Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in Escherichia coli

Long, Christopher P.; Gonzalez, Jacqueline E.; Feist, Adam M.; Palsson, Bernhard O.; Antoniewicz, Maciek R.

Published in:
Proceedings of the National Academy of Sciences of the United States of America

Link to article, DOI:
10.1073/pnas.1716056115

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in Escherichia coli

Christopher P. Long, Jacqueline E. Gonzalez, Adam M. Feist, Bernhard O. Palsson, and Maciek R. Antoniewicz

*Department of Chemical and Biomolecular Engineering, Metabolic Engineering and Systems Biology Laboratory, University of Delaware, Newark, DE 19716; †Department of Bioengineering, University of California, San Diego, La Jolla, CA 92039; and ‡Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark

Edited by Arnold L. Demain, Drew University, Madison, NJ, and approved November 6, 2017 (received for review September 12, 2017)

Unraveling the mechanisms of microbial adaptive evolution following genetic or environmental challenges is of fundamental interest in biological science and engineering. When the challenge is the loss of a metabolic enzyme, adaptive responses can also shed significant insight into metabolic robustness, regulation, and areas of kinetic limitation. In this study, whole-genome sequencing and high-resolution 13C-metabolic flux analysis were performed on 10 adaptively evolved pgiknockouts of Escherichia coli. Pgi catalyzes the first reaction in glycolysis, and its loss results in major physiological and carbon catabolism pathway changes, including an 80% reduction in growth rate. Following adaptive laboratory evolution (ALE), the knockouts increase their growth rate by up to 3.6-fold. Through combined genomic–fluxomic analysis, we characterized the mutations and resulting metabolic fluxes that enabled this fitness recovery. Large increases in pyridine cofactor transhydrogenase flux, correcting imbalanced production of NADPH and NADH, were enabled by direct mutations to the transhydrogenase genes nthA and nthB. The phosphotransferase system component crr was also found to be frequently mutated, which corresponded to elevated flux from pyruvate to phosphoenolpyruvate. The overall energy metabolism was found to be strikingly robust, and what has been previously described as latently activated Entner–Doudoroff and glyoxylate shunt pathways are shown here to represent no real increases in absolute flux relative to the wild type. These results indicate that the dominant mechanism of adaptation was to relieve the rate-limiting steps in cofactor metabolism and substrate uptake and to modulate global transcriptional regulation from stress response to catabolism.

Phosphoglucose isomerase (pgi) knockouts of Escherichia coli are of significant interest in metabolic engineering and have been the subject of many investigations (1). Pgi catalyzes the first reaction in glycolysis, the conversion of glucose 6-phosphate (G6P) to fructose 6-phosphate (F6P), which in the wild type during aerobic growth on glucose catabolizes ~70% of glucose (18, 19). Its loss results in a correspondingly severe growth impairment (70–80% lower growth rate) (13, 20) as the oxidative pentose phosphate pathway (oxPPP) and Entner–Doudoroff (ED) pathway must compensate. Several studies have used 13C-metabolic flux analysis (C-MFA) to characterize Δpgi, frequently describing the activation of normally latent (i.e., nonutilized) pathways and a redox imbalance caused by overproduction of NADPH in the pentose phosphate pathway (2, 21–24). The major flux, redox, and growth rate changes caused by loss of pgi make it a rich target for ALE experiments (13, 14). Previously, Charusanti et al. (14) adaptively evolved 10 strains in replicate experiments over 50 d of continuous culture in glucose minimal media, reporting significant growth recovery of 3.6-fold. However, no underlying intracellular fluxes have been reported for these strains or any similarly large-scale ALE study of genetic mutants.

To gain fundamental insight into the mechanisms and outcomes of adaptive evolution, both the mutations and the selected-upon phenotype (here, metabolism) must be measured. In this study, we applied high-resolution 13C-MFA and next-generation sequencing to the 10 evolved Δpgi strains and the parental strain reported previously (14). Novel mutations were identified in the phosphotransferase system component crr. Overcoming key bottlenecks, rather than a broad metabolic response, is the dominant mechanism of adaptation.

Significance

Understanding how microbes adapt to changing conditions is fundamental to biological science and engineering. For example, adaptation is a key driver of antimicrobial resistance, and adaptive laboratory evolution has become a key tool in biotechnology. Here, we present a comprehensive genetic and fluxomic analysis of 10 adaptively evolved Escherichia coli phosphoglucose isomerase (pgi) knockout strains. The loss of PGI, a key enzyme in glycolysis, results in massive redirection of carbon catabolic flux and reduction in growth rate. Adaptive evolution results in a 3.6-fold increase in growth rate, enabled by key mutations and metabolic flux rewiring. These include global transcriptional regulators, cofactor transhydrogenases, and the phosphotransferase system component crr. Overcoming key bottlenecks, rather than a broad metabolic response, is the dominant mechanism of adaptation.

Author contributions: C.P.L. and M.R.A. designed research; C.P.L. and A.M.F. performed research; J.E.G. contributed new reagents/analytic tools; C.P.L., J.E.G., A.M.F., B.O.P., and M.R.A. analyzed data; and C.P.L. and M.R.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

To whom correspondence should be addressed. Email: mranton@udel.edu.

1To whom correspondence should be addressed. Email: mranton@udel.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1716056115/-/DCSupplemental.
identified, and comparisons to recently reported wild-type ALE studies (5, 6) helped to identify mutations unique to Δpgi. Together with complimentary fluxomic information, a detailed picture of how Δpgi metabolically adapts to achieve faster growth is attained. Areas of convergence and divergence on the genetic and fluxomic levels highlight the large number of genetic solutions possible for achieving similar metabolic phenotypes as well as some differences in metabolic optima. In several cases, specific causal mutation-flux relationships were identified.

Results

Recovery of Growth Fitness in Δpgi ALE Strains Is Associated with Unique Mutations. In E. coli K-12 MG1655, the knockout of pgi results in a reduction in growth rate of ~80% compared with the wild type, from 0.72 h⁻¹ to 0.14 h⁻¹. Following ALE, a significant fraction of this growth rate can be recovered (46–71% of the wild-type growth rate, Fig. L4). These 2.4–3.6-fold increases in growth rate are quite large compared with, for example, similar ALE experiments with wild-type E. coli, which reported 1.6-fold increases in growth rate (5).

To assess the genetic basis of the large increases in growth rate, whole-genome sequencing was performed and sequences were mapped to the E. coli reference to identify mutations in 10 independent ALE experiments. Recent advances in sequencing allowed for improved determination of mutations in clones isolated from final populations, particularly of insertion sequences (IS elements), compared with a previous effort that utilized both a microarray hybridization-based method and an earlier Illumina short-read technology protocol (14). In all, 52 unique mutations were identified across the 10 ALE strains, spanning 34 different genetic regions (29 structural mutations—e.g., within an ORF—and 5 intergenic regions). The complete mutation table is in Supporting Information. A key advantage and reason for using replicate ALE experiments is to use mutation frequency to help differentiate causal mutations from genetic "hitch-hikers" that do not affect fitness (5). Fig. 1B lists the genes that were mutated in ≥2 of the 10 Δpgi ALE strains as well as how the genes frequently mutated in reported wild-type K-12 MG1655 ALE studies performed in similar conditions (5, 6). The two sets are striking in their lack of overlap; mutations that occur in almost every reported wild-type ALE experiment, in rpoB, pyrE/rph, and hnsIdk, occur rarely or not at all in Δpgi strains. Instead, the Δpgi ALE strains have a high frequency of mutations in the pyridine nucleotide transhydrogenases pntAB and shkA, the transcription factor rpoB, and the phospho-transporter system (PTS) sugar transport system component ccr. The distribution of these mutations across the 10 Δpgi ALE strains is also in Fig. 1A, showing that while some strains had many of these common mutations (ALE-2 had six out of the top seven), others had fewer (ALE-8 had only 1/7). This likely reflects less common but equally effective adaptive routes. It is worthwhile to mention that some genes or genetic regions had many unique mutations in parallel evolutions (e.g., six for rpoB and five for pntAB). In contrast, the IS element insertion in ccr was identical in five different strains. Both patterns clearly demonstrate evidence for causality. It was previously demonstrated that the combination of rpoB and shkA mutations are causal for increased growth in Δpgi and exhibit positive epistasis (14). The difference in mutation profiles demonstrate that there are unique selective pressures in Δpgi, which result in unique adaptive responses. To further investigate how these mutations enabled the large increases in growth rates from the initial perturbed Δpgi state, we next characterized the carbon metabolism of each strain using 13C-MFA.

Activation of Latent Pathways, or Not? To characterize the rewiring of central carbon metabolism in the parental and ALE Δpgi strains, high-resolution 13C-MFA was performed. The analysis consisted of two parallel labeling experiments with [1,2,13C] and [1,6,13C]glucose [an experimental design previously identified as providing optimal flux estimate precision (25)], and the simultaneous fitting of labeling from proteinogenic amino acids, the ribose moiety of RNA, and glucose moiety of glycogen (26) to estimate fluxes. For the wild type, data from parallel labeling experiments previously reported were refitted (19). The full network model, the measured isotopomer distributions, and the estimated metabolic fluxes are in Supporting Information.

In Fig. 2 A–C, the estimated intracellular fluxes of the wild-type, Δpgi parental strain, and Δpgi ALE-3 are summarized. ALE-3 was the fastest growing Δpgi strain (0.51 h⁻¹) and had an intracellular flux distribution typical of most of the ALE strains. The growth rates and glucose uptake rates for each strain are noted, and the fluxes shown are normalized to 100 units of glucose uptake. In all Δpgi strains, the forward and reverse fluxes of the PGI reaction were estimated to be zero, thus confirming the pgi knockout. The unevolved Δpgi (Fig. 2B) was found to utilize reactions and pathways that carry minimal flux in the wild type (Fig. 2A), including the ED pathway, glyoxylate shunt, and phosphoenolpyruvate carboxykinase (PCK) reaction [oxaloacetate to phosphoenolpyruvate (PEP)]. These flux changes have been noted in previous studies and have been described as “latent pathway activation” (2, 22–24, 27). After adaptive evolution, ALE-3 (Fig. 2C) and the other ALE strains significantly reduced the usage of these pathways; for example, the glyoxylate shunt flux was reduced from 25 to 6, and the PCK reaction from 22 to 3 in ALE-3. This “rerepression” following adaptive evolution has also been previously observed and been the focus of various speculations and computational analyses (13, 16). The previous terminology implies the presence of a transient regulatory response, activated in response to the stress caused by the knockout and then repressed during evolution to facilitate faster growth.

When interpreting metabolic fluxes, it is important to consider both normalized fluxes (e.g., relative to glucose uptake) and absolute fluxes (mmol·gDW⁻¹·h⁻¹), as each provide complimentary information. Several key fluxes are shown in both units in Fig. 2D, corresponding to the pathway map shown in Fig. 2E. The glucose uptake rates used to calculate absolute fluxes are in Fig. S1. In Fig. 2B and D, we can see that although the oxPPP is the dominant route of glucose catabolism in Δpgi, the absolute flux is reduced by roughly half in the unevolved strain relative to the wild type. It has been previously reported that G6P accumulates in Δpgi and that G6PDH (encoded zwf), the first step in the oxPPP, is likely rate limiting for growth due to allosteric inhibition caused by an elevated NADPH/NADP⁺ ratio (23).
Energy Metabolism Is Not Significantly Affected by Adaptive Evolution in \( \Delta pgi \). The measured metabolic fluxes can also inform a broader analysis of energy metabolism in these strains. In Fig. 3 A and B, oxygen uptake rates and acetate yields are shown. In the unevolved \( \Delta pgi \), the oxygen uptake is reduced to 4.3 mmol \( \text{g}_{DW}^{-1} \cdot \text{h}^{-1} \), down from 15 mmol \( \text{g}_{DW}^{-1} \cdot \text{h}^{-1} \) in the wild type, corresponding to the overall slowed metabolism and growth rate. The unevolved \( \Delta pgi \) does not produce acetate, as the citrate synthase (CS) flux can easily accommodate all of the flux from acetyl-CoA at less than half of its wild-type rate (Fig. 2D). In the ALE strains, oxygen uptake recovers to 62–91% of the wild-type flux and some strains produce acetate. This may represent a limit in TCA cycle or oxidative phosphorylation capacity that these strains encounter, above which excess glycolytic flux is diverted to acetate production. This acetate overflow effect is shown in Fig. 3C, which shows that the absolute flux through the pyruvate (Pyr) dehydrogenase (acetyl-CoA generation) strongly correlates with the acetate secretion flux in all strains. Fig. 3D and E show the normalized cofactor balances for ATP and NADH/FADH2 (the electron carriers used in oxidative phosphorylation for ATP production), with contributions to production and consumption by the various pathways and cell functions. One noticeable difference in the unevolved \( \Delta pgi \) strain is an increased contribution of the TCA cycle and oxidative phosphorylation to energy metabolism, leading to a slightly higher overall ATP yield. The ALE strains mostly reverted to normalized levels of total cofactor production and consumption that were very similar to the wild type. Overall, the profile of energy metabolism is remarkably
conserved in these strains. Perhaps the most significant difference between the wild-type and Δpga strains in cofactor metabolism is in the reversed role of transhydrogenase (Fig. 3D).

**Transhydrogenase Genes Are Mutated in Many but Not All Δpga ALE Strains.** *E. coli* is able to interconvert reduced cofactors with two pyridine nucleotide transhydrogenases, the membrane-bound PntAB, which primarily converts NADH to NADPH, and the soluble form SthA (also referred to as UdhA), which primarily converts NADPH to NADH (28). In the wild type, excess NADH produced in glycolysis and the TCA cycle is used to produce approximately half of the needed NADPH through the transhydrogenase (Figs. 3D and 4A and Fig. S2). In Δpga, the redirection of glucose flux in upper central carbon metabolism, away from glycolysis and into the oxPPP, results in a significant excess of NADPH generation. Some of this is utilized by elevated biosynthesis flux (i.e., biomass yield) (Fig. S1), but most of the imbalance must be corrected by a reversal of the transhydrogenase to convert NADPH to NADH.

The absolute rates of NADPH production and consumption are shown in Fig. 4A. Here again is shown the impact of the forcing of flux through the oxPPP, which generates a large excess of NADPH and necessitates the reversal of the transhydrogenase. This transhydrogenase flux is shown in Fig. 4B in both absolute and relative flux units with the 95% confidence intervals calculated from 13C-MFA. In the Δpga strains (both unevolved and evolved), the normalized flux does not change significantly (also Fig. S2), reflecting that other parts of central carbon metabolism were not re-wired to relieve the cofactor imbalance. Instead, the absolute transhydrogenase fluxes increase significantly in the ALE strains.

To gain insight into how these flux increases were achieved, the mutations directly related to the regions of the transhydrogenase genes in the ALE strains were compared with the flux changes (Fig. 4C). As noted above (Fig. 1), pntA, pntB, and sthA were some of the most frequently mutated genes in this study. Eight of the 10 ALE strains had at least one transhydrogenase mutation, with five having two. Based on the nature of the mutations, they presumably increase SthA activity, reduce PntAB activity, or both. The pntAB mutations were diverse, including deletions and a duplication, as well as an upstream IS element insertion. Three of the four mutations in the coding regions likely result in truncated, nonfunctional proteins: ALE-2 has a nonsense mutation in pntA, and ALE-4 and ALE-7 have major truncations in pntB and pntA, respectively (14). The sthA mutations were SNPs, including a commonly mutated site (five strains) 64 bp upstream. The exact effect of the upstream mutations on transcriptional regulation is uncertain, but we hypothesize that they increase sthA and reduce pntAB expression. Despite the high frequency of these mutations, two strains (ALE-5 and ALE-6) achieved the increased flux with no observed mutations directly in the transhydrogenase genes, raising questions about other possible mechanisms for cofactor rebalancing. Both ALE-5 and ALE-6 possess mutations in genes that directly affect transcription levels (e.g., bp and rpoA). In fact, pntAB has been identified as a regulatory target of RpoA (29).

**Mutations in PTS Component crr Are Associated with Increased Back Flux from Pyr to PEP.** Another frequently occurring, and more unexpected, mutation was an IS element insertion in crr in 5 of 10 ALE strains (Fig. 5). Crr encodes EIIA-GCK, the cytosolic subunit of enzyme II (EII) in the PTS glucose transport system. The PTS system is the primary mode of glucose uptake in *E. coli* and links the uptake and phosphorylation of glucose (by EI) to the glycolytic conversion of PEP to Pyr (by EI, linked by the intermediary phosphotransferase HPt) (30, 31) (Fig. 5A). Previous work has shown that EI of the PTS sugar transport system is reversible in vivo (32) and that in the wild type 10% of PEP is produced from Pyr in the reverse (i.e., gluconeogenic) direction by this mechanism. Previous work has also shown that in a Δcrr strain, this flux is increased by more than twofold due to perturbation of the PTS equilibrium as well as some possible activation of PPS, the gluconeogenic enzyme also able to facilitate the conversion of Pyr to PEP. Given the prevalence of the crr mutation and global metabolic perturbations in Δpga, we hypothesized that the PPS/EI flux (Pyr to PEP) would be altered in these ALE strains.

The PPS/EI flux was measured using an [U-13C]alanine tracer approach developed recently (32) (see Materials and Methods and Fig. S3). As shown in Fig. 5B, the extent of this flux varied widely among the ALE strains, and its magnitude corresponded strongly with the presence of the crr mutation. Expressed as normalized flux (relative to 100 units of glucose uptake), the PPS/EI flux was significantly elevated in the unevolved Δpga (from 18 in the wild type to 47) and was reduced subsequently in strains lacking the crr mutation but maintained at a high level in the strains with the mutated crr. In absolute terms, the flux was significantly reduced in all Δpga strains except for those with the crr mutation, where the flux was more similar to the wild type. Without further analysis of the activity of PPS in these strains, it is difficult to deduce the exact mechanism of these changes, but they are consistent with the previous observations in Δcrr (32). This result strongly supports a genetic–metabolic flux relationship between the crr IS element mutation and elevated PPS/EI flux.

The high frequency of this identical mutation indicates a strong
selective pressure for this mutation, but the exact mechanism of the fitness benefit is unclear. In addition to its direct role in glucose uptake, crr is also involved in signaling the global regulator Crp (33), which controls the transcription of over 100 genes (34).

Discussion

In this work, we have explored how recovery of growth rate in Δpgi is enabled by unique genetic mutations and significant metabolic rewiring. In this system, fitness recovery was driven by global transcriptional regulation (i.e., sigma factors and other RNA polymerase components) and relief of a rate-limiting step at the cofactor transhydrogenase. This led to increased absolute flux through the oxidative pentose phosphate pathway and a corresponding recovery of faster glucose uptake and growth rates. The usage of latent pathways including the ED pathway, glyoxylate shunt, and PCK reaction were shown to represent no increase in absolute flux relative to the wild type and in absolute terms did not appreciably change after adaptive evolution. The availability of these pathways, expressed at low levels in the wild type, may offer flexibility when facing changing conditions. However, they do not appear to be “activated” as part of a general stress response in this case. These insights and the dataset presented here should help advance predictive metabolic modeling (35–37). Overall, these results add to our understanding of adaptive evolution by elucidating how challenges to specific cellular subsystems—that is, central carbon metabolism and glycolysis—are overcome. Future application of this approach to other significant, growth-limiting mechanisms where the selected-for mutations depend on the state of the cell (in this case, perturbed Δpgi) versus a wild-type state, other mutations seen even more when starting with a wild-type strain, such as in pyrE/rdh, are likely relevant based on the overall growth state (the evolved strains started with wild type are significantly faster when evolved under the same conditions); thus, they may never be selected for in a “crippled” starting strain such as the Δpgi starting strain (42). However, such mutations and a convergence of mutations may be seen if evolved for more extended lengths of time under the same conditions (43).

Other unique mutations in Δpgi were in the transhydrogenases and in Δpgi. With regards to the former, a recent report of an ALE study of an oxPPP mutant (i.e., reduced NADPH production) found a high frequency of mutations in pmnt/AB that lead to increased activity (44). Given the importance of cofactors to the formation of metabolic products, the reported mutations from these two studies may provide valuable new candidates for rationally manipulating transhydrogenase activity in metabolic engineering (45). Lastly, the mechanism of fitness enhancement of the crr IS element mutation identified here requires further study. One possibility is that the reduced glycolytic flux in Δpgi could be sensed (46) by the PTS system—for example, via perturbation of the PEP/Pyr concentration ratio—and result in feedback inhibition of glucose uptake. Another is that the accumulated G6P (23) could induce a maladaptive regulatory response via CRP activation by P ~ EIIB\^5c. In these scenarios, a crr mutation may help to decouple feedback inhibition or limit the harmful regulatory effect. Ultimately, it will be desirable to confirm the identities and interactions of causal mutations by reproducing them synthetically. Previous work (14, 21) confirmed causality and epistasis for shhA and nps mutations but did not fully recapitulate the observed growth phenotypes of the evolved clones themselves. This likely points to a complex landscape that will become more feasible to explore as high-throughput genome engineering methods mature and many strains can be tested efficiently.

Materials and Methods

DNA Sequencing. Sequences were obtained using Illumina MiSeq. The breseq pipeline (47) was used to map sequenced reads and identify mutations relative to the E. coli K-12 MG1655 genome.

Tracer Experiments. For \(^13\)C-tracer experiments, strains were cultured aerobically in batch culture in M9 minimal medium at 37 °C in minibioreactors with 10 mL working volume (20). For the quantification of the Pyr to PEP flux, \(^13\)C-alanine tracer experiments were performed (32).

Gas Chromatography–Mass Spectrometry. Gas chromatography–mass spectrometry (GC-MS) analysis was performed on an Agilent 5977A mass spectrometer to measure labeling of proteinogenic amino acids (48, 49), glucose (derived from glycogen), and ribose (from RNA) (26, 50, 51). Mass isotopomer distributions (MIDs) were obtained by integration and corrected for natural isotope abundances.

\(^13\)C-MFA. \(^13\)C-MFA calculations were performed using the Metran software. For integrated analysis of parallel labeling experiments, the datasets were fitted simultaneously to a single flux model (18).

Data Availability. All data generated or analyzed during this study are included in this published article (and its Supporting Information files).
ACKNOWLEDGMENTS. This work was supported by NSF Grant MCB-1616332. C.P.L. and J.E.G. were also supported by the University of Delaware Graduate Fellows Award. A.M.F. and B.O.P. were supported under Novo Nordisk Foundation Center for Biosustainability Grant NNF10CC101617.


