro
480998	A	G	intergenic		
470464	T	C	non-synonymous	mdlA	Ser532Pro
469224	T	C	synonymous	mdlA	Ala118Ala
454086	A	G	non-synonymous	yajG	Ser28Pro
325640	C	T	synonymous	betA	Ala536Ala
320050	C	T	intergenic		
317391	T	C	intergenic		
300003	A	C	non-synonymous	paoC	Leu311Arg
280248	G	A	non-synonymous	insX	Leu39Phe
263683	T	C	intergenic		
236830	T	C	intergenic		
231010	G	A	intergenic		
164886	A	G	non-synonymous	mrcB	Lys53Glu
158495	C	T	non-synonymous	pcnB	Arg211His
112255	A	T	non-synonymous	zapD	Asp115Glu

References

[1] T. Ohtake, S. Pontrelli, W. A. Laviña, J. C. Liao, S. P. Putri, and E. Fukusaki, "Metabolomics-driven approach to solving a CoA imbalance for improved 1-butanol production in Escherichia coli," *Metab. Eng.*, vol. 41, no. April, pp. 135–143, 2017.

[2] K. Hammett, J. Kuchibhatla, C. Hunt, S. Holdread, and J. W. Brooks, "Developing Chemically Defined Media Through DOE: Complete Optimization with Increased Protein Production in Less than 8 Months," *Cell Technol. Cell Prod.*, pp. 683–691.

[3] J. Zhang and R. Greasham, "Chemically defined media for commercial fermentations," *Appl. Microbiol. Biotechnol.*, vol. 51, no. 4, pp. 407–421, 1999.

[4] "Statistical media optimization and alkaline protease production from Bacillus mojavensis in a bioreactor." .

[5] K. M. Desai, S. A. Survase, P. S. Saudagar, S. S. Lele, and R. S. Singhal, "Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan," *Biochem. Eng. J.*, vol. 41, no. 3, pp. 266–273, 2008.

[6] H. Mundhada *et al.*, "Increased production of L-serine in Escherichia coli through Adaptive Laboratory Evolution," *Metab. Eng.*, vol. 39, no. May 2016, pp. 141–150, 2017.

[7] S. Atsumi *et al.*, "Evolution, genomic analysis, and reconstruction of isobutanol tolerance in Escherichia coli.," *Mol. Syst. Biol.*, vol. 6, no. 449, p. 449, 2010.

[8] V. S. Cooper, A. F. Bennett, and R. E. Lenski, "Evolution of Thermal Dependence of Growth Rate of Escherichia Coli Populations During 20,000 Generations in a Constant Environment," *Evolution (N. Y).*, vol. 55, no. 5, pp. 889–896, 2001.

[9] S. S. Fong *et al.*, “In silico design and adaptive evolution of Escherichia coli for production of lactic acid,” *Biotechnol. Bioeng.*, vol. 91, no. 5, pp. 643–648, 2005.

[10] T. Sandberg, C. Lloyd, B. Palsson, and A. Feist, “Laboratory Evolution to Alternating Substrate Environments Yields Distinct Phenotypic and Genetic Adaptive Strategies,” *Appl. Environ. Microbiol.*, vol. 83, no. 13, pp. 1–15, 2017.

[11] C. D. Herring *et al.*, “Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale,” *Nat. Genet.*, vol. 38, no. 12, pp. 1406–1412, 2006.

[12] D. H. Lee and B. O. Palsson, “Adaptive evolution of escherichia coli K-12 MG1655 during growth on a nonnative carbon source, L-l,2-propanediol,” *Appl. Environ. Microbiol.*, vol. 76, no. 13, pp. 4158–4168, 2010.

[13] K. Jantama *et al.*, “Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate,” *Biotechnol. Bioeng.*, vol. 99, no. 5, pp. 1140–1153, 2008.

[14] P. P. Lin *et al.*, “Construction and evolution of an Escherichia coli strain solely relying on non-oxidative glycolysis for sugar catabolism,” *Proc. Natl. Acad. Sci.*

[15] S. Pontrelli *et al.*, “Metabolic Repair through Emergence of New Pathways,” *Nature*.

[16] D. Blank, L. Wolf, M. Ackermann, and O. K. Silander, “The predictability of molecular evolution during functional innovation.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 8, pp. 3044–9, 2014.

[17] K. Veeravalli, D. Boyd, B. L. Iverson, J. Beckwith, and G. Georgiou, “Laboratory evolution of glutathione biosynthesis reveals natural compensatory pathways,” *Nat. Chem.*

Biol., vol. 7, no. 2, pp. 101–105, 2011.

[18] M. Maruyama, T. Horiuchi, H. Maki, and M. Sekiguchi, "A dominant (mutD5) and a recessive (dnaQ49) mutator of Escherichia coli.," *J. Mol. Biol.*, vol. 167, no. 4, pp. 757–771, 1983.

[19] C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho, and J. C. Liao, "Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli," *Appl. Environ. Microbiol.*, vol. 77, no. 9, pp. 2905–2915, 2011.

[20] G. E. Degnen and E. C. Cox, "Conditional mutator gene in Escherichia coli: isolation, mapping, and effector studies," *J. Bacteriol.*, vol. 117, no. 2, pp. 477–487, 1974.

[21] D. Livingston, "Deoxyribonucleaic Acid Polymerase III of Escherichia coli," *J. Biol. Chem.*, vol. 250, no. 2, pp. 489–497, 1975.

[22] M. Maciag, D. Nowicki, A. Szalewska-Pałasz, and G. Wegrzyn, "Central carbon metabolism influences fidelity of DNA replication in Escherichia coli," *Mutat. Res. - Fundam. Mol. Mech. Mutagen.*, vol. 731, no. 1–2, pp. 99–106, 2012.

[23] S. J. Berríos-Rivera, G. N. Bennett, and K. Y. San, "Metabolic engineering of Escherichia coli: Increase of NADH availability by overexpressing an NAD+-dependent formate dehydrogenase," *Metab. Eng.*, vol. 4, no. 3, pp. 217–229, 2002.

[24] T. M. Conrad *et al.*, "RNA polymerase mutants found through adaptive evolution reprogram Escherichia coli for optimal growth in minimal media," *Proc. Natl. Acad. Sci.*, vol. 107, no. 47, pp. 20500–20505, 2010.

[25] S. W. Sowa *et al.*, "Integrative FourD omics approach profiles the target network of the carbon storage regulatory system," *Nucleic Acids Res.*, vol. 45, no. 4, pp. 1673–1686,

2017.

[26] M. Y. Liu *et al.*, “The RNA Molecule CsrB Binds to the Global Regulatory Protein CsrA and Antagonizes Its Activity in Escherichia coli The RNA Molecule CsrB Binds to the Global Regulatory Protein CsrA and Antagoniz,” *J. Biol. Chem.*, vol. 272, no. 28, pp. 17502–17510, 1997.

[27] R. G. Chavez, A. F. Alvarez, T. Romeo, and D. Georgellis, “The physiological stimulus for the BarA sensor kinase,” *J. Bacteriol.*, vol. 192, no. 7, pp. 2009–2012, 2010.

[28] K. A. Salmon *et al.*, “Global Gene Expression Profiling in *Escherichia coli* K12,” *J. Biol. Chem.*, vol. 280, no. 15, pp. 15084–15096, 2005.

[29] A. Sirko, E. V. a Zehelein, M. Freundlich, and G. Sawers, “Integration Host Factor Is Required for Anaerobic Pyruvate Induction of pfl Operon Expression in Escherichia coli,” *J. Bacteriol.*, vol. 175, no. 18, pp. 5769–5777, 1993.

[30] S. M. Arfin *et al.*, “Global gene expression profiling in Escherichia coli K12: The effects of integration host factor,” *J. Biol. Chem.*, vol. 275, no. 38, pp. 29672–29684, 2000.

[31] S. Iuchi and L. Weiner, “Cellular and molecular physiology of Escherichia coli in the adaptation to aerobic environments.,” *J. Biochem.*, vol. 120, no. 6, pp. 1055–1063, 1996.

[32] S. Alexeeva, K. J. Hellingwerf, and M. J. Teixeira de Mattos, “Requirement of ArcA for redox regulation in Escherichia coli under microaerobic but not anaerobic or aerobic conditions,” *J Bacteriol*, vol. 185, no. 1, pp. 204–209, 2003.

[33] S. Iuchi and E. C. C. Lin, “Signal transduction in the Arc system for control of operons encoding aerobic respiratory enzymes,” *Two-component signal Transduct.*, pp. 223–232, 1995.

[34] M. L. Mott and J. M. Berger, “DNA replication initiation: Mechanisms and regulation in bacteria,” *Nat. Rev. Microbiol.*, vol. 5, no. 5, pp. 343–354, 2007.

[35] and C. M. T. Espinosa, M., S. Cohen, M. Couturier, G. Del Solar, R. Diaz-Orejas, R. Giraldo, L. Janniere, C. Miller, M. Osborn, *Plasmid replication and copy number control." The horizontal gene pool: bacterial plasmids and gene spread*. 2000.

[36] M. Masters, M. D. Colloms, I. A. N. R. Oliver, L. I. N. He, E. J. Macnaughton, and Y. Charters, “The pcnB Gene of Escherichia coli , Which Is Required for ColEl Copy Number Maintenance , Is Dispensable,” vol. 175, no. 14, pp. 4405–4413, 1993.

[37] H. Tao *et al.*, “Functional Genomics : Expression Analysis of Escherichia coli Growing on Minimal and Rich Media Functional Genomics : Expression Analysis of Escherichia coli Growing on Minimal and Rich Media,” *J. Bacteriol.*, vol. 181, no. 20, p. 6425, 1999.

[38] J. S. Swedes, R. J. Sedo, and D. E. Atkinson, “Relation of growth and protein synthesis to the adenylate energy charge in an adenine-requiring mutant of Escherichia coli,” *J. Biol. Chem.*, vol. 250, no. 17, pp. 6930–6938, 1975.

[39] S. Atsumi *et al.*, “Metabolic engineering of Escherichia coli for 1-butanol production.,” *Metab. Eng.*, vol. 10, no. 6, pp. 305–11, Nov. 2008.

[40] A. Espah Borujeni, A. S. Channarasappa, and H. M. Salis, “Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites,” *Nucleic Acids Res.*, vol. 42, no. 4, pp. 2646–2659, 2014.

[41] K. a Datsenko and B. L. Wanner, “One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 12, pp. 6640–6645, 2000.

[42] D. G. Wernick, S. P. Pontrelli, A. W. Pollock, and J. C. Liao, “Sustainable biorefining in wastewater by engineered extreme alkaliphile Bacillus marmarensis,” *Sci. Rep.*, vol. 6, p. 20224, 2016.

[43] Y.-X. Huo *et al.*, “Conversion of proteins into biofuels by engineering nitrogen flux.,” *Nat. Biotechnol.*, vol. 29, no. 4, pp. 346–51, Apr. 2011.

[44] A. McKenna *et al.*, “The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data,” *Proc. Int. Conf. Intellect. Capital, Knowl. Manag. Organ. Learn.*, vol. 20, pp. 254–260, 2009.

[45] P. Cingolani *et al.*, “A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3,” *Fly (Austin).*, vol. 6, no. 2, pp. 80–92, 2012.

[46] M. Liu, T. Durfee, J. E. Cabrera, K. Zhao, D. J. Jin, and F. R. Blattner, “Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli,” *J. Biol. Chem.*, vol. 280, no. 16, pp. 15921–15927, 2005.