Assimilation of formic acid and CO_2 by engineered *Escherichia coli* equipped with reconstructed one-carbon assimilation pathways

Junho Bang[a,b] and Sang Yup Lee[a,b,c,d,1]

[a]Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 Plus Program), Institute for the BioCentury, Korea Advanced Institute of Science and Technology, 34141 Daejeon, Republic of Korea; [b]Systems Metabolic Engineering and Systems Healthcare Cross-Generation Collaborative Laboratory, Korea Advanced Institute of Science and Technology, 34141 Daejeon, Republic of Korea; [c]BioInformatics Research Center, Korea Advanced Institute of Science and Technology, 34141 Daejeon, Republic of Korea; and [d]BioProcess Engineering Research Center, Korea Advanced Institute of Science and Technology, 34141 Daejeon, Republic of Korea

Edited by James J. Collins, Massachusetts Institute of Technology, Boston, MA, and approved August 24, 2018 (received for review June 19, 2018)

Gaseous one-carbon (C1) compounds or formic acid (FA) converted from CO_2 can be an attractive raw material for bio-based chemicals. Here, we report the development of *Escherichia coli* strains assimilating FA and CO_2 through the reconstructed tetrahydrofolate (THF) cycle and reverse glycine cleavage (gcv) pathway. The *Methylobacterium extorquens* formate-THF ligase, methenyl-THF cyclohydrolase, and methylene-THF dehydrogenase genes were expressed to allow FA assimilation. The gcv reaction was reversed by knocking out the repressor gene (*gcvR*) and overexpressing the *gcvTHP* genes. This engineered strain synthesized 96% and 86% of proteinogenic glycine and serine, respectively, from FA and CO_2 in a glucose-containing medium. Native serine deaminase converted serine to pyruvate, showing 4.5% of pyruvate-forming flux comes from FA and CO_2. The pyruvate-forming flux from FA and CO_2 could be increased to 14.9% by knocking out *gcvR*, *pflB*, and *serA*, chromosomally expressing *gcvTHP* under *trc*, and overexpressing the reconstructed THF cycle, *gcvTHP*, and *lpd* genes in one vector. To reduce glucose usage required for energy and redox generation, the *Candida boidinii* formate dehydrogenase (Fdh) gene was expressed. The resulting strain showed specific glucose, FA, and CO_2 consumption rates of 370.2, 145.6, and 14.9 mg·g dry cell weight (DCW)$^{-1}$·h^{-1}, respectively. The C1 assimilation pathway consumed 21.3 wt% of FA. Furthermore, cells sustained slight growth using only FA and CO_2 after glucose depletion, suggesting that combined use of the C1 assimilation pathway and *C. boidinii* Fdh will be useful for eventually developing a strain capable of utilizing FA and CO_2 without an additional carbon source such as glucose.

tetrahydrofolate cycle | glycine cleavage pathway | formate dehydrogenase | formic acid | CO_2

Gaseous one-carbon (C1) compounds such as CO_2 and CH_4 are major greenhouse gases, which are largely responsible for global warming and climate change. With the aim of reducing greenhouse gases, biological conversion of C1 compounds has attracted much attention due to several advantages such as low energy requirements, environmental friendliness, and the possibility of converting greenhouse gases directly to value-added products (1–3). However, a major limitation of biological C1 conversion is the inefficiency of the natural C1 assimilation pathway such as the Calvin–Benson–Bassham (CBB) cycle (4). To address this problem, extensive studies have been performed to enhance biological C1 assimilation by developing de novo pathways using efficient carboxylases (5–9), introducing the CBB cycle into *Escherichia coli* (10–13), or by integration with an electrochemical method of reducing CO_2 to formic acid (FA) (14, 15). FA can be easily and efficiently synthesized from CO_2 by various chemical processes employing metal catalysts (16) or electricity (17). The use of FA as a fermentation substrate is advantageous as it can be easily stored and transported, it is soluble in water, and its biological assimilation is faster than that of CO_2 (18).

Due to these advantages of FA as a substrate, various studies have been conducted to biologically convert FA to value-added chemicals. For example, there has recently been a report on utilizing FA as a secondary carbon source to produce succinic acid (19). Numerous studies on natural and synthetic pathways for FA and CO_2 assimilation have suggested several promising candidate pathways including the serine pathway, the reductive acetyl-CoA pathway, the reductive glycine pathway, the serine–threonine pathway, the pyruvate formate lyase (Pfl)–threonine pathway, the formate tetrahydrofolate (THF) ligase-Pfl bicycle, the reductive glycine pathway with Pfl, the Pfl-phosphoketolase pathway, the formate reduction with ribulose monophosphate pathway, the 3-oxopropionyl-CoA pathway, and the synthetic enzyme formolase pathway (20–22). Among these candidate pathways, the THF cycle and glycine cleavage (gcv) pathway have several advantages, such as the oxygen tolerance of the enzymes involved, their high energy efficiency, and their independent operation with little overlapping of the central metabolism (23–25).

In a previous report, the THF cycle was established in *E. coli* by the expression of *Methylobacterium extorquens* formate-THF ligase (Ftl). FA assimilation through the THF cycle was verified using ^{13}C-labeled FA by analyzing ^{13}C-labeling patterns in amino

Significance

While biological utilization of one-carbon (C1) compounds has attracted much attention, previous studies have focused mainly on the utilization of CO_2. Here, we report development of *Escherichia coli* strains capable of assimilating formic acid (FA) and CO_2 through the C1 assimilation pathway, synthesizing pyruvate from FA and CO_2 by establishing the reconstructed tetrahydrofolate cycle and the reverse glycine cleavage pathway. To generate energy and redox while using less glucose, a heterologous formate dehydrogenase was introduced together with the C1 assimilation pathway. The resulting strain could utilize FA and CO_2 as sole carbon sources for sustaining growth. This work demonstrates that the combined use of the C1 assimilation pathway and formate dehydrogenase allows *E. coli* to utilize FA and CO_2 efficiently.

Author contributions: S.Y.L. designed research; J.B. performed research; J.B. and S.Y.L. analyzed data; and J.B. and S.Y.L. wrote the paper.

Conflict of interest statement: J.B. and S.Y.L. declare competing financial interests: The work described in this paper is covered by patents filed including but not limited to KR1020160180223, PCT/KR2017/013171, and KR1020170151642 and is of commercial interest. J.B. and S.Y.L. are holders of those patents.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: leesy@kaist.ac.kr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1810386115/-/DCSupplemental.

Published online September 17, 2018.

acids, including serine, glycine, and methionine (26). However, two major problems in using the THF cycle to efficiently assimilate FA by *E. coli* remained unsolved. First, the *E. coli* bifunctional enzyme Fol (encoded *folD*) possessing 5,10-methylene-THF dehydrogenase and 5,10-methylene-THF cyclohydrolase activities is allosterically inhibited by its reactant, 10-formyl THF (10-CHO-THF) (Fig. 1 and ref. 27). Second, the THF cycle requires glycine to sustain the assimilation of FA since it consumes one glycine molecule per cycle (Fig. 1). Thus, the THF cycle cannot operate continuously without an external glycine supply. To solve the glycine consumption problem, utilization of the gcv pathway was proposed in a previous review paper (20). In *E. coli*, the gcv pathway degrades glycine and produces 5,10-methylene THF (5,10-CH_2-THF), CO_2, NADH, and NH_3. However, in vitro enzyme assays suggested that the gcv reaction was reversible (28, 29), and thus it might be possible to synthesize glycine from 5,10-CH_2-THF, CO_2, NADH, and NH_3. Furthermore, a recent study reported a newly isolated uncultivated dissimilatory phosphite oxidation bacterium, *Candidatus phosphitivorax*, that could grow utilizing only CO_2 and $[HPO_3]^{2-}$ by utilizing the reverse gcv pathway (30). The metagenomics-guided analysis discovered that the bacterium lacked key genes of the known CO_2 assimilation pathways, but did contain genes involved in the gcv pathway, suggesting the potential relevance of the gcv pathway for the CO_2 assimilation process.

In this study, the aforementioned two key problems of efficiently utilizing the THF cycle in FA assimilation were solved by the reconstruction of the THF cycle and reversal of gcv pathway. The THF cycle was reconstructed using *E. coli* native serine hydroxymethyltransferase (GlyA) and three heterologous enzymes, Ftl, methenyl-THF cyclohydrolase (Fch), and methylene-THF dehydrogenase (Mtd) (Fig. 1), from *M. extorquens*. After confirming that the reconstructed THF (rTHF) cycle assimilated FA more efficiently than the previously reported THF cycle, reversal of the gcv pathway was conducted by knocking out the gcv pathway repressor gene (*gcvR*) to provide glycine from CO_2. An engineered *E. coli* strain harboring the rTHF cycle and the reverse gcv pathway was able to synthesize up to 96% of proteinogenic serine from FA and CO_2 in a medium also containing glucose. Serine was further converted to pyruvate by the native serine deaminase (Sda). After additional engineering, FA and CO_2 assimilation could be enhanced. To reduce the amount of glucose required for energy and redox generation, *Candida boidinii* formate dehydrogenase (Fdh) was additionally employed. The final engineered strain was able to sustain growth using only FA and CO_2 after glucose depletion.

Results

Use of the rTHF Cycle Improves FA Assimilation. The THF cycle assimilates FA utilizing THF as an intermediate by four enzymatic conversion steps catalyzed by the *M. extorquens* Ftl, *E. coli* Fol, and GlyA, which are encoded by *ftl*, *folD*, and *glyA*, respectively. Ftl assimilates FA together with THF, producing 10-CHO-THF. The bifunctional enzyme Fol first converts 10-CHO-THF to 5,10-methenyl THF (5,10-CH=THF) and eventually produces 5,10-CH_2-THF. Then GlyA converts 5,10-CH_2-THF and glycine to THF and serine, respectively. THF regenerated by GlyA is utilized again to form 10-CHO-THF by assimilating another FA in a new THF cycle (Fig. 1). However, the *E. coli* Fol is inefficient in FA assimilation due to allosteric inhibition by 10-CHO-THF. On the other hand, the *M. extorquens* Fch converts 10-CHO-THF to 5,10-CH=THF, and *M. extorquens* Mtd converts 5,10-CH=THF to 5,10-CH_2-THF (24). No allosteric inhibition has been reported from the studies on enzyme activities and reaction mechanisms of Fch and Mtd (31–33). Thus, the THF cycle was reconstructed by overexpressing the *M. extorquens ftl*, *fch*, and *mtd* genes encoding Ftl, Fch, and Mtd, respectively, and utilizing *E. coli* native GlyA (Fig. 1).

The rTHF cycle was capable of efficiently assimilating FA to cellular molecules and consumed more FA than the previously reported THF cycle in *E. coli*. To compare FA assimilation

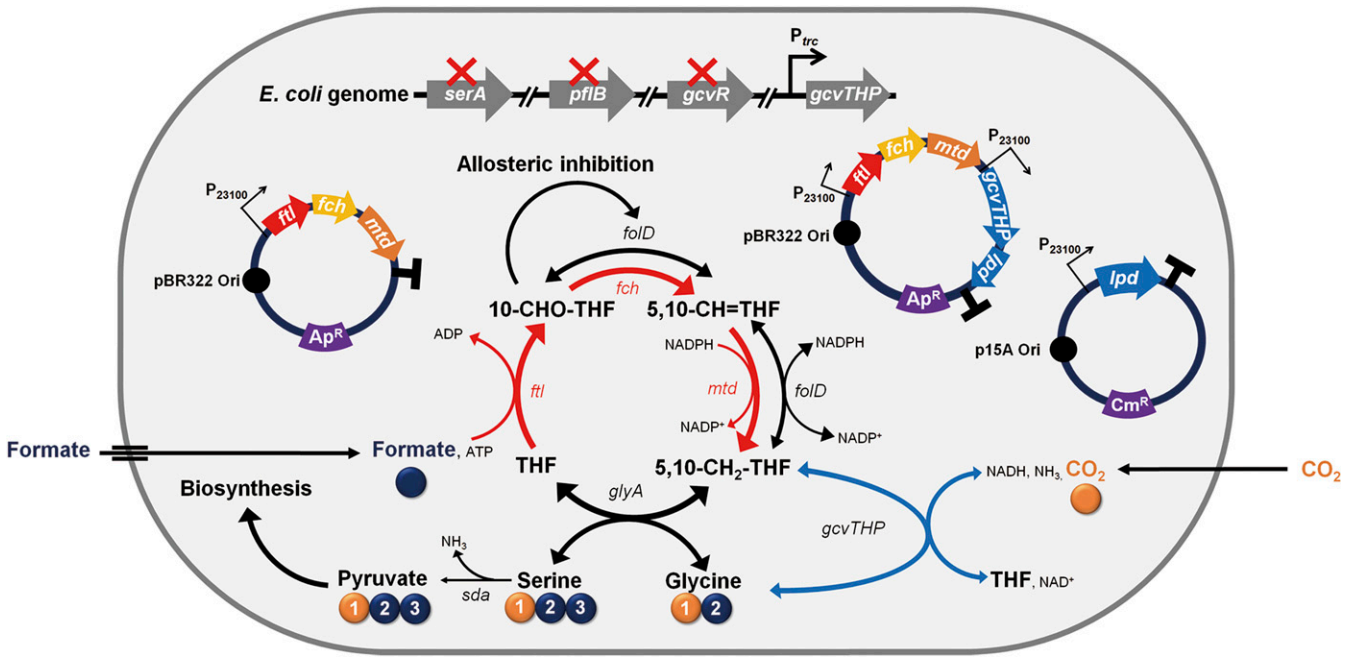

Fig. 1. The rTHF cycle and reverse gcv pathway established in *E. coli*. Red arrows indicate heterologous pathways, black arrows indicate native pathways, and blue arrows indicate overexpressed native pathways. The red Xs indicate genes knocked out. Carbon-labeling patterns in glycine, serine, and pyruvate based on FA and CO_2 carbons are indicated. Blue- and orange-colored carbons originated from FA and CO_2, respectively. Numbers in each carbon represent the carbon number in each compound. Recombinant plasmids designed and used in this study are also shown.

through the THF cycle and the rTHF cycle, the THF1, THF2, and THF4 strains equipped with the THF cycle, rTHF cycle, and THF cycle with additional *E. coli folD* overexpression, respectively, were constructed by introducing p100FTL, p100THF, and p100FTLFOL, respectively, in the DH5α strain (*SI Appendix*, Fig. S1 *B*, *C*, and *E* and Table S1). FA assimilation through the rTHF cycle was analyzed by performing ^{13}C isotope analysis and monitoring FA consumption profiles during cultivation. Since FA is toxic at levels above 4 g·L^{-1} (*SI Appendix*, Fig. S2), the MR minimal medium containing glucose was supplemented with ^{13}C-labeled sodium formate (sodium FA) at a concentration of 2.7 g·L^{-1} (the equivalent of 1.84 g·L^{-1} FA) to prevent FA toxicity. One glycine molecule is consumed per single THF cycle reaction. Thus, the culture medium was supplemented with 2 g·L^{-1} of glycine to ensure continuous operation of the THF cycle. ^{13}C isotope analysis was carried out by measuring ^{13}C labeling in proteinogenic serine and methionine. The carbon originating from FA (hereafter "FA carbon") is placed at the third carbon position of serine. Also, a portion of 5,10-CH_2-THF is used to synthesize methionine, and in this case the FA carbon is placed at the fifth carbon position of methionine (*SI Appendix*, Fig. S1 *A–E*). Thus, the relative FA assimilation flux through the rTHF cycle can be analyzed by determining the ratio of ^{13}C-labeled proteinogenic serine or ^{13}C-labeled proteinogenic methionine. In the THF1 strain, less than 10% of proteinogenic methionine and serine were labeled with ^{13}C (Fig. 2 *A* and *B*). The FA concentration at the end of cultivation did not change much from the initial concentration (Fig. 2*C*). On the other hand, most of the methionine and serine in the THF2 strain were labeled with ^{13}C (Fig. 2 *A* and *B*), and about half of the supplemented FA was consumed at the end of cultivation. In the THF4 strain, ^{13}C-labeled proteinogenic methionine and serine ratios were higher than in the THF1 strain but were lower than in the THF2 strain (Fig. 2 *A* and *B*).

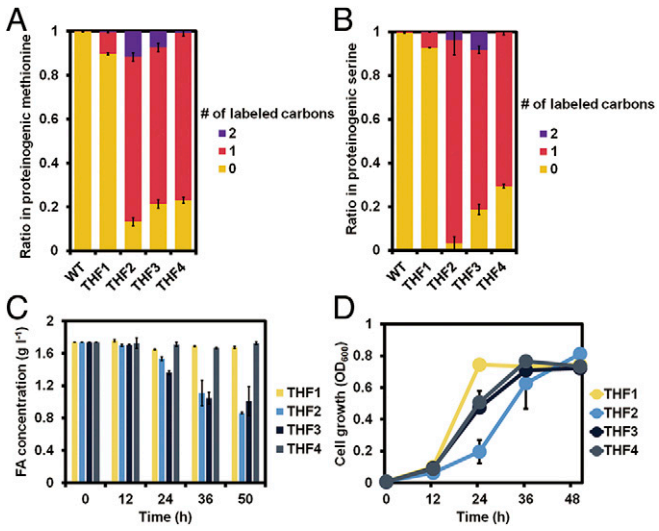

Fig. 2. ^{13}C isotope analysis of serine and methionine and FA consumption in the WT, THF1, THF2, THF3, and THF4 strains. The THF1 strain is equipped with the THF cycle, the THF2 strain is equipped with the rTHF cycle, the THF3 strain is equipped with a stronger rTHF cycle, and the THF4 strain is equipped with the THF cycle with additional overexpression of the native *E. coli folD* gene. The MR minimal medium containing glucose was supplemented with 2.7 g·L^{-1} ^{13}C-labeled sodium FA (the equivalent of 1.84 g·L^{-1} FA) and 2 g·L^{-1} glycine. (*A* and *B*) The number of labeled carbons and their ratios in proteinogenic methionine (*A*) and serine (*B*). (*C* and *D*) FA concentration profile (*C*) and growth profile (*D*) of the THF1, THF2, THF3, and THF4 strains during FA assimilation through the THF cycle. Data are shown as average values with error bars representing the SD obtained in duplicate experiments ($n = 2$).

To investigate the possibility of enhanced FA assimilation, the rTHF-cycle enzymes were further overproduced by expressing the genes under a strong *trc* promoter (*SI Appendix*, Fig. S3 *A* and *B*). The THF3 strain equipped with a stronger rTHF cycle overexpressing Ftl, Fch, and Mtd was constructed by introducing plasmid pTrcTHF in the DH5α strain (*SI Appendix*, Fig. S1*E*). However, the THF3 strain showed no significant change in the ^{13}C-labeled proteinogenic methionine and serine ratios compared with those of the THF2 strain (Fig. 2 *A* and *B*). FA consumption in the THF3 strain was also similar to that of the THF2 strain at the end of cultivation (Fig. 2*C*). Thus, further overexpression of the *ftl*, *fch*, and *mtd* genes was not effective in improving FA assimilation. Growth retardation was not observed in any of the strains (THF1, THF2, THF3, and THF4) equipped with the THF cycle or the rTHF cycle; a prolonged lag phase was observed for the THF2 strain (Fig. 2*D*).

To further enhance FA assimilation through the rTHF cycle, the effect of increased THF availability in FA assimilation was examined using the THF2 strain. THF availability was increased by the addition of folic acid, which can be converted to THF in vivo by the native dihydrofolate reductase encoded by the *folA* gene. However, no improvement in FA assimilation through the rTHF cycle was observed (*SI Appendix*, Fig. S4 *A–D*).

Increasing in Vivo Glycine Biosynthesis from FA and CO_2. In native *E. coli*, glycine is synthesized from serine by GlyA and is degraded to 5,10-CH_2-THF and CO_2 through the gcv pathway encoded by the gcv complex (*gcvTHP*). To examine possible reversal of the gcv pathway in vivo, reactants utilized in reverse reaction, including 5,10-CH_2-THF, CO_2, and NADH, were additionally supplied. The change in the metabolic flux in the gcv pathway was monitored by analyzing the ^{13}C-labeled ratio of the first carbon in proteinogenic glycine when the medium was supplemented with ^{13}C-labeled $NaHCO_3$, since CO_2 assimilated through the gcv pathway is positioned at the first carbon of glycine. However, the additional supply of reactants did not improve the reverse reaction flux (*SI Appendix*, Fig. S5 *A–C*). It has been known that *gcvTHP* genes, which encode the enzymes involved in the gcv reaction, are negatively regulated mainly by *gcvR*, which inhibits the transcription of *gcvTHP* genes (34, 35). Thus, *gcvTHP* genes were overexpressed by introducing the plasmid p184GCV, which harbors *gcvTHP* genes under the constitutive BBa 23100 promoter, into DH5α to make the GCV strain. Also, the RG1 strain was constructed by additionally knocking out the *gcvR* repressor gene in the GCV strain. The WT, GCV, and RG1 strains, which are expected to give low, medium, and high levels of *gcvTHP* expression, were cultured, and glycine biosynthesis through the reverse gcv pathway was analyzed. The ^{13}C-labeled first carbon ratios in proteinogenic glycine were gradually increased successively in the WT, GCV, and RG1 strains (*SI Appendix*, Fig. S5*D*). The ^{13}C-labeled ratio of the first carbon in proteinogenic glycine reached up to 52% of proteinogenic glycine in the RG1 strain by enhancing the reverse gcv reaction (Fig. 3*A*).

The rTHF cycle was additionally introduced in the above strains to supply 5,10-CH_2-THF from FA for synthesizing glycine from FA and CO_2 only; the THF2, THFGCV, and RG2 strains were constructed by introducing p100THF in the WT, GCV, and RG1 strains, respectively. Similar to the above results, the ^{13}C-labeled first carbon ratios of proteinogenic glycine were increased successively in the THF2, THFGCV, and RG2 strains. The ratios of ^{13}C-labeled first carbon in proteinogenic glycine increased much more upon coexpression of the rTHF cycle (*SI Appendix*, Fig. S5*E*). In the RG2 strain, the ^{13}C-labeled ratio of the first carbon in glycine reached as high as 94% of proteinogenic glycine (Fig. 3*A*). These results suggest that the RG2 strain is capable of synthesizing most of the glycine solely from FA and CO_2 by the combined use of rTHF cycle and the reverse gcv pathway.

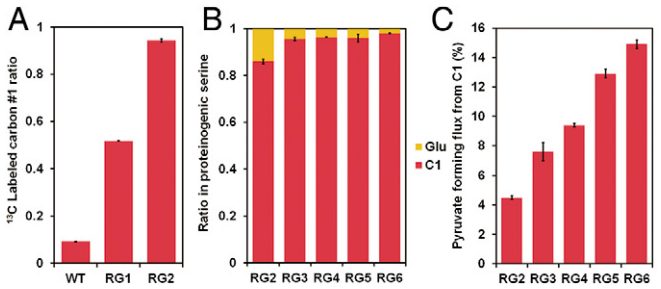

Fig. 3. ^{13}C isotope analysis of glycine-, serine-, and pyruvate-forming flux in the WT, RG1, RG2, RG3, RG4, RG5, and RG6 strains which have the rTHF and/or reverse gcv pathway. (*A*) ^{13}C isotope-labeled fraction in glycine carbon number one of the WT, RG1, and RG2 strains. ^{13}C-labeled $NaHCO_3$ (8.4 g·L^{-1}) was added to supply ^{13}C-labeled CO_2. (*B* and *C*) Ratio of proteinogenic serine synthesized only from C1 sources (FA and CO_2) (red) or glucose (yellow) in the RG2, RG3, RG4, RG5, and RG6 strains (*B*). Pyruvate-forming flux from the C1 pathway in the RG2, RG3, RG4, RG5, and RG6 strains. Pyruvate-forming flux from C1 pathway represents the percentage of carbon flux of the C1 assimilation pathway in total pyruvate-forming flux from the C1 pathway and glucose (*C*). For the RG2, RG3, and RG4 strains, 2.7 g·L^{-1} of ^{13}C-labeled sodium FA (the equivalent of 1.84 g·L^{-1} of FA) and 8.4 g·L^{-1} of ^{13}C-labeled $NaHCO_3$ were added to supply ^{13}C-labeled FA and ^{13}C-labeled CO_2, respectively. For the RG5 and RG6 strains, 2.7 g·L^{-1} of ^{13}C-labeled sodium FA (the equivalent of 1.84 g·L^{1} of FA) and 8.4 g·L^{-1} of unlabeled $NaHCO_3$ were added. Data are shown as average values with error bars representing ± SD obtained in duplicate experiments for the WT, RG1, RG2, and RG3 strains (n = 2) and in triplicate for the RG4, RG5, and RG6 strains (n = 3).

Biosynthesis of Serine and Pyruvate from FA and CO_2 Through the rTHF Cycle and Reverse Gcv Pathway. By the combined use of the rTHF cycle and reverse gcv pathway, glycine could be synthesized solely from FA and CO_2 in a glucose-containing medium. Also, serine could be synthesized from two FA molecules and one CO_2 molecule and sequentially converted to pyruvate by the *E. coli* Sda. Thus, one pyruvate (or serine) molecule was synthesized from two FA and one CO_2 molecules by consuming two ATPs, two NADPHs, and one NADH (Fig. 1). To confirm serine and pyruvate biosynthesis from FA and CO_2, the ^{13}C-labeled proteinogenic serine and alanine ratios were measured in the RG2 strain after culturing in a medium containing ^{13}C-labeled sodium FA and ^{13}C-labeled $NaHCO_3$. The pyruvate-forming flux from FA and CO_2 in a glucose-containing medium can be calculated from the ^{13}C-labeled proteinogenic alanine ratio, since alanine is synthesized from pyruvate by a one-step reaction. The RG2 strain synthesized 86% of the proteinogenic serine from FA and CO_2. On the other hand, only 4.5% of proteinogenic alanine was synthesized from FA and CO_2, which suggested that 4.5% of the total pyruvate-forming flux came from FA and CO_2 (Fig. 3*B*).

Compared with the ratio of proteinogenic serine synthesized from FA and CO_2, far less pyruvate-forming flux was observed (Fig. 3*C*). Thus, native *sdaA* was overexpressed to possibly improve pyruvate-forming flux from FA and CO_2. To examine the effect of overexpressing the *E. coli sdaA* gene, the THF2 strain was used rather than the RG2 strain because the former strain is capable of generating more serine and can clearly show the effect of *sdaA* overexpression on the serine-to-pyruvate conversion. Thus, *E. coli sdaA* was overexpressed in the THF2 strain to make the SDA strain. The THF2 and SDA strains were grown in a medium containing ^{13}C-labeled sodium FA and glycine. Unexpectedly, the SDA strain exhibited lower pyruvate-forming flux from FA and glycine than the THF2 strain; the THF2 and SDA strains contained 26% and 17% of ^{13}C-labeled proteinogenic alanine, respectively (*SI Appendix*, Fig. S6*A*).

To find the reason, the ratios of proteinogenic serine synthesized from glucose and from the THF cycle and glycine in the THF2 and SDA strains were determined (detailed results are presented in *SI Appendix*, *Text S1*). It was found that overexpression of *sdaA* gene promoted serine biosynthesis pathway flux from glucose (*SI Appendix*, Fig. S6*B*), causing a decrease in the proteinogenic serine ratio coming from the rTHF cycle. Consequently, the pyruvate-forming flux from the rTHF cycle and glycine was also decreased. Therefore, the bottleneck in improving the pyruvate-forming flux from FA and CO_2 seems to be insufficient flux through the reverse gcv pathway rather than the conversion of serine to pyruvate. When the pyruvate-forming flux from the C1 assimilation pathway became dependent only on the FA assimilation flux through the rTHF cycle by supplying excessive amounts of glycine, the pyruvate-forming flux from FA and glycine was 26% of the total pyruvate-forming flux in the THF2 strain. Glycine needs to be synthesized from CO_2 and 5,10-CH_2-THF, and half of the 5,10-CH_2-THF needs to be used for glycine biosynthesis. One mole of FA is converted to 0.5 mol of pyruvate, and thus 13% of the total pyruvate-forming flux can potentially come from FA and CO_2 if the reverse gcv flux matches the rTHF cycle flux.

Reinforcing the Reverse Gcv Pathway and Further Metabolic Engineering to Increase FA and CO_2 Assimilation into Pyruvate. To increase the pyruvate-forming flux from FA and CO_2, reverse gcv flux was enhanced by the optimized overexpression of *gcvTHP* genes. In the RG2 strain, the reverse gcv flux was enhanced by plasmid p184GCV-based overexpression of the *gcvTHP* operon. However, the use of multiple plasmids negatively affected cell growth and yielded large colony-to-colony variation. Thus, chromosomal overexpression of the *gcvTHP* operon was examined to determine whether cell growth retardation can be prevented and colony-to-colony variation can be reduced while maintaining strong glycine biosynthesis. For this, the native promoter of the *gcvTHP* operon in the *E. coli* genome was changed to a strong *trc* promoter (DH5α_GT strain) (*SI Appendix*, Table S1). Then, the rTHF cycle genes were introduced to the DH5α_GT strain to make the RG3 strain (*SI Appendix*, Table S1). The RG3 strain showed better growth with less colony-to-colony variation, improved serine biosynthesis (95.5% of proteinogenic serine) (Fig. 3*B*), and higher pyruvate-forming flux (7.6% of total pyruvate-forming flux) from FA and CO_2 compared with the RG2 strain (Fig. 3*C*). These results suggest that strong chromosomal overexpression of the *gcvTHP* operon allows better glycine biosynthesis and is more beneficial in increasing FA and CO_2 assimilation into pyruvate. To further improve pyruvate-forming flux from FA and CO_2, the pyruvate formate lyase gene (*pflB*) was knocked out to block undesired pyruvate degradation to FA. Also, the native lipoamide dehydrogenase (Lpd) gene (*lpd*) was overexpressed since Lpd supplies NADH to the enzyme complex of the reverse gcv reaction. The resulting RG4 strain, which is the RG3 strain with *pflB* knockout and *lpd* overexpression, showed similarly high serine biosynthesis (96.3% of proteinogenic serine) from FA and CO_2 (Fig. 3*B*). Also, 9.4% of the total pyruvate-forming flux came from FA and CO_2 (Fig. 3*C*). However, it should be noted that *lpd* overexpression has the negative consequence of releasing CO_2 as Lpd converts pyruvate to acetyl-CoA. Reflecting this, the acetic acid concentration increased slightly upon *lpd* overexpression. The RG3, RG3-1 (the RG3 strain with *pfl* knockout only), and RG4 strains produced 1.34 ± 0.11, 1.18 ± 0.01, and 0.83 ± 0.02 g·L^{-1} of acetic acid, respectively. The acetic acid yields on glucose were 1.25 ± 0.01, 0.92 ± 0.01, and 1.49 ± 0.04 mol acetic acid produced per mole consumed glucose, respectively (*SI Appendix*, Table S2). The RG4 strain generated 0.57 mol more CO_2 per mole consumed glucose than the RG3-1 strain, since one CO_2 molecule is generated during the conversion of pyruvate to acetyl-CoA. On the other hand, the amounts of CO_2 assimilated through the C1 assimilation

pathway per consumed glucose were calculated to be 0.23 mol and 0.25 mol in the RG4 and RG3-1 strains, respectively. These results suggest that *lpd* overexpression allows increased pyruvate-forming flux from FA and CO_2, but the total CO_2 production per consumed glucose became higher.

To further improve pyruvate biosynthesis from FA and CO_2, we designed another strategy. D-3-phosphoglycerate dehydrogenase encoded by *serA* is essential for *E. coli* growing in M9 minimal medium containing glucose since the *serA*-deleted strain cannot synthesize the 5,10-CH_2-THF needed for the biosynthesis of metabolites such as purines and methionine (5). Thus, we reasoned that the *serA*-deleted strain equipped with the C1 assimilation pathway would synthesize 5,10-CH_2-THF only through the C1 assimilation pathway, eventually allowing enhanced FA and CO_2 assimilation through the C1 assimilation pathway. The effect of *serA* knockout on FA and CO_2 assimilation was first examined using the RG4-1 strain (the *serA*-deleted RG4 strain). However, cell growth was severely retarded even in the complex Luria–Bertani medium. Thus, the effect of *serA* knockout on C1 assimilation was examined in the RG5 strain (the *serA*-deleted RG3-1 strain). It was found that knocking out *serA* was effective in improving FA and CO_2 assimilation into pyruvate. Pyruvate-forming flux from the C1 assimilation pathway reached up to 12.9% of the total pyruvate-forming flux in the RG5 strain, compared with 7.3% in the RG3-1 strain.

In the above studies, a two-plasmids-based expression system was employed. It has been known that in some cases a two-plasmids expression system can be detrimental to cell growth, depending on the genes on the plasmid. Thus, we tested a single-plasmid-based expression system. Plasmid p100THFGcvlpd was constructed to express the rTHF cycle, *gcvTHP*, and *lpd* genes from a single plasmid. The pyruvate-forming flux from FA and CO_2 in the resulting RG6 strain was increased to 14.9% of the total pyruvate-forming flux (Fig. 3*C*). In addition, the specific FA consumption rate reached 42.4 ± 3.6 mg·gram dry cell weight $(gDCW)^{-1}\cdot h^{-1}$, which was higher than the consumption rates obtained with the RG4 (20.7 ± 1.8 $mg\cdot gDCW^{-1}\cdot h^{-1}$) and RG5 (37.9 ± 5.1 $mg\cdot gDCW^{-1}\cdot h^{-1}$) strains (Table 1). The specific CO_2 consumption rate through the C1 assimilation pathway in the RG6 strain was calculated to be 20.3 $mg\cdot gDCW^{-1}\cdot h^{-1}$, which was also higher than those obtained with the RG4 (9.9 $mg\cdot gDCW^{-1}\cdot h^{-1}$) and RG5 (18.1 $mg\cdot gDCW^{-1}\cdot h^{-1}$) strains (Table 1). These results suggest that the use of single-plasmid-based overexpression of the rTHF cycle, *gcvTHP*, and *lpd* genes in the DH5α_GTPS strain allowed more efficient assimilation of FA and CO_2.

It was also confirmed by ^{13}C isotope analysis that pyruvate synthesized through the C1 assimilation pathway was converted to other cellular molecules by the native *E. coli* metabolism. Valine, leucine, isoleucine, and glutamate were selected as the target cellular metabolites as they are synthesized utilizing pyruvate and/or acetyl-CoA (*SI Appendix*, Fig. S7*A*). The ^{13}C isotope-labeling patterns observed were as expected (*SI Appendix*, Fig. S7*B*), and the ^{13}C-labeled proteinogenic amino acids ratios (*SI Appendix*, Fig. S7 *C–F*) were consistent with those expected from the ^{13}C-labeled pyruvate ratio. Detailed explanations are presented in *SI Appendix*, *Text S2*.

When the reaction is reversible, it is important to examine both the forward and reverse reaction fluxes. The major reversible reaction that affects the C1 assimilation pathway flux is the gcv reaction. The fluxes of forward and reverse gcv reactions in the RG6 strain were examined from the number of ^{13}C-labeled carbons, the ^{13}C-labeled position, and the ratio of proteinogenic glycine (detailed methods and procedures are presented in *SI Appendix*, *Text S3*). The flux of the reverse gcv reaction increased as the intracellular glycine pool decreased, while the flux of the forward gcv reaction decreased as intracellular glycine pool decreased (except for feeding 0.7–0.5 $g\cdot L^{-1}$ of 2-^{13}C–labeled glycine). These labeling experiments suggested that the forward gcv reaction flux must be very low (e.g., lower than that obtained with 0.01 $g\cdot L^{-1}$ of 2-^{13}C–labeled glycine supplementation) (*SI Appendix*, Fig. S8) because our experiments in assimilating FA and CO_2 to form pyruvate were performed without glycine supplementation. Thus, in the RG6 strain the reverse gcv pathway operates like an almost irreversible reaction, and labeling patterns are little, if at all, affected by the potential reversibility of the reaction.

Introduction of Fdh Further Increased FA Utilization with Less Glucose Consumption. Although thus far we were able to achieve the most efficient FA and CO_2 assimilation in engineered *E. coli* equipped with the rTHF cycle and reverse gcv pathway, the need to use relatively large amounts of glucose remained a problem. Cells still need to use glucose to generate energy and redox, although theoretically all the carbons can come from FA and CO_2. Thus, we reasoned that the generation of energy and redox from FA would reduce the amount of glucose needed. To achieve this goal, a heterologous NAD^+-utilizing Fdh, which converts FA to CO_2 while producing NADH, was introduced into *E. coli*. Two Fdhs from *M. extorquens* and *C. boidinii* reported to have outstanding enzyme activities (36, 37) were chosen in this study. Plasmids p184FDH and p184FDHcbo expressing the *fdh* genes of *M. extorquens* and *C. boidinii*, respectively, were transformed into the RG6 strain to construct the RG7 and RG8 strains,

Table 1. Specific glucose, FA, and CO_2 consumption rates and C1 utilization ratios of the RG4, RG5, RG6, and RG8 strains

Strain	Specific glucose consumption rate, $mg\cdot gDCW^{-1}\cdot h^{-1}$	Specific FA consumption rate, $mg\cdot gDCW^{-1}\cdot h^{-1}$	Specific CO_2 consumption rate*, $mg\cdot gDCW^{-1}\cdot h^{-1}$	C1 utilization ratio†, mol/mol
RG4	249.8 ± 31.4	20.7 ± 1.8	9.9	0.076 ± 0.015
RG5	319.4 ± 5.1	37.9 ± 5.1	18.1	0.104 ± 0.013
RG6	319.7 ± 1.9	42.4 ± 3.6	20.3	0.115 ± 0.008
RG8	232.2 ± 6.0	28.6 (by C1 pathway)‡ 80.3 (by Fdh)‡	13.9	0.076§
RG8¶	370.2	31.0 (by C1 pathway)‡ 114.6 (by Fdh)‡	14.9	0.076§

*The specific CO_2 consumption rate is not a measured value. It is calculated from the specific FA consumption rate using the stoichiometry of the C1 assimilation pathway; the C1 assimilation pathway assimilates two FA molecules and one CO_2 molecule, thus, the specific CO_2 consumption rate was first calculated as mole basis and later converted to mass basis.

†The C1 utilization ratio is calculated by (carbon moles from FA and CO_2)/(carbon moles from FA, CO_2, and glucose). The theoretical maximum value of the C1 utilization ratio is 0.627.

‡The RG8 strain consumes FA through the C1 assimilation pathway (C1 pathway) as well as Fdh. Thus, the specific FA consumption rates through the C1 pathway and Fdh are indicated separately.

§The C1 utilization ratio is a predicted value. The C1 utilization ratio of the RG8 strain is assumed to be similar to that of the RG4 strain because these two strains showed a similar percentage of the total pyruvate-forming flux coming from FA and CO_2.

¶Values were obtained in the bioreactor cultivation.

respectively. The RG7 strain exhibited instability and larger colony-to-colony variations, but the RG8 strain did not. The instability of the RG7 strain seems to be due to the inconsistent expression of a large *M. extorquens* Fdh complex (~185 kDa) comprising four subunits encoded by an operon of large size (5.3 kb). On the other hand, the *C. boidinii* Fdh is a homodimeric protein (~74 kDa) encoded by a 1.1-kb gene (38). Thus, we used the RG8 strain for further experiments.

To confirm that generation of energy and redox from FA would reduce the amount of glucose needed and potentially sustain cell growth only from FA and CO_2, three independent flask cultures were performed in two stages; this was because FA was completely consumed during the flask cultivation as described below. At the first stage of cultivation, the RG8 strain was cultivated in M9 minimal medium supplemented with 5 $g{\cdot}L^{-1}$ glucose and 3.7 $g{\cdot}L^{-1}$ sodium FA (the equivalent of 2.5 $g{\cdot}L^{-1}$ FA). The RG8 strain consumed FA rapidly, and FA was completely depleted when glucose was depleted (Fig. 4*A*). In contrast, the RG6 strain consumed FA more slowly, with FA remaining after glucose depletion (Fig. 4*B*). Flask culture of the RG8 strain showed specific glucose and FA consumption rates of 232.2 ± 6.0 and 108.9 ± 3.2 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$, respectively. Also, the pyruvate-forming flux from FA and CO_2 was measured after supplementation with ^{13}C-labeled FA as described earlier for the RG2–RG6 strains (Fig. 3*C*). In the RG8 strain 9.6% of the total pyruvate-forming flux came from FA and CO_2, which was similar to the percentage (9.4%) in the RG4 strain. However, it was lower than the percentage (14.9%) in the RG6 strain due to the consumption of FA by Fdh, which lowered the FA flux toward C1 assimilation pathway. It was assumed that the C1 utilization ratio of the RG8 strain would be similar to that of the RG4 strain (0.076) (Table 1) because these two strains showed similar percentages of the total pyruvate-forming flux coming from FA and CO_2. The specific CO_2 consumption rate of the RG8 strain was calculated to be 13.9 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$. The C1 assimilation pathway and Fdh reaction consumed 26.3 wt% and 73.7 wt% of FA, respectively, suggesting that three times more FA was consumed for energy and redox generation. Nonetheless, the specific FA consumption rate was increased by 160%, while the specific glucose consumption rate decreased by 27% compared with the RG6 strain. Thus, the generation of energy and redox from FA by employing the Fdh reaction indeed reduced the amount of glucose needed by cells. These results led us to examine whether cells can sustain growth utilizing only FA and CO_2.

After the depletion of the initially supplied glucose, the second stage of cultivation was started from the P1 point (Fig. 4*A*). To examine whether cells can sustain growth utilizing only FA and CO_2, the M9 minimal medium was manually supplemented with FA when FA was depleted during the cultivation. Initially, the M9 minimal medium was supplemented with 3.7 $g{\cdot}L^{-1}$ of ^{13}C-labeled sodium FA (the equivalent of 2.5 $g{\cdot}L^{-1}$ of FA) at the P1 point and again at the P2 point when the added ^{13}C-labeled FA had been consumed (Fig. 4*A*). Flask culture was ended at the P3 point. Cells were sampled at the P1, P2, and P3 points; the P1 and P2 samples were taken right before supplementation with ^{13}C-labeled FA (Fig. 4*A*). FA assimilation into amino acids synthesized from the C1 assimilation pathway (glycine and serine), TCA cycle (glutamate and isoleucine), pentose phosphate pathway (phenylalanine), and pyruvate (alanine, valine and leucine) (Fig. 4*C*) were analyzed by ^{13}C isotope analysis. The ^{13}C-labeled proteinogenic amino acids ratios at the P1 point should be zero, but some showed slight labeling. This was due to the incomplete resolution of compounds, hydrogen abstraction, and incomplete resolution in the time or *m/z* domain, causing inaccurate isotopomer ratios in proteinogenic amino acids (Fig. 4*C* and ref. 39). At the P2 point, the ^{13}C-labeled ratios of the above proteinogenic amino acids were increased compared with those at the P1 point (Fig. 4*C*). The OD_{600} at the P2 point was 10% higher than that at the P1 point. Cells taken at the P3 point showed slightly higher ^{13}C-labeled proteinogenic amino acid ratios than those at the P2 point, while the OD_{600} at P3 was slightly lower than that at the P2 point; however, the difference between the OD_{600} values at P2 and P3 is within the error range in flask cultivation (see below for bioreactor cultivation giving clear results). These results suggest that cells were able to sustain growth on FA and CO_2 as the sole carbon sources by the combined use of the C1 assimilation pathway and *C. boidinii* Fdh even after glucose was depleted.

To examine whether continuous gas supplementation with an air–CO_2 mixture would enhance cell growth and FA consumption, the RG8 strain was cultivated in a 1.3-L bioreactor (initial working volume of 300 mL) with continuous sparging of the air–CO_2 mixture (*SI Appendix*, *Materials and Methods*), since both of O_2 and CO_2 are required for cell growth (to produce ATP through aerobic respiration and to synthesize glycine through the C1 assimilation pathway, respectively). The bioreactor cultivation was also performed in two stages, as in flask cultivation described above. At the first stage, the RG8 strain was cultivated in M9 minimal medium supplemented with 5 $g{\cdot}L^{-1}$ glucose and 3.7 $g{\cdot}L^{-1}$ sodium FA (the equivalent of 2.5 $g{\cdot}L^{-1}$ FA). Cell growth of the RG8 strain was improved in bioreactor cultivation, resulting in a shorter lag phase than in flask cultivation (Fig. 4 *A* and *D*). The RG8 strain consumed FA rapidly, and FA was almost depleted when glucose was depleted (Fig. 4*D*). In contrast, the bioreactor-cultured control RG6 strain consumed FA much more slowly, and FA remained present after glucose depletion (Fig. 4*E*). Bioreactor cultivation of the RG8 strain showed the specific glucose and FA consumption rates of 370.2 and 145.6 mg $gDCW^{-1}{\cdot}h^{-1}$, respectively, which were 59% and 34% higher than those obtained in flask cultivation. The specific CO_2 consumption rate was calculated to be 14.9 mg $gDCW^{-1}{\cdot}h^{-1}$. The C1 assimilation pathway and Fdh reaction consumed 21.3 wt% and 78.7 wt% of FA, respectively. The percentage of FA consumed by the Fdh reaction was slightly higher than that observed in flask cultivation, suggesting that supplementation with the air–CO_2 mixture improved FA consumption by the Fdh reaction.

After the depletion of initially supplied glucose, the second stage of bioreactor cultivation was started from the F1 point (Fig. 4*D*). The M9 minimal medium was supplemented with a mixture of ^{13}C-labeled FA and unlabeled FA (in a 1:1 molar ratio) at the F1 point; the FA mixture was used in the bioreactor experiment to reduce the cost of the rather expensive ^{13}C-labeled FA. Then the FA mixture was supplemented once more when FA was almost depleted. Unlabeled FA was supplemented at the F2 point and once more afterward, as indicated in the fermentation profile (Fig. 4*D*); this experiment supplying unlabeled sodium FA was designed to examine decreased ^{13}C labeling in proteinogenic amino acids, which would clearly prove that FA was consumed to sustain the growth of the RG8 strain in the absence of glucose. Samples for ^{13}C isotope analysis were taken at the F1, F2, and F3 points; the F1 and F2 samples were taken right before FA mixture or unlabeled FA was supplemented (Fig. 4*D*).

The ^{13}C-labeled ratios of proteinogenic amino acids at the F1 point should be zero, but, as in the flask cultivation results, some amino acids showed 1–2% of ^{13}C-labeled proteinogenic amino acids ratios resulting from inaccurate isotopomer ratios in proteinogenic amino acids (see above and Fig. 4*C*). Compared with the samples at the F1 point, the ^{13}C-labeled ratios of proteinogenic amino acids were slightly higher in the samples at the F2 point (Fig. 4*C*). The OD_{600} at the F2 point was 9% higher than that at the F1 point (Fig. 4*D*). The ^{13}C-labeled proteinogenic amino acids ratios were not as high as those in flask cultivation, since the FA mixture rather than pure ^{13}C-labeled FA was used. Also, all the ^{13}C-labeled ratios of proteinogenic amino acids, except for glutamate (which did not change), were decreased at F3 compared with those at the F2 point (Fig. 4*C*); the absence of change in glutamate labeling is likely due to the initial

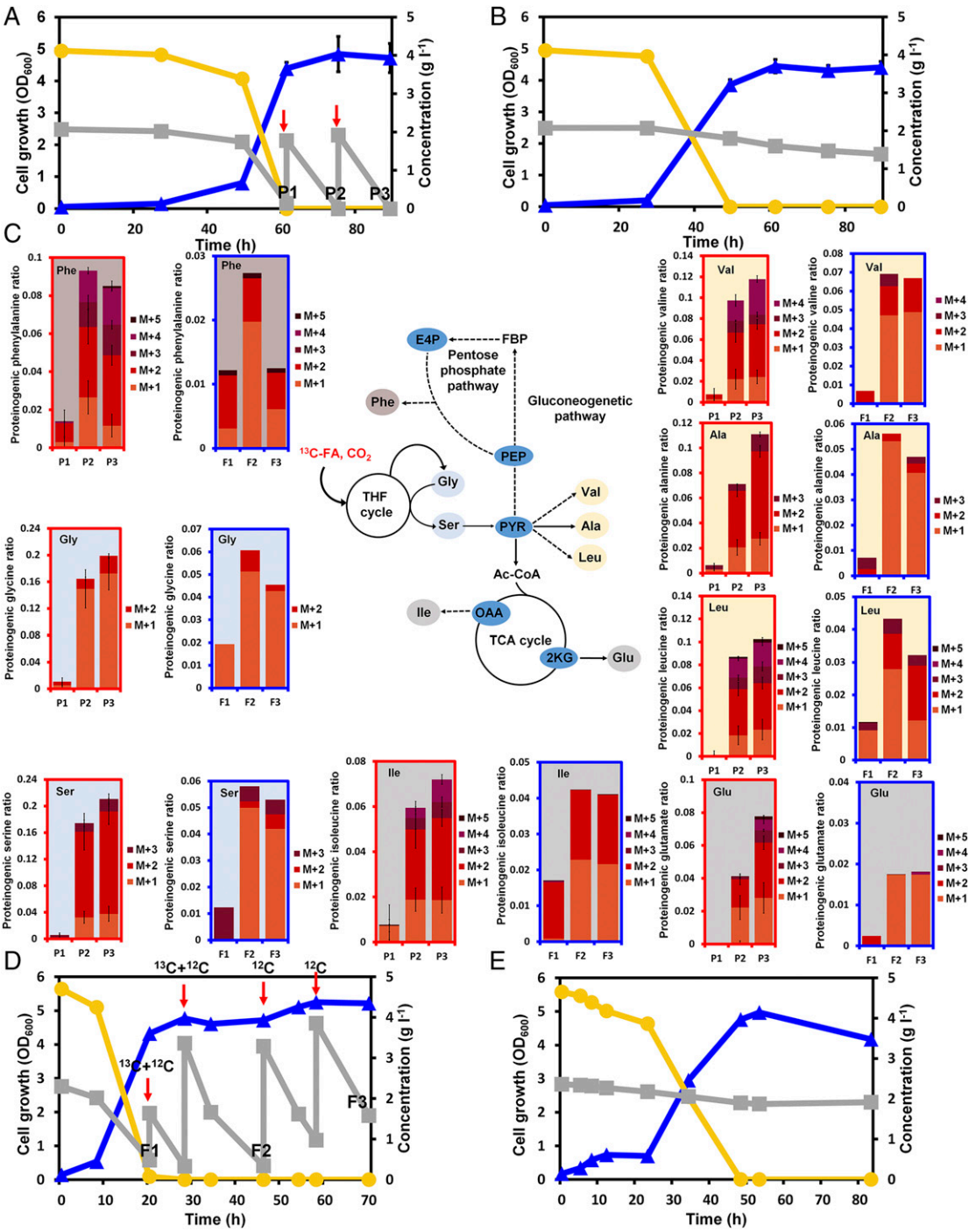

Fig. 4. (*A* and *B*) Cell growth and glucose and FA consumption profiles of the RG8 strain (*A*) and RG6 strain (*B*) in flask cultures. Cells were grown using M9 minimal medium initially supplemented with 5 g·L^{-1} of glucose and 3.7 g·L^{-1} of sodium FA (the equivalent of 2.5 g·L^{-1} of FA). For the RG8 strain, the medium was supplemented with 3.7 g·L^{-1} of ^{13}C-labeled sodium FA (the equivalent of 2.5 g·L^{-1} of FA) at the points indicated by the red arrows. Cells for ^{13}C isotope analysis were taken at the P1, P2, and P3 points. (*C*, *Center*) Biosynthesis route from FA and CO_2 to various amino acids during cell growth from only FA and CO_2. Solid lines and arrows represent single-step pathways; dashed lines and arrows represent multiple-step pathways. 2KG, alpha-ketoglutarate; Ac-CoA, acetyl-CoA; Ala, alanine; E4P, erythrose 4-phosphate; FBP, fructose 1,6-bisphosphate; Glu, glutamate; Gly, glycine; Ile, isoleucine; Leu, leucine; OAA, oxaloacetate; PEP, phosphoenolpyruvate; Phe, phenylalanine; PYR, pyruvate; Ser, serine. (*Left*, *Right*, and *Bottom Row*) ^{13}C-labeled proteinogenic amino acids ratios at three different points obtained from flask cultivation (red-outlined graphs) and bioreactor cultivation (blue-outlined graphs). Data from the flask cultivation are shown as average values with error bars representing ± SD obtained in triplicate (n = 3). (*D* and *E*) Cell growth and glucose and FA consumption profiles from the bioreactor cultivations of the RG8 strain (*D*) and the RG6 strain (*E*). Cells were grown in a bioreactor using M9 minimal medium initially supplemented with 5 g·L^{-1} glucose and 3.7 g·L^{-1} sodium FA (the equivalent of 2.5 g·L^{-1} FA). For the RG8 strain, the medium was supplemented with a mixture of ^{13}C-labeled FA and unlabeled FA (1:1 ratio; mole/mole) when the initially supplemented FA was almost depleted, and then the FA mixture was supplemented once again. After that, the medium was supplemented twice with unlabeled FA at the indicated time points. Cells for ^{13}C isotope analysis were taken at the F1, F2, and F3 points. Blue lines with closed triangles indicate cell growth. Yellow lines with closed circles indicate glucose concentration. Gray lines with closed squares indicate FA concentration.

low ^{13}C labeling (less than 2% at the F2 point). The OD_{600} was 11% higher at the F3 point than at the F2 point (Fig. 4*D*). Except for glutamate, all the ^{13}C-labeled ratios of proteinogenic amino acids were increased by ^{13}C-labeled FA supplementation and were decreased by unlabeled FA supplementation. On the other hand, cell growth in the RG8 strain decreased when not supplemented with additional FA after glucose depletion (*SI Appendix*, Fig. S9*A*). Thus, it can be concluded that cell growth after glucose depletion was due to carbon and energy metabolism utilizing only FA and CO_2.

Next, we wanted to confirm that cell growth of the RG8 strain in M9 minimal medium without glucose was truly due to the use of FA and CO_2 rather than cellular materials (glycogen, among others.). The RG8 strain was cultivated in M9 minimal medium containing glucose, washed with fresh M9 minimal medium, and transferred to M9 minimal medium with or without FA supplementation. When cells were cultured in M9 minimal medium with FA supplementation, the OD_{600} was increased by 57% in 66 h, consuming 1.65 $g{\cdot}L^{-1}$ FA (*SI Appendix*, Fig. S9*B*). On the other hand, in cells cultivated in M9 minimal medium without FA supplementation the OD_{600} increased by 21% in 25 h, after which no more cell growth was observed (*SI Appendix*, Fig. S9*B*). For comparison, in cells cultivated in M9 medium containing FA the OD_{600} increased by 32% in 25 h (the time point when cells stopped growing in the absence of FA), suggesting that cell growth occurred using FA in addition to cellular materials. Also, as mentioned above, cell growth continued after 25 h, and the OD_{600} increased by 57% in 66 h. These results clearly suggest that cells grew using FA and CO_2 (in addition to cellular materials) as the sole carbon and energy sources.

To further elaborate on some growth of cells in the absence of FA, glycogen concentrations in the RG8 strain cultivated with and without FA were measured. However, glycogen was not detected in either case. It should be noted that glycogen accumulation was not observed unless the glycogen biosynthetic pathway was amplified (40), even though the culture condition is not exactly the same as ours. Thus, the 21% increase in OD_{600} seems to be due to the turnover of RNA, proteins, and other cellular materials, including cellular metabolites. In summary, the growth of the RG8 strain observed in a bioreactor in the presence of FA and CO_2 but without glucose was truly due to the consumption of FA and CO_2 as the sole carbon and energy sources (in addition to some use of cellular material turnover).

Discussion

Although previous studies demonstrated the use of THF cycle for FA assimilation, the efficiencies of FA assimilation have been rather low. In this study, we report the development of an engineered *E. coli* strain equipped with the rTHF cycle and reverse gcv pathway for efficient FA and CO_2 assimilation into serine and pyruvate. The THF cycle was reconstructed using *E. coli* native GlyA and three heterologous enzymes, Ftl, Fch, and Mtd, from *M. extorquens*. After confirming that the rTHF cycle assimilated FA more efficiently than the previously reported THF cycle, a reverse gcv pathway was established by knocking out *gcvR* and overexpressing *gcvTHP* genes to synthesize glycine from CO_2. An engineered *E. coli* strain equipped with the rTHF cycle and reverse gcv pathway was able to synthesize up to 96% of proteinogenic serine from FA and CO_2 in a medium also containing glucose. Serine was further converted to pyruvate by the native Sda. After additional engineering, the pyruvate-forming flux from FA and CO_2 was increased up to 14.9% of total pyruvate-forming flux. Establishment of the *C. boidinii* Fdh reaction reduced the amount of glucose required for energy and redox generation. The final engineered RG8 strain showed specific glucose, FA, and CO_2 consumption rates of 370.2, 145.6, and 14.9 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$, respectively. It seemed possible that the RG8 strain was capable of sustaining growth from only FA and CO_2 in the absence of glucose.

Then, the C1 consumption efficiency of the strain developed in this study was compared with those of other studies. The total specific C1 consumption rate was determined by adding the specific FA consumption rate and specific CO_2 consumption rate; since FA can be considered as a liquefied form of CO_2, FA and CO_2 are assumed to be equivalent. The RG6 strain showed a total specific C1 consumption rate of 62.7 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$, which was 2.6 and 2.8 times higher than the specific CO_2 consumption rates of cyanobacteria [24.1·mg $gDCW^{-1}{\cdot}h^{-1}$ (41)] and *E. coli* equipped with the CBB cycle [22.5 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$ (11)], respectively. Consequently, the RG6 strain equipped with the C1 assimilation pathway assimilates FA and CO_2 more efficiently than cyanobacteria or *E. coli* equipped with the CBB cycle. In the case of the RG8 strain showing reduced glucose dependence, the total specific C1 consumption rate was 160.5 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$ (45.9 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$ for the C1 assimilation pathway and 114.6 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$ for the Fdh reaction), which shows much higher C1 consumption efficiency than in previous reports.

While we were revising this paper, two studies on the use of the THF cycle and reverse gcv pathway to synthesize pyruvate or serine from FA and CO_2 appeared (42, 43). In one study (42), *Clostridium ljungdahlii* Ftl, Fch, and Fol were used to construct the THF cycle. In addition, *gcvTHP* and *sdaA* were overexpressed, and native *serA* was deleted. However, the efficiency of FA and CO_2 assimilation into serine or pyruvate (~10% of proteinogenic serine was synthesized only from FA and CO_2) was much lower than that reported here (~98% of proteinogenic serine was synthesized only from FA and CO_2), likely because of the lower efficiencies of *C. ljungdahlii* enzymes in FA assimilation. In the other report (43), the same *M. extorquens* Ftl, Fch, and Mtd were utilized to construct the THF cycle. In addition, *gcvTHP* and *lpd* were overexpressed, and native *serA* was deleted. The efficiency of serine synthesis from FA and CO_2 (~90% of proteinogenic serine) was high but was slightly lower than that reported in our study (98% of proteinogenic serine). It seems that the additional metabolic engineering strategies, including *gcv* operon repressor (*gcvR*) and *pflB* knockout, the change of the native promoter of *gcv* operon to a strong *trc* promoter, and overexpressing *ftl*, *fch*, *mtd*, *gcvTHP*, and *lpd* using a single plasmid, were effective in improving serine synthesis from FA and CO_2 in our study.

Theoretically, pyruvate can be synthesized from only FA and CO_2 if the C1 assimilation pathway continues to function. This was why glucose was used to supply the cellular energy and redox needed. To synthesize 1 mol of pyruvate from 2 mol of FA and 1 mol of CO_2, 0.297 mol of glucose is required (the detailed calculation is presented in *SI Appendix*, *Text S4*). Thus, the C1 utilization ratio can reach up to 0.627, since 3 mol of carbons come from FA and CO_2, and 1.782 mol of carbons come from glucose. However, the highest C1 utilization ratio achieved with the RG6 strain (0.115) is still much lower than the theoretical maximum (0.627) because the C1 assimilation pathway is much less efficient than the glycolytic pathway, as shown by the significantly lower specific C1 consumption rate (62.7 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$) compared with the specific glucose consumption rate (319.7 $mg{\cdot}gDCW^{-1}{\cdot}h^{-1}$). This is the reason for the large gap between the percentage of serine biosynthesis (98%, corresponding to 100% as the labeled FA was 98% ^{13}C-labeled) and pyruvate-forming flux (14.9%) from the C1 assimilation pathway; a large portion of pyruvate is produced from glycolysis, while serine is synthesized only from FA and CO_2 (for example in the RG5 and RG6 strains).

To develop an *E. coli* strain capable of growing solely from FA and CO_2 without glucose supplementation, further improvement of FA and CO_2 assimilation into pyruvate is needed. In our study, the C1 assimilation pathway flux has been improved mainly by enhancing the reverse gcv pathway. Thus, future study is needed to further improve FA and CO_2 assimilation into pyruvate by enhancing the FA and CO_2 assimilation efficiency of the rTHF cycle. One possible approach is achieving balanced

overexpression of the *M. extorquens ftl*, *fch*, and *mtd* genes, since *fch* and *mtd* were not overexpressed as much as *ftl* (*SI Appendix*, Fig. S3 *A* and *B*). Also, it is known that enzymes involved in the reverse gcv reaction have low affinities toward CO_2 and NH_3 (21, 28, 29), which are the reactants of the reverse reaction. Thus, improvement of enzyme affinities toward CO_2 and NH_3 can be another promising solution to enhance FA and CO_2 assimilation by further improving the reverse gcv pathway flux. In addition, the *E. coli* central carbon metabolism needs to be modified for better growth on FA and CO_2. For example, improving metabolic flux from pyruvate to phosphoenolpyruvate might be a solution to better synthesize intermediates of upper-glycolysis utilized for synthesizing various cellular molecules, including nucleotides, amino acids, and cell walls.

In conclusion, we report the development of engineered *E. coli* strains capable of efficiently utilizing FA and CO_2 by introducing the rTHF cycle and reverse gcv pathway. The possibility of reducing the amount of glucose needed was also demonstrated by establishing the Fdh reaction. The strategies and *E. coli* strains developed in this study will be useful for further advancing C1 biorefinery toward the production of chemicals and materials from FA and CO_2.

Materials and Methods

All the materials and methods conducted in this study are detailed in *SI Appendix, Materials and Methods*, including reagents, media compositions, plasmid construction, gene knockout and native promoter change of the genomic DNA, protein expression experiments, FA assimilation analyses through the rTHF cycle, analysis of FA and CO_2 assimilation using ^{13}C isotopes, flux analysis of the serine synthesis pathway, the FA tolerance test, determination of specific consumption rates of glucose, FA, and CO_2, bioreactor cultivation, and other analytical procedures. The data supporting the findings of this study are available in *SI Appendix* and Datasets S1–S3.

ACKNOWLEDGMENTS. We thank Professor H. U. Kim, Dr. J. Y. Ryu, J. S. Cho, and J. H. Ahn at the Systems Metabolic Engineering and Systems Healthcare Cross-Generation Collaborative Laboratory of the Korea Advanced Institute of Science and Technology for valuable discussions during manuscript preparation, and Professor Jinwon Lee at Sogang University for reviewing our manuscript and giving valuable advice. This work was supported by the C1 Gas Refinery Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT Grant NRF-2016M3D3A1A01913250.

1. Ho S-H, Chen C-Y, Lee D-J, Chang J-S (2011) Perspectives on microalgal CO_2-emission mitigation systems–A review. *Biotechnol Adv* 29:189–198.
2. Kumar A, et al. (2010) Enhanced $CO(_2)$ fixation and biofuel production via microalgae: Recent developments and future directions. *Trends Biotechnol* 28:371–380.
3. Machado IM, Atsumi S (2012) Cyanobacterial biofuel production. *J Biotechnol* 162: 50–56.
4. Walker BJ, VanLoocke A, Bernacchi CJ, Ort DR (2016) The costs of photorespiration to food production now and in the future. *Annu Rev Plant Biol* 67:107–129.
5. Bouzon M, et al. (2017) A synthetic alternative to canonical one-carbon metabolism. *ACS Synth Biol* 6:1520–1533.
6. Schwander T, Schada von Borzyskowski L, Burgener S, Cortina NS, Erb TJ (2016) A synthetic pathway for the fixation of carbon dioxide *in vitro*. *Science* 354:900–904.
7. Siegel JB, et al. (2015) Computational protein design enables a novel one-carbon assimilation pathway. *Proc Natl Acad Sci USA* 112:3704–3709.
8. Bogorad IW, Lin T-S, Liao JC (2013) Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature* 502:693–697.
9. Shih PM, Zarzycki J, Niyogi KK, Kerfeld CA (2014) Introduction of a synthetic CO_2-fixing photorespiratory bypass into a cyanobacterium. *J Biol Chem* 289:9493–9500.
10. Antonovsky N, et al. (2016) Sugar synthesis from CO_2 in *Escherichia coli*. *Cell* 166: 115–125.
11. Gong F, et al. (2015) Quantitative analysis of an engineered CO_2-fixing *Escherichia coli* reveals great potential of heterotrophic CO_2 fixation. *Biotechnol Biofuels* 8:86.
12. Bonacci W, et al. (2012) Modularity of a carbon-fixing protein organelle. *Proc Natl Acad Sci USA* 109:478–483.
13. Zhuang Z-Y, Li S-Y (2013) Rubisco-based engineered *Escherichia coli* for in situ carbon dioxide recycling. *Bioresour Technol* 150:79–88.
14. Li H, et al. (2012) Integrated electromicrobial conversion of CO_2 to higher alcohols. *Science* 335:1596.
15. Jabeen G, Farooq R (2016) Bio-electrochemical synthesis of commodity chemicals by autotrophic acetogens utilizing CO_2 for environmental remediation. *J Biosci* 41: 367–380.
16. Jeletic MS, Mock MT, Appel AM, Linehan JC (2013) A cobalt-based catalyst for the hydrogenation of CO_2 under ambient conditions. *J Am Chem Soc* 135:11533–11536.
17. Innocent B, et al. (2009) Electro-reduction of carbon dioxide to formate on lead electrode in aqueous medium. *J Appl Electrochem* 39:227–232.
18. Joó F (2008) Breakthroughs in hydrogen storage–Formic acid as a sustainable storage material for hydrogen. *ChemSusChem* 1:805–808.
19. Ahn JH, Bang J, Kim WJ, Lee SY (2017) Formic acid as a secondary substrate for succinic acid production by metabolically engineered *Mannheimia succiniciproducens*. *Biotechnol Bioeng* 114:2837–2847.
20. Bar-Even A (2016) Formate assimilation: The metabolic architecture of natural and synthetic pathways. *Biochemistry* 55:3851–3863.
21. Bar-Even A, Noor E, Flamholz A, Milo R (2013) Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim Biophys Acta* 1827:1039–1047.
22. Cotton CA, Edlich-Muth C, Bar-Even A (2018) Reinforcing carbon fixation: CO_2 reduction replacing and supporting carboxylation. *Curr Opin Biotechnol* 49:49–56.
23. Kikuchi G (1973) The glycine cleavage system: Composition, reaction mechanism, and physiological significance. *Mol Cell Biochem* 1:169–187.
24. Crowther GJ, Kosály G, Lidstrom ME (2008) Formate as the main branch point for methylotrophic metabolism in *Methylobacterium extorquens* AM1. *J Bacteriol* 190: 5057–5062.
25. Yishai O, Lindner SN, Gonzalez de la Cruz J, Tenenboim H, Bar-Even A (2016) The formate bio-economy. *Curr Opin Chem Biol* 35:1–9.
26. Yishai O, Goldbach L, Tenenboim H, Lindner SN, Bar-Even A (2017) Engineered assimilation of exogenous and endogenous formate in *Escherichia coli*. *ACS Synth Biol* 6:1722–1731.
27. Dev IK, Harvey RJ (1978) A complex of N5,N10-methylenetetrahydrofolate dehydrogenase and N5,N10-methenyltetrahydrofolate cyclohydrolase in *Escherichia coli*. Purification, subunit structure, and allosteric inhibition by N10-formyltetrahydrofolate. *J Biol Chem* 253:4245–4253.
28. Fujiwara K, Okamura-Ikeda K, Motokawa Y (1984) Mechanism of the glycine cleavage reaction. Further characterization of the intermediate attached to H-protein and of the reaction catalyzed by T-protein. *J Biol Chem* 259:10664–10668.
29. Fujiwara K, Motokawa Y (1983) Mechanism of the glycine cleavage reaction. Steady state kinetic studies of the P-protein-catalyzed reaction. *J Biol Chem* 258: 8156–8162.
30. Figueroa IA, et al. (2018) Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural CO_2 fixation pathway. *Proc Natl Acad Sci USA* 115:E92–E101.
31. Marx CJ, Laukel M, Vorholt JA, Lidstrom ME (2003) Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. *J Bacteriol* 185:7169–7175.
32. Pomper BK, Vorholt JA, Chistoserdova L, Lidstrom ME, Thauer RK (1999) A methenyl tetrahydromethanopterin cyclohydrolase and a methenyl tetrahydrofolate cyclohydrolase in *Methylobacterium extorquens* AM1. *Eur J Biochem* 261:475–480.
33. O'Brien WE, Brewer JM, Ljungdahl LG (1973) Purification and characterization of thermostable 5,10-methylenetetrahydrofolate dehydrogenase from *Clostridium thermoaceticum*. *J Biol Chem* 248:403–408.
34. Ghrist AC, Stauffer GV (1995) Characterization of the *Escherichia coli gcvR* gene encoding a negative regulator of *gcv* expression. *J Bacteriol* 177:4980–4984.
35. Wilson RL, Stauffer LT, Stauffer GV (1993) Roles of the *GcvA* and *PurR* proteins in negative regulation of the *Escherichia coli* glycine cleavage enzyme system. *J Bacteriol* 175:5129–5134.
36. Laukel M, Chistoserdova L, Lidstrom ME, Vorholt JA (2003) The tungsten-containing formate dehydrogenase from *Methylobacterium extorquens* AM1: Purification and properties. *Eur J Biochem* 270:325–333.
37. Lu Y, et al. (2010) Alteration of hydrogen metabolism of ldh-deleted Enterobacter aerogenes by overexpression of NAD+-dependent formate dehydrogenase. *Appl Microbiol Biotechnol* 86:255–262.
38. Schirwitz K, Schmidt A, Lamzin VS (2007) High-resolution structures of formate dehydrogenase from *Candida boidinii*. *Protein Sci* 16:1146–1156.
39. Antoniewicz MR, Kelleher JK, Stephanopoulos G (2007) Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Anal Chem* 79: 7554–7559.
40. Dauvillée D, et al. (2005) Role of the *Escherichia coli glgX* gene in glycogen metabolism. *J Bacteriol* 187:1465–1473.
41. Zhou J, Zhu T, Cai Z, Li Y (2016) From cyanochemicals to cyanofactories: A review and perspective. *Microb Cell Fact* 15:2.
42. Tashiro Y, Hirano S, Matson MM, Atsumi S, Kondo A (2018) Electrical-biological hybrid system for CO_2 reduction. *Metab Eng* 47:211–218.
43. Yishai O, Bouzon M, Döring V, Bar-Even A (2018) In vivo assimilation of one-carbon via a synthetic reductive glycine pathway in *Escherichia coli*. *ACS Synth Biol*, 10.1021/acssynbio.8b00131.