UC San Diego

UC San Diego Previously Published Works

Title

Growth Adaptation of gnd and sdhCB Escherichia coli Deletion Strains Diverges From a Similar Initial Perturbation of the Transcriptome

Permalink

https://escholarship.org/uc/item/95b5m5hn

Authors

McCloskey, Douglas
Xu, Sibei
Sandberg, Troy E
et al.

Publication Date

2018

DOI

10.3389/fmicb.2018.01793

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

frontiers in Microbiology

ORIGINAL RESEARCH
published: 07 August 2018
doi: 10.3389/fmicb.2018.01793

Growth Adaptation of *gnd* and *sdhCB Escherichia coli* Deletion Strains Diverges From a Similar Initial Perturbation of the Transcriptome

Douglas McCloskey[1,2], *Sibei Xu*[1], *Troy E. Sandberg*[1], *Elizabeth Brunk*[1], *Ying Hefner*[1], *Richard Szubin*[1], *Adam M. Feist*[1,2] *and Bernhard O. Palsson*[1,2]*

[1] *Department of Bioengineering, University of California, San Diego, San Diego, CA, United States,* [2] *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark*

Adaptive laboratory evolution (ALE) has emerged as a new approach with which to pursue fundamental biological inquiries and, in particular, new insights into the systemic function of a gene product. Two *E. coli* knockout strains were constructed: one that blocked the Pentose Phosphate Pathway (*gnd* KO) and one that decoupled the TCA cycle from electron transport (*sdhCDAB* KO). Despite major perturbations in central metabolism, minimal growth rate changes were found in the two knockout strains. More surprisingly, many similarities were found in their initial transcriptomic states that could be traced to similarly perturbed metabolites despite the differences in the network location of the gene perturbations and concomitant re-routing of pathway fluxes around these perturbations. However, following ALE, distinct metabolomic and transcriptomic states were realized. These included divergent flux and gene expression profiles in the *gnd* and *sdhCDAB* KOs to overcome imbalances in NADPH production and nitrogen/sulfur assimilation, respectively, that were not obvious limitations of growth in the unevolved knockouts. Therefore, this work demonstrates that ALE provides a productive approach to reveal novel insights of gene function at a systems level that cannot be found by observing the fresh knockout alone.

Keywords: adaptive laboratory evolution, sdhD gene knockouts, systems biology, casual mutations, gnd, sdhA, sdhB, sdhC

OPEN ACCESS

Edited by:
Sergey M. Stolyar, University of Idaho, United States

Reviewed by:
Christopher L. Hemme, University of Rhode Island, United States
Yinjie Tang, Washington University in St. Louis, United States

***Correspondence:**
Bernhard O. Palsson
bpalsson@ucsd.edu

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *21 May 2018*
Accepted: *17 July 2018*
Published: *07 August 2018*

Citation:
McCloskey D, Xu S, Sandberg TE, Brunk E, Hefner Y, Szubin R, Feist AM and Palsson BO (2018) Growth Adaptation of gnd and sdhCB Escherichia coli Deletion Strains Diverges From a Similar Initial Perturbation of the Transcriptome. Front. Microbiol. 9:1793. doi: 10.3389/fmicb.2018.01793

INTRODUCTION

Biological systems contain many overlapping metabolic and regulatory networks that contribute to robustness against perturbations. Systems robustness to perturbation is due to regulatory adjustments that can coordinate re-routing of flux by adjusting enzyme levels to compensate for metabolic pathway disruption (Zhao et al., 2004; Ishii et al., 2007; Nicolas et al., 2007; Nakahigashi et al., 2009). A systems level understanding of how regulatory adjustments are made can be gained by knocking out a gene product and studying the downstream metabolic and regulatory shifts that occur. Genes encode two major metabolic reactions were removed from a pre-evolved strain of *E. coli* K-12 MG1655. The two reactions were GND (*gnd*, 6-phosphogluconate dehydrogenase) and SUCDi (genes *sdhA*, *sdhB*, *sdhC*, and *sdhD* corresponding to the enzyme Succinate Dehydrogenase). GND carries out the final NADPH generating step of the oxidative

Pentose Phosphate Pathway (oxPPP) to produce D-ribulose-5-phosphate (ru5p-D), which is a necessary precursor for nucleotide biosynthesis. SUCDi converts succinate (succ) to fumarate (fum) in the TCA cycle and also directly charges and donates quinones to the electron transport chain (ETC) via Complex II, thus directly coupling the TCA cycle to respiration.

Only minimal changes in growth rate were found when disabling the GND and SUCDi reactions. It has been shown previously that disrupting the *gnd* gene resulted in non-appreciable changes to growth rate, but induces major changes in fluxes through the PPP and TCA cycle (Zhao et al., 2004; Ishii et al., 2007; Nakahigashi et al., 2009). Similar observations have also been made of *zwf* mutants (Nicolas et al., 2007; Olavarria et al., 2014), which demonstrate that metabolic networks can rapidly adjust enzyme level and flux without an appreciable cost to growth rate. The same ability to re-route fluxes has not been demonstrated for disruptions to the TCA cycle as has been done for the PPP. Importantly, the question of whether or not the immediate regulatory and flux shifts were the most optimal has not been explored.

The adaptive response to gene loss can be studied using adaptive laboratory evolution (ALE). ALE is an experimental method that introduces a selection pressure (e.g., growth rate selection) in a controlled environmental setting (Dragosits and Mattanovich, 2013; Plucain et al., 2014; Tenaillon et al., 2016). Using ALE, organisms can be perturbed from their evolutionary optimized homeostatic states, and their re-adjustments can be studied during the course of adaptation to reveal novel and non-intuitive gene product functions and interactions (Chou et al., 2015). When applied to KO strains in *E. coli*, it has been shown that the growth rate loss from the removal of key metabolic genes can be overcome through the selection of beneficial mutations (Ibarra et al., 2002; Fong et al., 2005, 2006). The rate of accumulation of compensatory adaptation is often associated with the amount of initial growth rate lost (Moore et al., 2000). The relationship between rate of compensatory adaptation and relative growth rate change would imply that little to no compensatory adaptation would be expected to occur when the environmental or genetic change resulted in minimal growth rate loss. The relationship between rate of compensatory adaptation and relative growth rate change would also imply that little to no changes in the regulatory or metabolic network would be found post-evolution.

In this study, the evolutionary drivers in the absence of significant growth rate loss were explored. Starting from a pre-optimized *E. coli* K-12 MG1655 strain, two major metabolic perturbations were inflicted that blocked the oxPPP and decoupled the TCA cycle from the ETC **(Figure 1)**. Minimal changes in growth rate were found in the KO strains and evolved KO endpoints. Detailed omics analysis demonstrated that despite minimal changes in growth rate, massive changes in metabolic flux, gene expression, and metabolite levels occurred in both KO and evolved strains (**Figures 1–8**, **Figures S1–S2, Tables S1–S8**). Interestingly, many of the changes found in the regulatory network were shared by both knockout strains despite major differences in the location of the metabolic perturbation. Investigation of mutations in the evolved end points indicated that in the absence of a significant change in growth rate, selection pressures existed that resulted in major adaptations in regulatory and metabolic networks. These mutations led to divergent adjustments at the regulatory and metabolic networks in both knockout strains. These results demonstrated even when genetic perturbations induce little cost to growth rate, major optimizations to regulatory and metabolic network structure can be found after evolution.

RESULTS

Strain Selection: Starting With a Growth Optimized Strain

In order to eliminate any confounding effects between the adaptation to the growth conditions used in the experiment and the loss of a gene product, we chose a wild-type *E. coli* K-12 MG1655 strain previously evolved under glucose minimal media at 37°C (LaCroix et al., 2015) (denoted as "Ref"). Based on this rationale, this reference strain has been used as the basis for several ALE studies (Sandberg et al., 2014, 2016).

Blockage of the Oxidative PPP and Decoupling of the TCA Cycle From the ETC Resulted in Minimal Fitness Loss

GND (*gnd*, 6-phosphogluconate dehydrogenase) and SUCDi (genes *sdhA*, *sdhB*, *sdhC*, and *sdhD* corresponding to the enzyme Succinate Dehydrogenase) were removed from Ref to generate strain uGnd and uSdhCB (denoted "unevolved *gnd* and *sdhCB* knockout strain"). The initial growth rate of uGnd and uSdhCB were minimally changed (9 and 6% decrease in growth rate, respectively) (**Figure 1D**). Three uGnd and Three uSdhCB stains from independently inoculated starting cultures were simultaneously evolved on glucose minimal media at 37°C in an automated ALE platform (Sandberg et al., 2014; LaCroix et al., 2015) denoted "evolved *gnd* and *sdhCB* knockout strains" or "eGnd and eSdhCB." A non-significant and minimal increase in final growth rate was found in all endpoints of the eSdhCB and eGnd lineages (ave $\pm$ stdev 3 $\pm$ 4, 5 $\pm$ 3% increase in growth rate) compared to the unevolved knockout strains (uSdhCB and uGnd lineages), respectively (Student's *t*-test, *p*-value < 0.05). In addition, no significant changes in glucose uptake, and acetate secretion rate were found in uGnd and eGnd strains (**Figure 1D**, **Table S2**). Two of the three eSdhCB strains were found to have a statistically significant increase in acetate secretion, and a statistically significant decrease in succinate secretion rates compared to uSdhCB (**Figure 1D**, **Table S2**).

Genome-scale models were used to compute the flux map in Ref. The results indicated that GND and SUCDi represent two of the highest flux reactions when Ref was grown on glucose minimal media (data not shown). The minimal changes in growth rate after the loss of these two reactions were surprising given that massive flux rerouting would have to take place in each of the uKO strains. Detailed -omics analysis of all uGnd, eGnd, uSdhCB, and eSdhCB strains were carried out to better

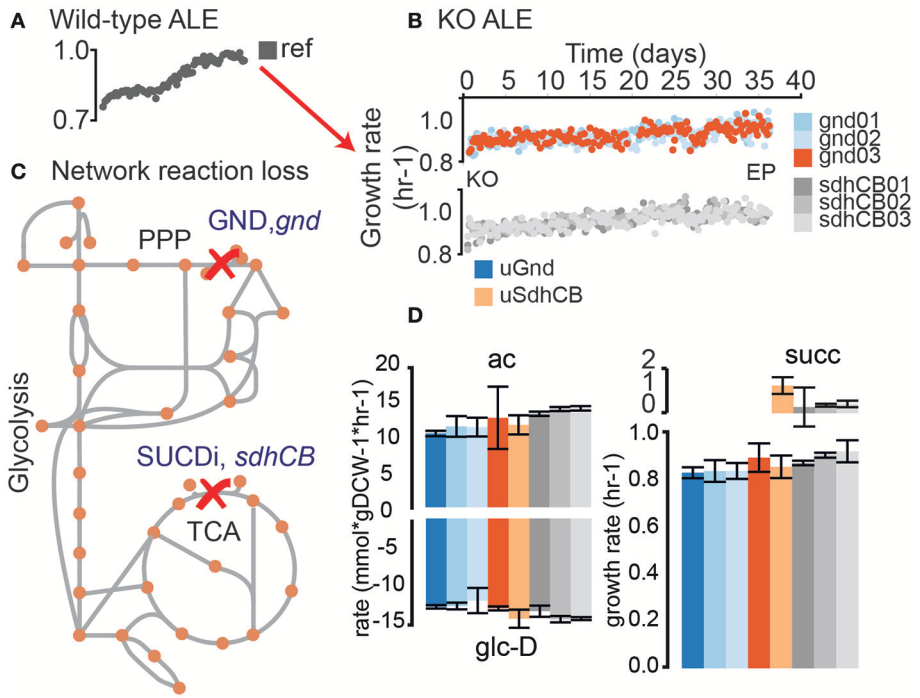

FIGURE 1 | Evolution of knockout (KO) strains from a pre-evolved (i.e., optimized) wild-type strain. **(A)** Wild-type (wt) *E. coli* (MG1655 K-12) was previously evolved on glucose minimal media at 37°C (LaCroix et al., 2015). An isolate from the endpoint of the evolutionary experiment was selected as the starting strain for subsequent KOs of *gnd* and *sdhCB* and adaptive laboratory evolution (ALE). **(B)** Adaptive laboratory evolution trajectories of the evolved knockout lineages. -Omics data collected included metabolomics, fluxomics, physiology, DNA resequencing, and transcriptomics. **(C)** 6-phosphogluconate dehydrogenase (GND) and Succinate Dehydrogenase (SUCDi) were disabled by the gene KOs. GND converts 6-phosphogluconate (6pgc) to ribulose 5 phosphate (ru5p-D) in the Pentose Phosphate Pathway (PPP), and SUCDi converts succinate (succ) to fumarate (fum) in the TCA cycle. **(D)** Growth rate and uptake and secretion rates for unevolved KO (uGnd and uSdhCB) and evolved KOs (eGnd and eSdhCB). The growth rate for uGnd, uSdhCB, eGnd strains, and eSdhCB strains are shown in the plot on the bottom right. Secretion rates for acetate (ac) and succinate (succ) are shown on the top plots. The glucose (glc-D) uptake rates are shown on the bottom left plot.

understand the drivers for flux rerouting in the absence of substantial growth rate loss.

Blocked Flux Through the Oxidative PPP in uGND Overflowed Into the ED Pathway

6-phosphogluconate dehydrogenase (GND) encoded by *gnd* catalyzes the decarboxylation of D-gluconate-6-phosphate (6pgc) to D-ribulose-5-phosphate (ru5p-D) while regenerating NADPH. Loss of *gnd* created a flux bottleneck at the 6pgc node, which resulted in a massive buildup of 6pgc and depletion of ribose-5-phosphate (r5p) in uGnd. 6pgc was 3.3 log2 fold higher and r5p was −0.6 log2 fold lower in concentration in uGnd than Ref. The relatively large proportion of flux compared to Ref remained through the first steps of the OxPPP, but the flux that would flow through GND instead spilled over into the Entner–Doudoroff (ED) pathway in uGnd (**Figure 2A**, **Table S6**). The spillover amounted to a 1.3 log2 fold higher absolute flux in ED in uGnd compared to Ref.

An overall decrease in OxPPP flux was also accompanied by an increase in upper glycolytic flux by necessity of balancing the glucose-6-phosphate (g6p) node (**Figure 2A**, **Table S6**). Thus, a 1.1 log2 fold higher absolute flux through upper glycolysis was found in uGnd compared to Ref. The increase in glycolytic flux also increased the levels of glycolytic metabolites (i.e., glucose-6-phosphate, fructose-6-phosphate, fructose 1,6-bisphosphate, and dihydroxyacetone phosphate, **Table S3**). For example g6p was 1.3 log2 fold higher in uGnd compared to Ref.

The r5p node was regenerated by re-routing flux through the non oxidative Pentose Phosphate Pathway (NonOxPPP) (**Figure 2A**, **Table S6**). These flux changes resulted in a flip in flux direction through the transketolase enzymes (TKT1 and TKT2) as well as the transaldolase enzyme (TALA).

Purine and Pyrimidine Biosynthetic Pathways Were Downregulated Due to Transcription Factor Activation in Both uGnd and uSdhCB Strains

The building blocks for purines and pyrimidines is r5p, generated from the PPP. Activity of the majority of purine and pyrimidine biosynthetic genes are feedback inhibited by end products including AMP, GMP, and UMP (Berg et al., 2002). In addition, gene expression of the majority of purine and pyrimidine biosynthetic genes are negatively regulated by the PurR transcription factor (TF), which is itself activated by hypoxanthine (hxan) (Choi and Zalkin, 1992; Cho et al., 2011).

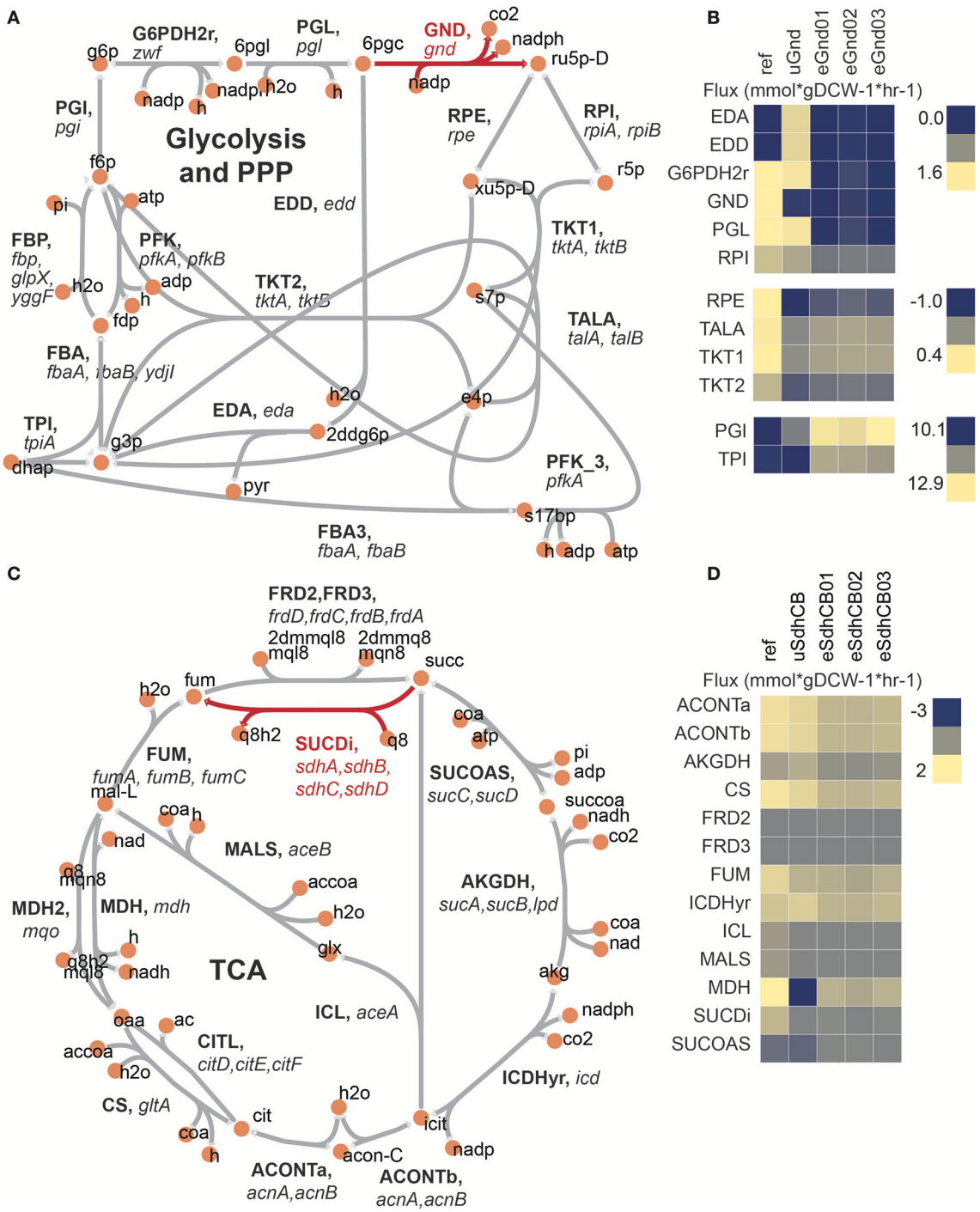

FIGURE 2 | Flux maps of *gnd* and *sdhCB* knockout strains at the vicinity of the network perturbation. **(A)** Network diagram of upper glycolysis, the oxidative pentose phosphate pathway (oxPPP), and the non oxidative pentose phosphate pathway (nonOxPPP). The KO reaction is highlighted in red. **(B)** Measured fluxes for Ref, uGnd, and eGnd strains. **(C)** Network diagram of the TCA cycle. The KO reaction is highlighted in red. **(D)** Measured fluxes for Ref, uSdhCB, and eSdhCB. All reaction and metabolite identifications shown are based on the iJO1366 model of *E. coli*.

Interestingly, all *de novo* purine and pyrimidine biosynthetic genes were downregulated in uGnd and uSdhCB and subsequently restored to levels similar to Ref in eGnd and eSdhCB strains (**Table S4**). Down regulation of purine and pyrimidine biosynthetic genes appears non-intuitive because the levels of the purine and pyrimidine building block, r5p,

and flux through the PPP were at dramatically different levels in uGnd and uSdhCB strains. However, the levels of the PurR TF activator, hxan, was significantly elevated in both uGnd and uSdhCB strains. hxan was 4.2 and 4.6 log2 fold higher in uGnd and uSdhCB, respectively, compared to Ref. Given the down regulation of *de novo* purine and pyrimidine biosynthetic genes and non-significant change in *purR* expression (**Table S4**), the similar expression profiles of purine and pyrimidine genes in uGnd and eGnd appeared to be driven by similar levels of the PurR TF activator, hxan. hxan was restored to levels similar to Ref in eGnd and eSdhCB strains.

Decoupling of the TCA From the ETC in uSdhCB Resulted in Flux Cycling and Increased Flux Toward Nitrogen Assimilation

SUCDi catalyzes the conversion of succinate (succ) to fumarate (fum) while contributing electrons directly to the ETC. Removal of *sdhA*, *sdhB*, *sdhC*, and *sdhD* resulted in a massive build-up of succ that was largely excreted into the medium (**Figure 1D**). Even with the loss of Succinate Dehydrogenase (SUCDi), a small amount of conversion from succ to fumarate (fum) was found (**Figures 2C,D**, **Table S6**). 0.09 mmol*gDCW-1*hr-1 absolute flux (or −1.4 log2 fold change in uSdhCB compared to Ref) was found through SUCDi. A massive increase in gene expression of the *frdABCD* operon, which encodes the Fumarate Reductase (FRD) enzyme, was found (Cecchini et al., 2002). FRD is an iron-sulfur protein that is optimally active under anaerobic conditions (Cecchini et al., 2002). However, even when oxygen is present, FRD is able to catalyze the oxidation of succinate at a ratio of 1:1.5 compared to reduction of fumarate (Hirsch et al., 1963). For reference, SUCDi catalyzes the oxidation of succinate at a ratio of 25:1 compared to reduction of fumarate (Hirsch et al., 1963). The large expression levels of FRD and the massively elevated levels of succ indicated that the small amount of conversion from succ to fum was either due to mass action or suboptimal activity of FRD. Succ was 6.3 log2 fold higher in uSdhCB compared to Ref.

A larger, yet depleted, amount of flux exiting the fum node through FUM appeared to be driven primarily by recycling of fum from peripheral metabolism (**Figures 2C,D**, **Table S6**). Flux from succ to fum accounted for less than a quarter of the flux through FUM. In addition, a reduced amount of flux through the glyoxylate shunt was found (−0.7 log2 fold less in uSdhCB compared to Ref).

Surprisingly, an overall increase in flux entering the TCA cycle was found in uSdhCB (**Figures 2C,D**, **Table S6**). The increase in flux was driven by an increase in flux through Phosphoenolpyruvate Carboxylase (PPC) and maintenance of flux through Citrate Synthase (CS) leading to a 0.3 log2 fold change in Isocitrate Dehydrogenase (ICDHyr) flux in uSdhCB compared to Ref. Approximately half of the flux leaving the alpha-ketoglutarate (akg) node was directed toward nitrogen metabolism.

A cycle between flux entering the TCA through PPC and flux leaving the TCA through the Malic Enzymes (MEs) was also found (**Figures 2C,D**, **Table S6**). The cycle was made possible by a complete reversal of Malate Dehydrogenase (MDH) from 1.3 mmol*gDCW-1*hr-1 to −1.4 mmol*gDCW-1*hr in uSdhCB. This cycle appeared to deplete levels of oxaloacetate (oaa) based on the measured levels of L-aspartate (asp-L) while maintaining the levels of Malate (mal-L) (**Table S3**).

Common Expression Profiles in uGnd and uSdhCB in TCA Cycle Genes Were Driven by Transcription Factor and Two Component System Activation by Metabolites

Despite loss of enzymatic activity and major flux re-routing in distal network locations, expression profiles of TCA cycle genes were found to be remarkably similar (**Figure 3**). In particular, the expression levels of genes that encode fumarate reductase (*frdA*, *frdB*, *frdC*, and *frdD*) and succinate dehydrogenase (*sdhA*, *sdhB*, *sdhC*, and *sdhD*) were almost identical in uGnd and uSdhCB strains. In fact, expression of the *sdh* operon in uGnd was downregulated to near the levels measured in the KO uSdhCB (**Figure 3**, **Table S4**).

Upregulation of expression of the *frd* operon in uGnd and uSdhCB was found to be attributed to a common elevation in four-carbon acids (i.e., succinate, fumarate, and malate) (**Figures 3B–D**). Expression of the *frd* operon was most likely activated by the DcuR TF, which was most likely phosphorylated and activated by two component (TC) system pair DcuS in response to elevations in four-carbon acids (Janausch et al., 2004). succ, fum, and mal-L were 3.7, 2.9, and 1.21 log2 fold higher in uGnd compared to Ref, respectively, and 6.3, 0.7, and −0.1 log2 fold change in uSdhCB, respectively, compared to Ref. Downregulation of the *sdh* operon in uGnd and uSdhCB was found to be attributed to an attenuated anaerobic response that involved a complex interaction of TFs ArcA, CRP, SoxR, SoxS, Fur, and Fnr, and small RNAs *fnrS* and *ryhB* (**Figures 3E–H**). A shift in the oxidized status of the membrane bound quinones ubiquinone (q8 and q8h2) and menaquinones (mql8, mqn8, 2dmmql8, and 2dmmq8), and anaerobic metabolite (i.e., lac-D) most likely triggered the ArcAB TC, which phosphorylated and activated the ArcA TF.

The ensuing regulatory cascade following ArcA activation in uGnd and uSdhCB strains culminated in the similar downregulation of *fnr*. Coupled with a common activation of CRP-cAMP through increased levels of cAMP resulted in the downregulation of *soxS* and upregulation of *fnrS*. Unique to uGnd was the activation of *rhyB*, the small regulatory RNA, which may have attributed to the downregulation of *sdhABCD* genes.

Previous work has shown that removal of the ubiquinone pathway led to increased phosphorylation (i.e., activation) of ArcA under aerobic conditions (van Beilen and Hellingwerf, 2016). In contrast, removal of the menaquinone pathways resulted in a non-significant change in ArcA phosphorylation, but decreased levels of ArcA phosphorylation during anaerobiosis (van Beilen and Hellingwerf, 2016). In addition *arcA* upregulation and *fnr* downregulation has been found to lead to decreased expression of TCA cycle genes and an overall increase in fermentative metabolism (e.g., increased flux through

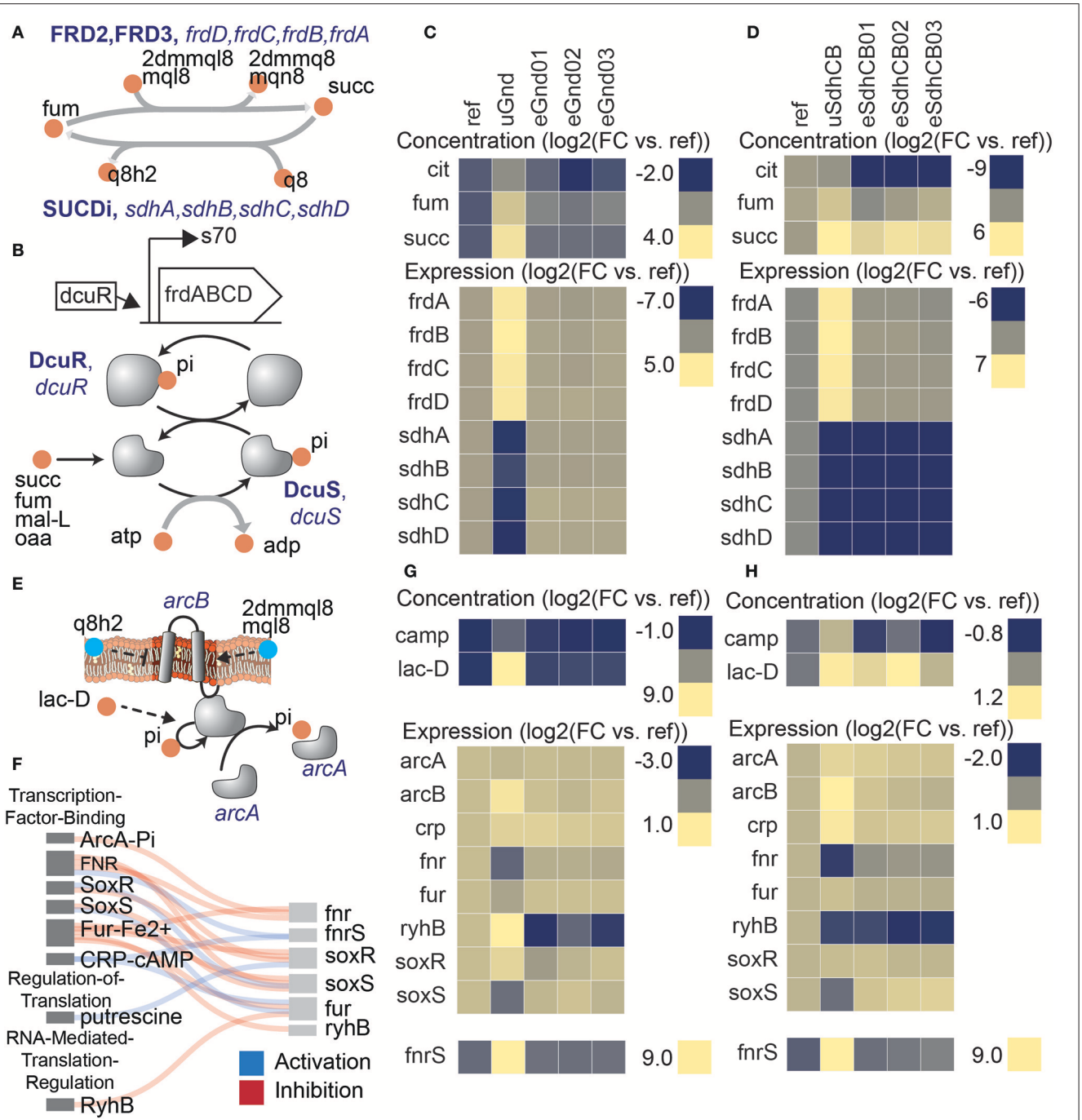

FIGURE 3 | Perturbations in separate network locations yielded similar expression states in the uKOs due to similar metabolite levels. Removal of the succinate dehydrogenase complex (sdhCB), which decoupled the TCA cycle from oxidative phosphorylation, and removal of 6-phosphogluconate dehydrogenase (gnd), which re-routed flux through upper glycolysis, the ED pathway, and the pentose phosphate pathway (PPP), resulted in similar expression profiles of TCA cycle genes as a result of increased levels of intracellular four-carbon acids **(A–D)**. **(A)** Reactions catalyzed by succinate dehydrogenase (SUCDi) and fumarate reductase (FRD2, FRD3) in the TCA cycle. **(B)** Schematic of the *frd* operon and regulation by DcuR. Also shown is a schematic of the dcuRS two component system (Janausch et al., 2004). Elevation in four-carbon acids (e.g., succinate, fumarate, malate, and oxaloacetate) were detected by the dcuRS two-component system in the uKO strains. Phosphorylated DcuR activated expression of the fumarate reductase operon genes. Metabolite levels of fumarate (fum), succinate (succ), and citrate (cit), and expression levels of fumarate reductase (*frdA*, *frdB*, *frdC*, *frdD*) and succinate dehydrogenase (*sdhA*, *sdhB*, *sdhC*, *sdhD*) genes for gnd **(C)** and sdhCB **(D)**. The similar metabolite levels in the uGnd and uSdhCB activated a network response that resulted in the upregulation of *frdABCD* genes in uGnd and uSdhCB, and downregulation of *sdhABCD* genes in uGnd. De-coupling of the TCA cycle from oxidative phosphorylation triggered an attenuated anaerobic response in gnd and sdhCB that involved a complex interaction of TFs ArcA, CRP, SoxR, SoxS, Fur, and Fnr, and small RNAs *fnrS* and *ryhB* **(E–H)**. **(E)** Regulatory schematic of the signal

(Continued)

FIGURE 3 | transduction cascade triggered by the oxidized status of the membrane bound quinones ubiquinone (q8 and q8h2) and menaquinols (mql8, mqn8, 2dmmql8, and 2dmmq8), and anaerobic metabolites (e.g., lac-D). **(F)** Regulatory interaction diagram between the different regulators. Metabolite and expression profiles of key components involved in the regulatory cascade for gnd **(G)** and sdhCB **(H)**. Note the similar downregulation of *fnr* in response to ArcA activation through increased levels of lac-L and changes in the oxidized status of the membrane bound quinones, the upregulation of *fnrS* in response to activation of CRP-cAMP through increased levels of cAMP, and the downregulation of *soxS* in uGnd and uSdhCB.

glycolysis and increased levels of organic acids) (Kumar and Shimizu, 2011; Basan et al., 2017). These finding are consistent with the expression profiles shown and highlight the regulatory shifts caused by the ArcAB TC and FNR regulators.

Differences in uGnd and uSdhCB OxPPP and TCA Cycle Fluxes and Derived Metabolites Were Found

In contrast to uGgnd, an overall increase in flux through the OxPPP was found in uSdhCB. In contrast to uSdhCB, an increase in flux through the lower half of the TCA cycle was found in uGnd (and also in eGnd strains) (**Table S6**). In particular, the increased flux through ICDHyr was most likely to compensate for the loss of NADPH generation in the oxPPP. NADPH was −1.3 log2 fold lower in uGnd compared to Ref. However, flux at the akg node was not significantly diverted toward nitrogen assimilation as was the case in uSdhCB. The insignificant divergence of flux toward nitrogen assimilation is consistent with previous work that has found an upregulation in TCA cycle flux toward succ in response to disruption of the *gnd* gene (Zhao et al., 2004; Mienda et al., 2016). These differences, among several others, set the stage for physiological advantages to be had through evolution and compensatory mutations.

Mutations at the *pntA* Promoter Boosted NADPH Levels in eGnd

Intergenic mutations that increased gene expression of the NADP(H) binding component of the insoluble pyridine nucleotide transhydrogenase (THD2pp) were found in eGnd strains (**Figure 4**). Specifically, single nucleotide polymorphisms (SNP) mutations in the vicinity of the *pntA* transcription start site (TSS) were found in all eGnd strains. THD2pp is composed of two subunits encoded by *pntA* and *pntB*. The former contains the NADP(H) binding domain while the latter contains the proton pumping transmembrane domain (Johansson et al., 2005). Significantly elevated expression of *pntA* was found in all eGnd strains. The expression levels of *pntA* were 2.5, 2.1, and 3.2 log2 fold higher in eGnd01, 02, and 03, respectively, compared to uGnd. THD2pp has been shown to provide an important source of NADPH in *E. coli* (Sauer et al., 2003), and it is highly likely that increased expression of *pntA* contributed to restoring the levels of NADPH in the eGnd strains.

Mutations Affecting Nitrogen Assimilation in eSdhCB Strains May Have Coordinated Increased Flux Out of the TCA Cycle With Nitrogen Regulation

The loss of SUCDi in uSdhCB created a bottleneck in the TCA cycle which led to an elevated amount of flux exiting the TCA in the steps leading to the production of succinate. As described above, this would lead to an increased amount of flux directed toward nitrogen assimilation, which resulted in a significant elevation of key nitrogen sensing metabolites, akg and gln-L (p-value < 0.05). akg and gln-L were 1.2 and 1.0 fold higher in uSdhCB compared to Ref. Several mutations were found in eSdhCB strains that altered enzymatic activity of metabolic genes and key regulators in nitrogen assimilation pathways that potentially balanced increased flux with nitrogen generation. These included mutations in *allD* and *glnD* that are described below.

A mutation in the active site of ureidoglycolate dehydrogenase (URDGLYCD) in eSdhCB01 was found that potentially provided an auxiliary means to metabolize excess glyoxylate (glx) and/or balance nitrogen levels (**Figure 5**). The *allDCE* operon encodes genes involved in converting allantoin (alln) to glyoxylate (glx) and oxaluric acid (oxur) (Cusa et al., 1999) (**Figure 5B**). oxur can then be broken down to oxymate (oxam), and carbamoyl phosphate (cbp); the latter provides a source of ATP and nitrogen (Cusa et al., 1999) (**Figure 5B**). The *allDCE* operon is positively regulated by AllS, whose gene expression is negatively regulated by AllR (Rintoul et al., 2002) (**Figure 5A**). *allA* and *allB* are also negatively regulated by AllR (Rintoul et al., 2002) (**Figure 5A**). Glyoxylate enzymatically inhibits AllR repression of *allS*, *allA*, and *allB* (Rintoul et al., 2002*)* (**Figure 5A**). A deletion (DEL) mutation was found in eSdhCB01 in the active site of URDGLYCD that would most likely affect binding of NAD and other substrates (**Figure 5C**). Interestingly, the gene expression profiles of all allonoin related genes are consistent with high glyoxylate levels in uSdhCB, which were found to be 3.3 log2 fold higher in uSdhCB compared to Ref (**Figure 5D**, **Tables S3, S4**). Given the high amounts of glyoxylate and also increased flux toward nitrogen metabolism, the mutation in eSdhCB01 may have conferred a fitness advantage by providing an additional means to metabolize glyoxylate and/or balance the levels of ammonia and nitrogen in the cell.

Mutations were also found in *glnD*, which encodes the primary nitrogen status sensor PII uridylyl-/deuridylyl-transferase, that fixed in eSdhCB02 and that overtook a majority of the population in eSdhCB03 (**Figure 6**). Nitrogen assimilation is heavily regulated in *E. coli* (see van Heeswijk et al., 2013 for a review) (**Figures 6A,B**). PII senses nitrogen levels primarily through the concentration of L-glutamine (gln-L). Increased levels of gln-L increase the deuridylylation activity of glnD and decrease the uridylyltransferase activity of glnD (Jiang et al., 2012). PII-ump stimulates deadenylation of GLNS via glnE while PII stimulates adenylation of GLNS (Rhee et al., 2006; van Heeswijk et al., 2013); the removal of amp enhances GLNS activity (Rhee et al., 2006; van Heeswijk et al., 2013).

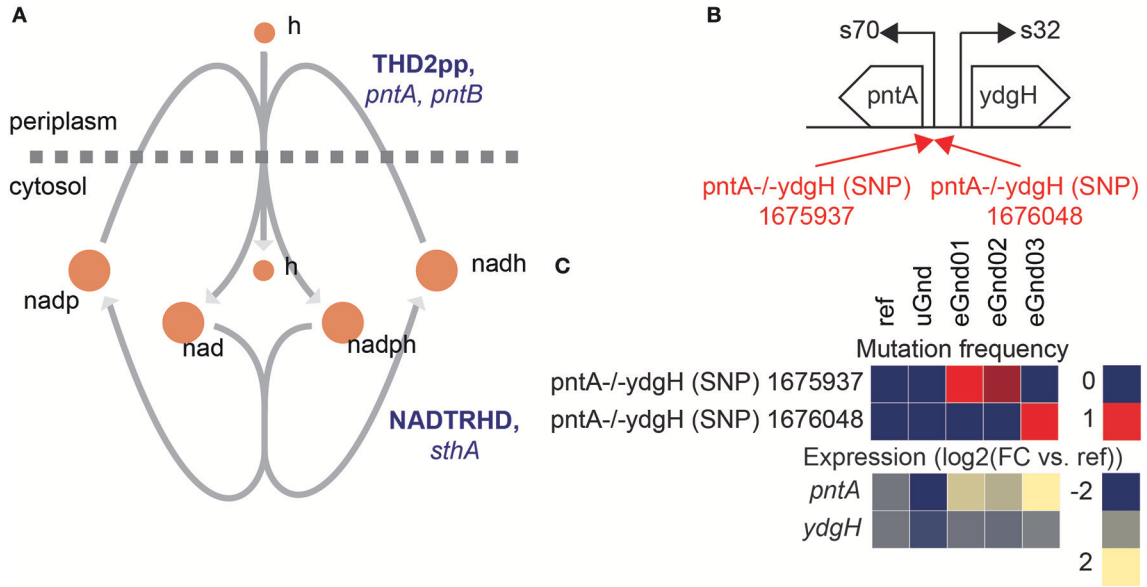

FIGURE 4 | Intergenic mutations that increased gene expression of the NADP(H) binding component of the insoluble pyridine nucleotide transhydrogenase in eGnd strains. **(A)** Network diagram of the insoluble (THD2pp) and soluble (NADTRHD) pyridine nucleotide transhydrogenase. THD2pp has been shown to provide an important source of NADPH in *E. coli* (Sauer et al., 2003). The insoluble pryidine nucleotide transhydrogenase is composed of two subunits encoded by *pntA* and *pntB*. The former contains the NADP(H) binding domain while the latter contains the proton pumping transmembrane domain (Johansson et al., 2005). **(B)** Operon diagram of *pntA*. Mutations that affected gene expression of *pntA* were found in all eGnd strains. The 1675937 SNP is 62 nucleotides downstream of the *pntA* transcription start site (TSS) and the 1676048 SNP is 49 nucleotides downstream of the pntA TSS. Mutations are annotated in red. **(C)** Mutation frequency and expression levels of *pntA* and *ydgH*. Note that all eGnd strains significantly elevated expression of *pntA*.

The mutations in *glnD* were located in the ACT 1 and 2 domains (**Figure 6C**), which are believed to play a direct role in binding and sensing gln-L (Zhang et al., 2010; Jiang et al., 2012). Given the increased flux toward nitrogen assimilation out of the TCA cycle and elevated gln-L levels (**Tables S3, S6**), the mutations potentially provided a fitness advantage by altering regulation of nitrogen assimilation in eSdhCB strains.

Altered TCA Cycle Flux Perturbed Sulfur Metabolism Gene Expression in uSdhCB

Major perturbations in sulfur metabolic pathway gene expression were found in uSdhCB (Figure S2). The sulfur metabolic pathway converts sulfate (so4), asp-L, L-serine (ser-L), and Succinyl-CoA (succoa) to L-cysteine (cys-L), which is then converted to L-methionine (met-L). The pathway is also regulated enzymatically and transcriptionally by multiple intermediates and end-products of the pathway. Major changes included a significant decrease in gene expression of *cysDNC* and *cysJIH* operons in uSdhCB (**Figures 2A,B,E,F**). *cysDNC* and *cysJIH* operons encode enzymes involved in sulfate reduction, and are controlled by the TF CysB. A significant decrease in gene expression of *thrA* and *metL* were found (**Figures S2C,D**). *thrA* and *metL* encode the fused aspartate kinase/homoserine dehydrogenase 1 and 2 (ASPK and HSDy), respectively. ASPK and HSDy catalyze the conversion of asp-L to L-aspartyl-4-phosphate (4pasp) and L-aspartate-semialdehyde (aspsa) to L-homoserine (hom-L) (Falcoz-Kelly et al., 1969; Starnes et al., 1972). *metL* gene expression is repressed by high met-L levels (Patte et al., 1967). The levels of met-L were 0.82 log2 fold higher in uSdhCB compared to Ref.

Gene expression changes in the eSdhCB strains resulted in an increased flux toward hom-L biosynthesis (**Figure S2D**). ASPK, ASAD, and HSDy flux increased by 0.7, 0.7, and 0.9 log2 fold in eSdhCB strains, respectively, compared to Ref. A significant decrease in gene expression of *cysE*, *cysK*, and *cysM* genes were found (**Figures S2E,F**). *cysE*, *cysK*, and *cysM* encode enzymes involved in L-cysteine biosynthesis. In conjunction with *cysE*, *cysK*, and *cysM* down regulation, a significant decrease in gene expression of *metB*, *metC*, *metH*, and *metE* was also found. *metB*, *metC*, *metH*, and *metE* encode enzymes involved in met-L biosynthesis.

In contrast to the down regulation of the majority of sulfur metabolic genes, a significant increase in gene expression of *metA* was found (**Figures S2E,F**). *metA* encodes the first step in *de novo* L-methionine biosynthesis catalyzed by homoserine O-succinyltransferase (HSST). One of the substrates of HSST, succinyl-CoA (succoa), is derived from the TCA cycle and two steps removed from SUCDi. Interestingly, other inputs into sulfur metabolism, asp-L and ser-L, were perturbed due to changed metabolic flux (**Table S3**). These results indicated that the perturbed gene expression of the majority of sulfur metabolic genes could be attributed to changed pathway flux and levels of key signaling metabolites.

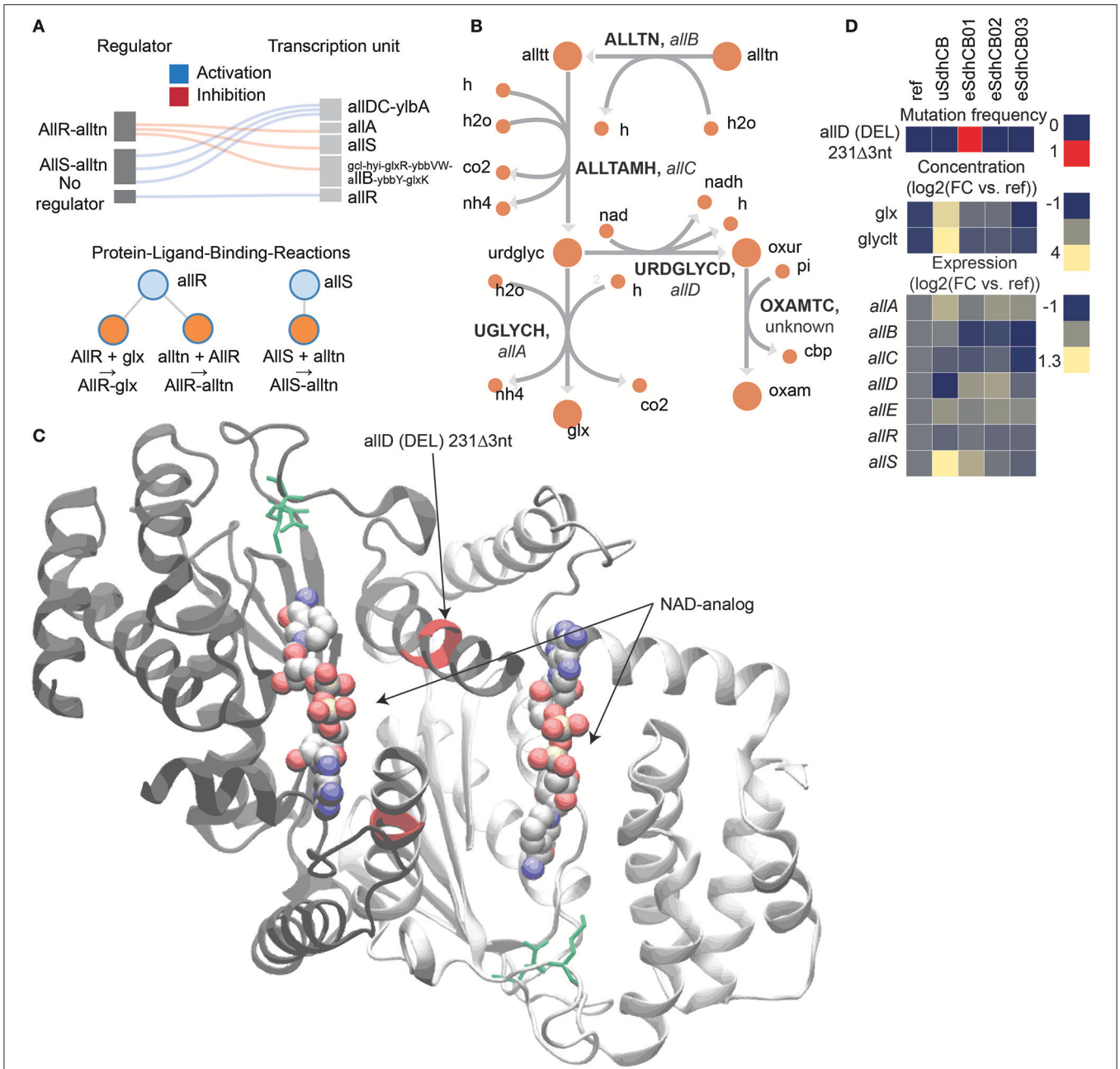

FIGURE 5 | A mutation in the active site of ureidoglycolate dehydrogenase (URDGLYCD) in eSdhCB01 that potentially provided an auxiliary means to metabolize excess glyoxylate (glx) and/or balance nitrogen levels. **(A)** Regulatory diagram and protein-ligand-binding reactions. The *allDCE* operon is positively regulated by AllS, whose gene expression is negatively regulated by AllR (Rintoul et al., 2002). *allA* and *allB* are also negatively regulated by AllR (Rintoul et al., 2002). Glyoxylate enzymatically inhibits AllR repression of *allS*, *allA*, and *allB* (Rintoul et al., 2002). **(B)** Network diagram of allantoin assimilation. **(C)** Crystal structure of AllD. Mutations are highlighted in red. Active site residues are highlighted in green. An NAD-analog where NAD would bind is shown. Note that the deletion (DEL) mutation in eSdhCB01 would affect active site residues, and most likely binding of NAD and other substrates. **(D)** Mutation frequency, metabolite levels, and gene expression of components involved in allantoin assimilation. Note that gene expression profiles are consistent with high glyoxylate levels in uSdhCB.

Intergenic Mutations Were Found That Enhanced Expression of Sulfur Metabolic Genes in eSdhCB Strains

Expression of sulfur metabolic pathway genes were restored to levels similar to Ref in all eSdhCB strains. The restoration of expression was most likely due to restoration of key regulator metabolites including met-L (**Table S3**). However, significantly elevated expression of the *cysDNC* operon in eSdhCB03 was found that most likely resulted from an intergenic mutation (**Figure 7**). Two intergenic mutations were found that targeted

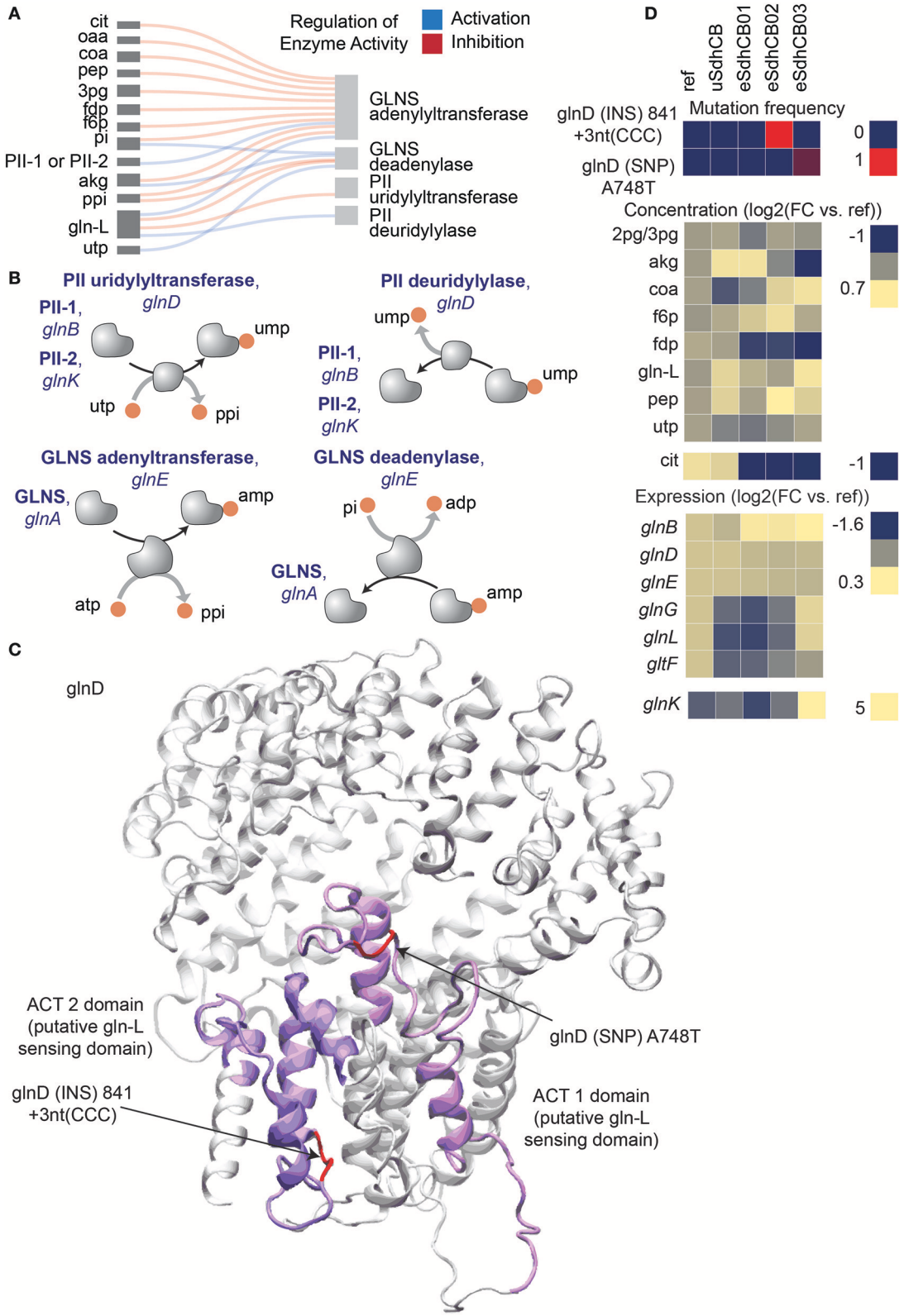

FIGURE 6 | Mutations that potentially affect regulation of nitrogen assimilation in eSdhCB strains. Mutations in *glnD*, which encodes the primary nitrogen status sensor PII uridylyl-/deuridylyl-transferase, fixed in eSdhCB02 and overtook a majority of the population in eSdhCB03. **(A)** Regulatory diagram of allosteric regulation

(Continued)

FIGURE 6 | of PII uridylyl-/deuridylyl-transferase and GLNS adenylyltransferase/GLNS deadenylase. **(B)** Protein-protein interaction diagram of PII uridylyl-/deuridylyl-transferase and GLNS adenylyltransferase/GLNS deadenylase. Increased levels of gln-L increase the deuridylylation activity of glnD and decrease the uridylyltransferase activity of glnD (Jiang et al., 2012). PII-ump stimulates deadenylation of GLNS via glnE while PII stimulates adenylation of GLNS (Rhee et al., 2006; van Heeswijk et al., 2013), the removal of amp enhances GLNS activity (Rhee et al., 2006; van Heeswijk et al., 2013). **(C)** Crystal structure of PII uridylyl-/deuridylyl-transferase [Reference]. Mutations are highlighted in red. The ACT 1 and 2 domains are highlighted in dark and light magenta, respectively. The ACT 1 and 2 domains are believed to play a role in sensing the intracellular levels of gln-L (Zhang et al., 2010; Jiang et al., 2012). Note that the mutations were located in the ACT 1 and 2 domains. **(D)** Mutation frequency, metabolite levels, and gene expression levels of components involved in *E. coli* nitrogen assimilation.

the RNAP binding site (eSdhCB02) and the CysB binding site (eSdhCB03) (**Figure 7A**). *cysDNC* is positively regulated by the CysB TF. *cysD* and *cysN* encode subunits of sulfate adenylyltransferase (SADT2) which converts sulfate (so4) to adenosine 5′-phosphosulfate (aps) (**Figure 7B**). *cysC* encodes adenylyl-sulfate kinase (ADSK) which converts aps to 3′-phosphoadenylyl-sulfate (paps) (**Figure 7B**). The eSdhCB03 was found to have significantly higher expression levels of *cysDNC* than all other strains (**Figure 7C**).

Mutations in AKGDH Decreased TCA Cycle Flux and Reduced Succinate Secretion in All eSdhCB Strains

Succinate secretion was significantly reduced in almost all eSdhCB strains compared to uSdhCB (**Figure 1D**). Overall flux from citrate (cit) to succ was also significantly reduced in all eSdhCB strains (**Figure 2D**). A major contributor to reduced TCA cycle flux and succinate secretion were mutations that affected 2-oxoglutarate Dehydrogenase (AKGDH) expression and enzymatic activity leading to significantly depressed flux in all eSdhCB strains (**Figure 8**). AKGDH is a multimer, composed of multiple subunits encoded by *sucA*, *sucB*, and *lpdA*, which converts akg to succoa while generating NADH (**Figures 8A,B**). Gene expression of *sucAB* in all eSdhCB strains was upregulated compared to uSdhCB except for eSdhCB01, which was significantly depressed compared to all other strains. The drop in gene expression of eSdhCB01 was most likely due to an intergenic mutation (**Figure 8C**, **Table S4**). In addition, flux appeared to be reduced in all strains due to mutations that affected substrate binding and complex formation (**Figure 8C**). A DEL mutation in *sucA* in eSdhCB03 occurred at residues that affect substrate binding, while SNP mutations in multiple strains occurred in regions that could affect *sucA* homomer formation (**Figure 8D**). A DEL mutation in *sucB* in eSdhCB01 cleaved residues 405 to 273, which are located in the active site (**Figure 8D**). In total, these mutations appeared to confer a fitness advantage in all eSdhCB strains by modulating the flux through AKGDH, which ultimately reduced succ byproduct secretion.

Adaptive Evolution Re-wired the Flux Map of eGnd and eSdhCB Strains Despite an Insignificant Change in Growth Rate

While insignificant changes in growth rate were found in most eGnd and eSdhCB strains, massive changes in metabolic flux were found (**Table S6**). These shifts in flux were a result of mutations that were selected for during evolution (**Figures 3**–**8**, **Table S8**) as described above. For example, while flux was initially diverted through the ED pathway in uGnd, all eGnd strains completely inactivated oxPPP and ED flux in exchange for increased flux through upper glycolysis and the nonOxPPP (**Figures 2A,B**). This shift also entailed another reversion of flux through the transketolases and transaldolase reactions (**Figures 2A,B**). The inactivation of the oxPPP was made possible by increased flux through ICDHyr and mutations in THD2pp (**Figure 4**) that made up for the lack of NADPH production in the oxPPP. Major flux shifts were also found in all eSdhCB strains. Most notably, TCA cycle flux was significantly reduced in all eSdhCB strains (**Figures 2C,D**). This shift was made possible by mutations affecting nitrogen metabolism (**Figures 5**,**6**), sulfur metabolism (**Figure 7**), and AKGDH in the TCA cycle (**Figure 8**) that appeared to coordinate the amount of flux out of the TCA cycle with generation of precursor metabolites derived from the TCA cycle.

CONCLUSION

While disruption of the GND and SUCDi reactions in the oxPPP and TCA cycles, respectively, resulted in minimal changes to growth rate, massive changes in metabolite levels and metabolic flux were found. Interestingly, many similarities in gene expression states between both strains despite the large difference in network location of the perturbations were found. Commonalities in gene expression profiles could be traced back to known metabolite regulators that were similarly perturbed in both strains. Minimal changes in growth rate were found after ALE, but mutations that lead to major divergence in flux and gene expression states from the unevolved strains were found. In the *gnd* evolutions, mutations were found that compensated for the reduced ability to generate NADPH through the oxPPP; in the *sdhCB* evolutions, mutations were found that reduced TCA cycle flux while balancing the regulation of nitrogen and sulfur assimilation. The divergence of regulatory states after evolution demonstrates that while metabolic and regulatory networks are robust to perturbation, the initial adjustments are often not optimal even when minimal changes in fitness occured. It is only after adaptation that the optimal regulatory and flux states were revealed.

MATERIALS AND METHODS

Biological Material, Analytical Reagents, and Experimental Conditions

Biological Material

A glucose, 37°C, evolved *E.coli* derived from *E. coli* K-12 MG1655 (ATCC 700926)(Ishii et al., 2007; Nakahigashi et al., 2009)

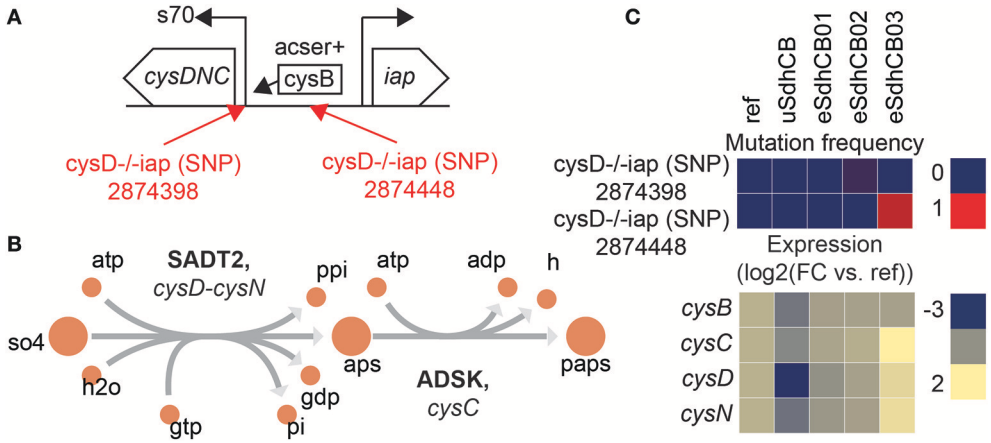

FIGURE 7 | Intergenic mutations that enhance expression of the *cysDNC* operon in eSdhCB strains. **(A)** Operon diagram of the *cysDNC* and *iap* operons. *cysDNC* is positively regulated by the transcription factor CysB, and encodes enzymes involved in sulfate reduction. Mutations are annotated in Red. The 2874398 SNP is located at the RNAP binding site. The 2874448 SNP is located at the CysB binding site. **(B)** Network schematic of the reactions encoded by *cysD*, *cysN*, and *cysC*. *cysD* and *cysN* encode subunits of sulfate adenylyltransferase (SADT2) which converts sulfate (so4) to adenosine 5′-phosphosulfate (aps). *cysC* encodes adenylyl-sulfate kinase (ADSK) which converts aps to 3′-phosphoadenylyl-sulfate (paps). **(C)** Mutation frequency and expression levels for genes involved in expression of the *cysDNC* operon. Note that eSdhCB03 has significantly higher expression levels of *cysDNC* than all other strains.

served as the starting strain. Lambda-red mediated DNA mutagenesis (Zhao et al., 2004) was used to create the knockout strains (DNA mutagenesis and PCR confirmation primers are given in **Table S2**). Knockouts were confirmed by PCR and DNA resequencing. Genes gnd, ptsH, ptsI, crr, sdhC, sdhA, sdhD, sdhC, tpiA, and pgi encoding for the reactions of 6-phosphogluconate dehydrogenase (GND), phosphotransferase sugar import (GLCptspp), succinate dehydrogenase complex (SUCDi), triophosphate isomerase (TPI), and phosphoglucose isomerase (PGI) were removed. PPC was also deleted, but resulted in an auxotrophy for asp-L, and was not included in the study. Genes aceE, aceF, zwf, and atpI-A encoding for the reactions of PDH, G6PDH2r, and ATPS4rpp could not be removed using the method of Datsenko et al. All cultures were grown in 25 mL of unlabeled or labeled glucose M9 minimal media (Nicolas et al., 2007) with trace elements (Olavarria et al., 2014) and sampled from a heat block in 50 mL autoclaved tubes that were maintained at 37°C and aerated using magnetics.

Materials and Reagents

Uniformly labeled ^{13}C glucose and 1-^{13}C glucose was purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA). Unlabeled glucose and other media components were purchased from Sigma-Aldrich (St. Louis, MO). LC-MS reagents were purchased from Honeywell Burdick and Jackson® (Muskegon, MI), Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich (St. Louis, MO).

Reaction Knockout Selection

iJO1366 (Tenaillon et al., 2016) was used as the metabolic model for *E. coli* metabolism; GLPK (version 4.57) was used as the linear program solver. MCMC sampling (Plucain et al., 2014) was used to predict the flux distribution of the optimized reference strain. Uptake, secretion, and growth rates were constrained to the measured average value ± SD. Potential reaction deletions were ranked by (1) averaged sampled flux, (2) the number of immediate upstream and downstream metabolites that could be measured, (3) the number of genes required to produce a functional enzyme. Reactions involved in sampling loops, that were spontaneous, were computationally or experimentally essential, or were not actively expressed under the experimental growth conditions were not included in the analysis. Also, reactions that would require more than one genetic alteration to abolish activity were excluded. The top 9 reactions deletions from the rank ordered set of reactions that met the above criteria were chosen for implementation.

Adaptive Laboratory Evolution (ALE)

Cultures were serially propagated (100 μL passage volume) in 15 mL (working volume) flasks of M9 minimal medium with 4 g/L glucose, kept at 37°C and well-mixed for full aeration. An automated system passed the cultures to fresh flasks once they had reached an OD_{600} of 0.3 (Tecan Sunrise plate reader, equivalent to an OD_{600} of ∼1 on a traditional spectrophotometer with a 1 cm path length), a point at which nutrients were still in excess and exponential growth had not started to taper off (confirmed with growth curves and HPLC measurements). Four OD_{600} measurements were taken from each flask, and the slope of $\ln(OD_{600})$ vs. time determined the culture growth rates. A cubic interpolating spline constrained to be monotonically increasing was fit to these growth rates to obtain the fitness trajectory curves.

Multi-Omics Data Processing

Phenomics

All cultures were grown grown on the same media used for the ALE experiments, and in a tumble stir system described in LaCroix et al. (2015) set to maintain a temperature of 37°C and optimal aeration. Physiological measurements for

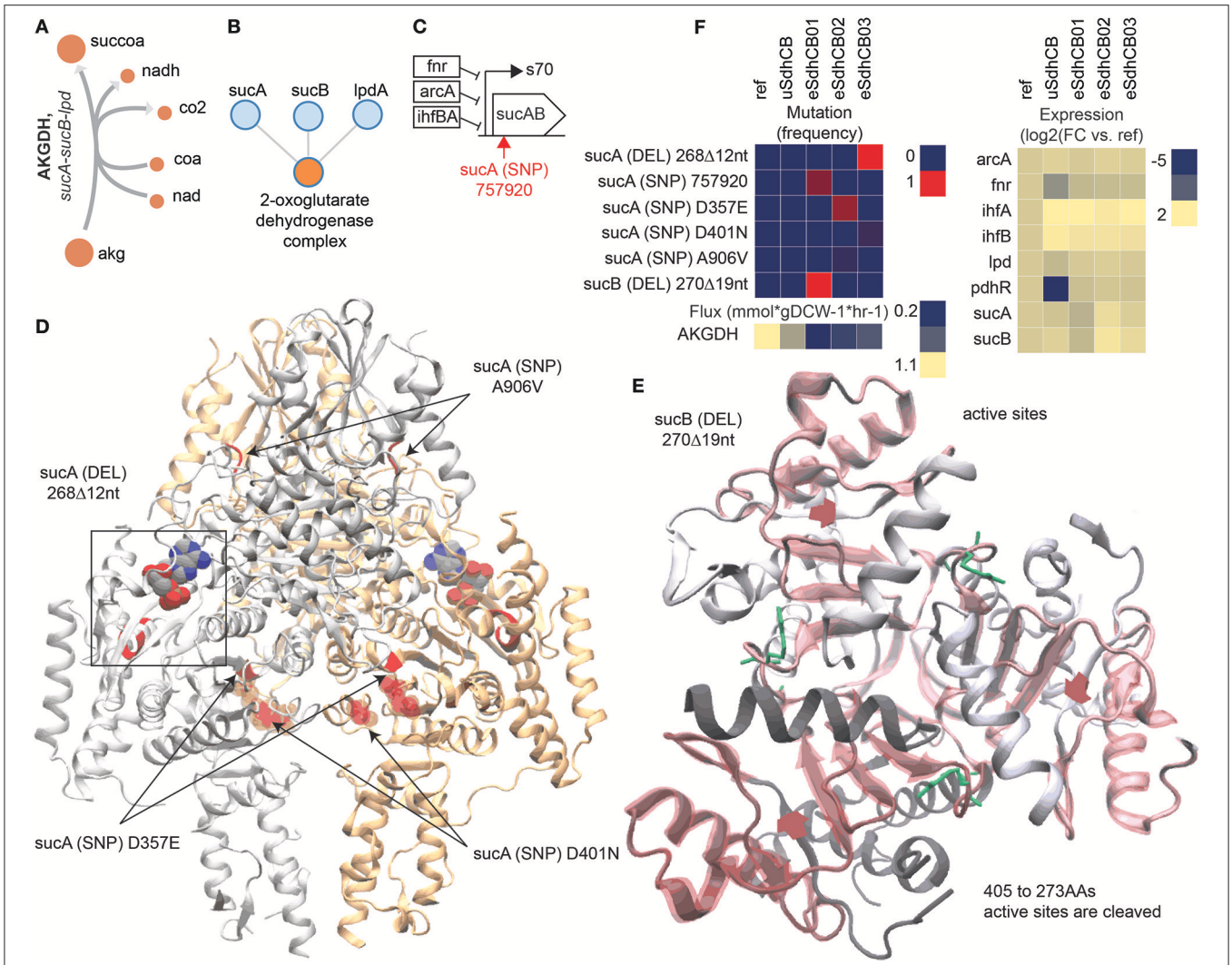

FIGURE 8 | Mutations that affect 2-oxoglutarate Dehydrogenase (AKGDH) enzymatic activity and expression in eSdhCB strains. **(A)** Network diagram of the AKGDH reaction. **(B)** Protein-protein binding diagram of the 2-oxoglutarate Dehydrogenase complex formation. **(C)** Operon diagram of the *sucAB* operon. The position of the intergenic mutation is highlighted in red. Gene expression of *sucAB* in eSdhCB01 is significantly depressed compared to all other strains, most likely due to the intergenic mutation (**Table S4**). In contrast, *sucAB* is upregulated in eSdhCB02/03 compared to uSdhCB (**Table S4**). **(D)** Crystal structure of SucA. Mutations are highlighted in Red. A substrate analog is shown. The deletion (DEL) mutation in eSdhCB03 occurred at residues that affect substrate binding. The sucA D357E, A906V, and D40N SNPs all occur near homodimer interface regions. **(E)** Crystal structure of SucB. Mutations are highlighted in red. Cleaved regions are outlined in red. Substrate binding residues are highlighted in green. The DEL mutation cleaved residues 405–273, which are all in the active site. **(F)** Mutation frequency, gene expression levels, and flux of components involved in the AKGDH reaction. Note the reduced flux through AKGDH in all eSdhCB strains.

culture density were measured at 600 nm absorbance with a spectrophotometer and correlated to cell biomass. A minimum of 8 time points starting from OD_{600} of 0.1 to ~1.0 (mid exponential phase) were used to calculate the growth rate. Samples to determine substrate uptake and secretion were filtered through a 0.22 μm filter(PVDF, Millipore) and measured using refractive index (RI) detection by HPLC (Agilent 12600 Infinity) with a Bio-Rad Aminex HPX87-H ion exclusion column (injection volume, 10 ul) and 5 mM H2SO4 as the mobile phase (0.5 ml/min, 45°C). Growth, uptake, and secretion rates were calculated from a minimum of four steady-state time-points.

LC-MS/MS Instrumentation and Data Processing

Metabolites were acquired and quantified on an AB SCIEX Qtrap® 5500 mass spectrometer (AB SCIEX, Framingham, MA) and processed using MultiQuant® 3.0.1 as described previously(Dragosits and Mattanovich, 2013). Mass isotopomer distributions (MIDs) were acquired on the same instrument and processed using MultiQuant® 3.0.1 and PeakView® 2.2 as described previously (Chou et al., 2015).

Metabolomics

Internal standards were generated as described previously (Fong et al., 2006). All samples and calibrators were spiked with the

same amount of internal standard taken from the same batch of internal standards. Calibration curves were ran before and after all biological and analytical replicates. The consistency of quantification between calibration curves was checked by running a Quality Control sample that was composed of all biological replicates twice a day. Solvent blanks were injected every ninth sample to check for carryover. System suitability tests were injected daily to check instrument performance.

Metabolomics samples were acquired from triplicate cultures (1 mL of cell broth at an OD600 $\sim$ 1.0) using a previously described method(Ibarra et al., 2002). A pooled sample of the filtered medium that was re-sampled using the FSF filtration technique and processed in the same way as the biological triplicates was used as an analytical blank. Extracts obtained from triplicate cultures and re-filtered medium were analyzed in duplicate. The intracellular values reported, unless otherwise noted, are derived from the average of the biological triplicates ($n = 6$). Metabolites in the pooled filtered medium with a concentration greater than 80% of that found in the triplicate samples were not analyzed. In addition, metabolites that were found to have a quantifiable variability (RSD $\geq$ 50%) in the Quality Control samples or any individual components with an RSD $\geq$ 80 were not used for analysis.

Missing values were imputed using a bootstrapping approach as coded in the R package Amelia II (Fong et al., 2005) (version 1.7.4, 1000 imputations). Remaining missing values were approximated as ½ the lower limit of quantification for the metabolite normalized to the biomass of the sample. Prior to statistical analyses, metabolite concentrations were log normalized to generate an approximately normal distribution using the R package LMGene (Moore et al., 2000) (version 3.3, "mult" = "TRUE," "lowessnorm" = "FALSE"). A Bonferroni-adjusted p-value cutoff of 0.01 as calculated from a Student's t-test was used to determine significance between metabolite concentration levels. The glog-normalized values or the median-normalized values to the reference strain (FC-median vs. ref) were used for downstream statistical analyses.

Fluxomics

Fluxomics samples were acquired from triplicate cultures (10 mL of cell broth at an OD600 $\sim$ 1.0) using a modified version of the FSF technique as described previously(Chou et al., 2015). MIDs were calculated from biological triplicates ran in analytical duplicate ($n = 6$). MIDs with an RSD greater than 50 were excluded. In addition, MIDs with a mass that was found to have a signal greater than 80% in unlabeled or blank samples were excluded. A previously validated genome-scale MFA model of *E. coli* with minimal alterations was used for all MFA estimations using INCA(LaCroix et al., 2015) (version 1.4) as described previously(Sandberg et al., 2016). The model was constrained using MIDs as well as measured growth, uptake, and secretion rates. Best flux values that were used to calculate the 95% confidence intervals were estimated from 500 restarts.

The 95% confidences intervals were used as lower and upper bound reaction constraints for further constraint-based analyses. MFA derived constraints that violated optimality were discarded and resampled. The descriptive statistics (i.e., mean, median, interquartile ranges, min, max, etc.) for each reaction for each model were calculated from 5,000 points sampled from 5,000 steps using optGpSampler (Sandberg et al., 2014) (version 1.1), which resulted in an approximate mixed fraction of 0.5 for all models. A permuted p-value < 0.05 and geometric fold-change of sampled flux values > 0.001 were used to determine differential flux levels, differential metabolite utilization levels, and differential subsystem utilization levels between models. Demand reactions and reactions corresponding to Unassigned, Transport; Outer Membrane Porin, Transport; Inner Membrane, Inorganic Ion Transport and Metabolism, Transport; Outer Membrane, Nucleotide Salvage Pathway, Oxidative Phosphorylation were excluded from differential flux analysis. The geometric fold-change of the mean between models and the reference model were used for hierarchical clustering; the median, interquartile ranges, min, and max values of each sampling distribution for each reaction and model were used as representative samples for downstream statistical analyses.

Transcriptomics

Total RNA was sampled from triplicate cultures (3 mL of cell broth at an OD600 $\sim$ 1.0) and immediately added to 2 volumes Qiagen RNA-protect Bacteria Reagent (6 mL), vortexed for 5 s, incubated at room temperature for 5 min, and immediately centrifuged for 10 min at 17,500 RPMs. The supernatant was decanted and the cell pellet was stored in the $-80°$C. Cell pellets were thawed and incubated with Readylyse Lysozyme, SuperaseIn, Protease K, and 20% SDS for 20 min at 37°C. Total RNA was isolated and purified using the Qiagen RNeasy Mini Kit columns and following vendor procedures. An on-column DNase-treatment was performed for 30 min at room temperature. RNA was quantified using a Nano drop and quality assessed by running an RNA-nano chip on a bioanalyzer. The rRNA was removed using Epicenter's Ribo-Zero rRNA removal kit for Gram Negative Bacteria. a KAPA Stranded RNA-Seq Kit (Kapa Biosystems KK8401) was used following the manufacturer's protocol to create sequencing libraries with an average insert length of around $\sim$300 bp for two of the three biological replicates. Libraries were ran on a MiSeq and/or HiSeq (illumina).

RNA-Seq reads were aligned using Bowtie (Berg et al., 2002) (version 1.1.2 with default parameters). Expression levels for individual samples were quantified using Cufflinks (Choi and Zalkin, 1992) (version 2.2.1, library type fr-firststrand) Quality of the reads was assessed by tracking the percentage of unmapped reads and expression level of genes that mapped to the ribosomal gene loci rrsA-F and rrlA-F. All samples had a percentage of unmapped reads less than 7%. Differential expression levels for each condition ($n = 2$ per condition) compared to either the starting strain or initial knockout strain were calculated using Cuffdiff (Choi and Zalkin, 1992) (version 2.2.1, library type fr-firststrand, library norm geometric). Genes with an 0.05 FDR-adjusted p-value less than 0.01 were considered differentially expressed. Expression levels for individual samples for all combinations of conditions tested in down-stream statistical analyses were normalized using Cuffnorm (Choi and Zalkin, 1992) (version 2.2.1, library type fr-firststrand, library norm

geometric). Genes with unmapped reads were imputed using a bootstrapping approach as coded in the R package Amelia II (version 1.7.4, 1000 imputations). Remaining missing values were filled using the minimum expression level of the data set. Normalized FPKM values for gene expression were log2 normalized to generate an approximately normal distribution prior to any statistical analysis. All replicates for a given condition were found to have a pair-wise Pearson correlation coefficient of 0.95 or greater.

DNA Resequencing

Total DNA was sample from an overnight culture (1 mL of cell broth at an OD600 of ∼2.0) and immediately centrifuged for 5 min at 8,000 RPMs. The supernatant was decanted and the cell pellet was frozen in the −80°C. Genomic DNA was isolated using a Nucleospin Tissue kit (Macherey Nagel 740952.50) following the manufacturer's protocol, including treatment with RNase A. Resequencing libraries were prepared using a Nextera XT kit (Illumina FC-131-1024) following the manufacturer's protocol. Libraries were ran on a MiSeq (illumina).

DNA resequencing reads were aligned to the *E. coli* reference genome (U00096.2, genbank) using Breseq (Cho et al., 2011)(version 0.26.0) as populations. Mutations with a frequency of less than 0.1, *p*-value greater than 0.01, or quality score less than 6.0 were removed from the analysis. In addition, genes corresponding to crl, insertion elements (i.e., insH1, insB1, and insA), and the rhs and rsx gene loci were not considered for analysis due to repetitive regions that appear to cause frequent miscalls when using Breseq. mRNA and peptide sequence changes were predicted using BioPython (https://github.com/biopython/biopython.github.io/). Large regions of DNA (minimum of 200 consecutive indices) where the coverage was two times greater than the average coverage of the sample were considered duplications.

Structural Analysis

Corresponding PDB files for genes with a mutation of interested were downloaded from PDB (Hirsch et al., 1963; Cecchini et al., 2002). Structural models for genes for which there were no corresponding PDB files were taken from I-TASSER generated homology models (Janausch et al., 2004) or generated using the I-TASSER protocol (Mienda et al., 2016). The BioPython predicted sequence changes and important protein features as listed in EcoCyc (Johansson et al., 2005) were visualized and annotated using VMD (Sauer et al., 2003).

AUTHOR CONTRIBUTIONS

DM designed the experiments, generated the strains, conducted all aspects of the metabolomics, fluxomics, phenomics, transcriptomics, and genomics experiments, performed all multi-omics statistical, graph, and modeling analyses, and wrote the manuscript. TS ran the ALE experiments. EB assisted with structural analysis. RS processed the DNA and RNA samples. SX assisted with metabolomics and fluxomics data collection, sample processing, and peak integration. YH assisted with fluxomics data collection and sample processing. AF designed and supervised the evolution experiments, and contributed to the data analysis and the manuscript. BP conceived and outlined the study, supervised the data analysis, and co-wrote the manuscript.

FUNDING

This work was funded by the Novo Nordisk Foundation Grant Number NNF10CC1016517.

ACKNOWLEDGMENTS

We thank José Utrilla for helpful discussion and guidance when implementing the knockouts in the pre-evolved strain. We thank Jamey Young for helpful discussions throughout the MFA analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01793/full#supplementary-material

REFERENCES

Basan, M., Hui, S., and Williamson, J. R. (2017). ArcA overexpression induces fermentation and results in enhanced growth rates of *E. coli*. *Sci. Rep*. 7:11866. doi: 10.1038/s41598-017-12144-6

Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002). *Biochemistry, 5th Edn*. New York, NY: W. H. Freeman.

Cecchini, G., Schröder, I., Gunsalus, R. P., and Maklashina, E. (2002). Succinate dehydrogenase and fumarate reductase from *Escherichia coli*. *Biochim. Biophys. Acta* 1553, 140–157. doi: 10.1016/S0005-2728(01)00238-9

Cho, B.-K., Federowicz, S. A., Embree, M., Park, Y. S., Kim, D., and Palsson BØ. (2011). The PurR regulon in *Escherichia coli* K-12 MG1655. *Nucleic Acids Res*. 39, 6456–6464. doi: 10.1093/nar/gkr307

Choi, K. Y., and Zalkin, H. (1992). Structural characterization and corepressor binding of the *Escherichia coli* purine repressor. *J. Bacteriol*. 174, 6207–6214. doi: 10.1128/jb.174.19.6207-6214.1992

Chou, H.-H., Marx, C. J., and Sauer, U. (2015). Transhydrogenase promotes the robustness and evolvability of *E. coli* deficient in NADPH production. *PLoS Genet*. 11:e1005007. doi: 10.1371/journal.pgen.1005007

Cusa, E., Obradors, N., Baldomà, L., Badía, J., and Aguilar, J. (1999). Genetic analysis of a chromosomal region containing genes required for assimilation of allantoin nitrogen and linked glyoxylate metabolism in *Escherichia coli*. *J. Bacteriol*. 181, 7479–7484.

Dragosits, M., and Mattanovich, D. (2013). Adaptive laboratory evolution – principles and applications for biotechnology. *Microb. Cell Fact*. 12:64. doi: 10.1186/1475-2859-12-64

Falcoz-Kelly, F., van Rapenbusch, R., and Cohen, G. N. (1969). The methionine-repressible homoserine dehydrogenase and aspartokinase activities of *Escherichia coli* K 12. Preparation of the homogeneous protein catalyzing the two activities. Molecular weight of the native enzyme and of its subunits. *Eur. J. Biochem*. 8, 146–152. doi: 10.1111/j.1432-1033.1969.tb00507.x

Fong, S. S., Burgard, A. P., Herring, C. D., Knight, E. M., Blattner, F. R., Maranas, C. D., et al. (2005). *In silico* design and adaptive evolution of

Escherichia coli for production of lactic acid. *Biotechnol. Bioeng.* 91, 643–648. doi: 10.1002/bit.20542

Fong, S. S., Nanchen, A., Palsson, B. O., and Sauer, U. (2006). Latent pathway activation and increased pathway capacity enable *Escherichia coli* adaptation to loss of key metabolic enzymes. *J. Biol. Chem.* 281, 8024–8033. doi: 10.1074/jbc.M510016200

Hirsch, C. A., Rasminsky, M., Davis, B. D., and Lin, E. C. (1963). A fumarate reductase in *Escherichia coli* distinct from succinate dehydrogenase. *J. Biol. Chem.* 238, 3770–3774.

Ibarra, R. U., Edwards, J. S., and Palsson, B. O. (2002). *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature* 420, 186–189. doi: 10.1038/nature01149

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., et al. (2007). Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* 316, 593–597. doi: 10.1126/science.1132067

Janausch, I. G., Garcia-Moreno, I., Lehnen, D., Zeuner, Y., and Unden, G. (2004). Phosphorylation and DNA binding of the regulator DcuR of the fumarate-responsive two-component system DcuSR of *Escherichia coli*. *Microbiology* 150, 877–883. doi: 10.1099/mic.0.26900-0

Jiang, P., Zhang, Y., Atkinson, M. R., and Ninfa, A. J. (2012). The robustness of the *Escherichia coli* signal-transducing UTase/UR-PII covalent modification cycle to variation in the PII concentration requires very strong inhibition of the UTase activity of UTase/UR by glutamine. *Biochemistry* 51, 9032–9044. doi: 10.1021/bi3005736

Johansson, T., Oswald, C., Pedersen, A., Törnroth, S., Okvist, M., Karlsson, B. G., et al. (2005). X-ray structure of domain I of the proton-pumping membrane protein transhydrogenase from *Escherichia coli*. *J. Mol. Biol.* 352, 299–312. doi: 10.1016/j.jmb.2005.07.022

Kumar, R., and Shimizu, K. (2011). Transcriptional regulation of main metabolic pathways of cyoA, cydB, fnr and fur gene knockout *Escherichia coli* in Climited and N-limited aerobic continuous cultures. *Microb. Cell Fact.* 10:3. doi: 10.1186/1475-2859-10-3

LaCroix, R. A., Sandberg, T. E., O'Brien, E. J., Utrilla, J., Ebrahim, A., Guzman, G. I., et al. (2015). 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. doi: 10.1128/AEM.02246-14

Mienda, B. S., Shamsir, M. S., and Illias, R. M. (2016). Model-guided metabolic gene knockout of gnd for enhanced succinate production in *Escherichia coli* from glucose and glycerol substrates. *Comput. Biol. Chem.* 61, 130–137. doi: 10.1016/j.compbiolchem.2016.01.013

Moore, F. B., Rozen, D. E., and Lenski, R. E. (2000). Pervasive compensatory adaptation in *Escherichia coli*. *Proc. Biol. Sci.* 267, 515–522. doi: 10.1098/rspb.2000.1030

Nakahigashi, K., Toya, Y., Ishii, N., Soga, T., Hasegawa, M., Watanabe, H., et al. (2009). Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol. Syst. Biol.* 5:306. doi: 10.1038/msb.2009.65

Nicolas, C., Kiefer, P., Letisse, F., Krömer, J., Massou, S., Soucaille, P., et al. (2007). Response of the central metabolism of *Escherichia coli* to modified expression of the gene encoding the glucose-6-phosphate dehydrogenase. *FEBS Lett.* 581, 3771–3776. doi: 10.1016/j.febslet.2007.06.066

Olavarria, K., De Ingeniis, J., Zielinski, D. C., Fuentealba, M., Muñoz, R., McCloskey, D., et al. (2014). Metabolic impact of an NADH-producing glucose-6-phosphate dehydrogenase in *Escherichia coli*. *Microbiology* 160, 2780–2793. doi: 10.1099/mic.0.082180-0

Patte, J. C., Le Bras, G., and Cohen, G. N. (1967). Regulation by methionine of the synthesis of a third aspartokinase and of a second homoserine dehydrogenase in *Escherichia coli* K 12. *Biochim. Biophys. Acta* 136, 245–247. doi: 10.1016/0304-4165(67)90069-4

Plucain, J., Hindré, T., Le Gac, M., Tenaillon, O., Cruveiller, S., Médigue, C., et al. (2014). Epistasis and allele specificity in the emergence of a stable polymorphism in *Escherichia coli*. *Science* 343, 1366–1369. doi: 10.1126/science.1248688

Rhee, S. G., Chock, P. B., and Stadtman, E. R. (2006). "Regulation of *Escherichia coli* glutamine synthetase: meister/advances in advances," in *Enzymology and Related Areas of Molecular Biology*, ed A. Meister (John Wiley and Sons, Inc), 37–92.

Rintoul, M. R., Cusa, E., Baldomà, L., Badia, J., Reitzer, L., and Aguilar, J. (2002). Regulation of the *Escherichia coli* allantoin regulon: coordinated function of the repressor AllR and the activator AllS. *J. Mol. Biol.* 324, 599–610. doi: 10.1016/S0022-2836(02)01134-8

Sandberg, T. E., Long, C. P., Gonzalez, J. E., Feist, A. M., Antoniewicz, M. R., and Palsson, B. O. (2016). Evolution of *E. coli* on [U-13C]glucose reveals a negligible isotopic influence on metabolism and physiology. *PLoS ONE* 11:e0151130. doi: 10.1371/journal.pone.0151130

Sandberg, T. E., Pedersen, M., LaCroix, R. A., Ebrahim, A., Bonde, M., Herrgard, M. J., et al. (2014). Evolution of *Escherichia coli* to 42°C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. *Mol. Biol. Evol.* 31, 2647–2662. doi: 10.1093/molbev/msu209

Sauer, U., Canonaco, F., Heri, S., Perrenoud, A., and Fischer, E., (2003). The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in nadph metabolism of *Escherichia coli*. *J. Biol. Chem.* 279, 6613–6619. doi: 10.1074/jbc.M311657200

Starnes, W. L., Munk, P., Maul, S. B., Cunningham, G. N., Cox, D. J., and Shive, W. (1972). Threonine-sensitive aspartokinase-homoserine dehydrogenase complex amino acid compositions, molecular weight, and subunit composition of the complex. *Biochemistry* 11, 677–687. doi: 10.1021/bi00755a003

Tenaillon, O., Barrick, J. E., Ribeck, N., Deatherage, D. E., Blanchard, J. L., Dasgupta, A., et al. (2016). Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature* 536, 165–170. doi: 10.1038/nature18959

van Beilen, J. W. A., and Hellingwerf, K. J. (2016). All three endogenous quinone species of *Escherichia coli* are involved in controlling the activity of the Aerobic/Anaerobic response regulator ArcA. *Front. Microbiol.* 7:1339. doi: 10.3389/fmicb.2016.01339

van Heeswijk, W. C., Westerhoff, H. V., and Boogerd, F. C. (2013). Nitrogen assimilation in *Escherichia coli*: putting molecular data into a systems perspective. *Microbiol. Mol. Biol. Rev.* 77, 628–695. doi: 10.1128/MMBR.00025-13

Zhang, Y., Pohlmann, E. L., Serate, J., Conrad, M. C., and Roberts, G. P. (2010). Mutagenesis and functional characterization of the four domains of GlnD, a bifunctional nitrogen sensor protein. *J. Bacteriol.* 192, 2711–2721. doi: 10.1128/JB.01674-09

Zhao, J., Baba, T., Mori, H., and Shimizu, K. (2004). Global metabolic response of *Escherichia coli* to gnd or zwf gene-knockout, based on 13C-labeling experiments and the measurement of enzyme activities. *Appl. Microbiol. Biotechnol.* 64, 91–98. doi: 10.1007/s00253-003-1458-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 McCloskey, Xu, Sandberg, Brunk, Hefner, Szubin, Feist and Palsson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.