Improving ethylene glycol utilization in *Escherichia coli* fermentation

Smaranika Panda[1], Vincent Yuen Kin Fung[1], Jie Fu J Zhou[1], Hong Liang[1], Kang Zhou[1 *]

[1]*Department of Chemical and Biomolecular Engineering, National University of Singapore*

**Corresponding author:*

Kang Zhou, Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore, 117585. Email: kang.zhou@nus.edu.sg

Abstract

Ethylene glycol (EG) is used in the manufacture of polyester plastics such as polyethylene terephthalate (PET). EG has a potential to become an abundant, renewable substrate in the future to fuel microbial production of value-added chemicals. It could be obtained from hydrolysis of the plastic wastes, or hydrogenolysis of cellulosic wastes. *Escherichia coli* is a workhorse in metabolic engineering applications, but most *E. coli* strains are unable to metabolize EG. In this study, we have successfully engineered *E. coli* to grow on EG as the major carbon source by overexpressing two dehydrogenases (FucO and AldA) that oxidize EG into glycolate, which can be metabolized via the glycerate pathway. Overexpression of *fucO* and *aldA* with a constitutive promoter (P_gyrA) improved the cell growth on EG. EG utilization was further improved with the supplementation of a mixture of amino acids at low concentration. In an optimized medium, the engineered strain consumed up to 20 g/L of EG. With further development in the future, this strain can potentially produce valuable chemicals from raw EG derived from plastic and cellulosic waste streams.

Keywords

Ethylene glycol, *Escherichia coli*, Glycolaldehyde, Glycolate, Metabolic engineering

1. Introduction

Ethylene glycol (EG) is an important chemical that has a wide range of industrial applications. EG can be directly used as an anti-freezing agent or a coolant, and it is also a substrate to produce polymers such as polyethylene terephthalate (PET) [1]. Over 25 million tons of PET are synthesized annually for the manufacturing of plastic bottles [2]. Due to its increasing global demand and persistent negative effect on the environment, the accumulation of PET plastics in nature is a growing concern [3, 4]. If PET wastes can be cost-effectively hydrolysed, the resultant hydrolysed PET monomers would be a source of sustainable and abundant feedstock to produce useful molecules. Specifically, EG obtained from PET hydrolysis may be utilized by microorganisms to produce recombinant proteins and small molecule products.

EG could also be obtained from cellulosic biomass through chemical reactions. The catalytic conversion of cellulose into EG and other fine chemicals has been intensively studied [5-11]. Cellulose was converted into EG either by a one-pot reaction using a multifunctional catalyst or a two-step process that first hydrolysed cellulose into glycolaldehyde, and subsequently hydrogenated the glycolaldehyde into EG [6, 8]. Other polyols such as glycerol, propylene glycol, xylitol, sorbitol, mannitol were obtained as by-products. The main bottleneck of commercializing this process is the lack of cost-effective technologies for purifying EG from the mixture of other polyols [11]. One possible workaround is the usage of microbes to convert EG and other polyols generated from this process into a homogenous, easy to purify product. Thus, microbial conversion of this EG into microbial biomass and a wide spectrum of products would play important roles in upcycling plastic and cellulosic wastes.

Whereas some bacteria such as *Pseudomonas aeruginosa* can naturally utilize EG [12], genetic engineering of microbes could improve the EG utilization rate in these microbes or enable the workhorse strains used in biotechnology to utilize EG. For example, a recent study engineered *Pseudomonas putida* KT2440 to convert EG into cellular mass [13].

Escherichia coli is a commonly used host organism for metabolic engineering [14, 15]. Yet the utilization of EG has not been intensively explored in *E. coli*. Although wildtype *E. coli* is unable to utilize EG as a sole carbon source, an early work established that EG-utilizing *E. coli* strain can be isolated through adaptive evolution from the mutants that could grow on propylene glycol [16]. The mutant had shown higher activities of L-1,2-propanediol oxidoreductase (FucO) and aldehyde dehydrogenase A (AldA) [16]. Thus, EG assimilation in *E. coli* was hypothesized to involve the oxidation of EG into glycolaldehyde and subsequently to glycolate (**Fig. 1a**). Glycolate can be further metabolized into glyoxylate by the glycolate dehydrogenase. Glyoxylate could be metabolized via the glyoxylate shunt by condensation with acetyl-CoA to form malate. Glyoxylate may also be condensed with succinate by an isocitrate lyase (encoded by *aceA* gene) to form isocitrate (**Fig. 1c**). However, one molecule of acetyl-CoA per glyoxylate is borrowed via the glyoxylate shunt and the most likely way to return it would be to oxidatively decarboxylate malate into pyruvate and then into acetyl-CoA. Similarly, for the isocitrate lyase-catalyzed reaction, a succinate is borrowed per glyoxylate, and the most obvious pathway of returning it is to oxidatively decarboxylate isocitrate into succinate. Hence, both pathways are cyclic, and the net reactions oxidize one glyoxylate into two CO_2, which would not lead to biomass formation. Therefore, the linear glycerate pathway must be used to convert glyoxylate towards

acetyl-CoA by condensing two glyoxylate into one tartronate semialdehyde (TSA) and further reduce it into glycerate (**Fig. 1a, 1b**).

In this study, we improved EG utilization in *E. coli* by optimizing gene expression (*fucO* and *aldA*) and growth medium. We found that the optimized strain had to require a very low concentration of glycerol (0.1 g/L) or a mixture of amino acids to efficiently utilize EG. Under the best condition we have found, the *E. coli* strain overexpressing *fucO* and *aldA* could consume up to 20 g/L EG in a chemically defined medium. This engineered strain is a good base strain for future works to convert raw EG obtained from hydrolysed PET waste and cellulosic biomass into value-added chemicals.

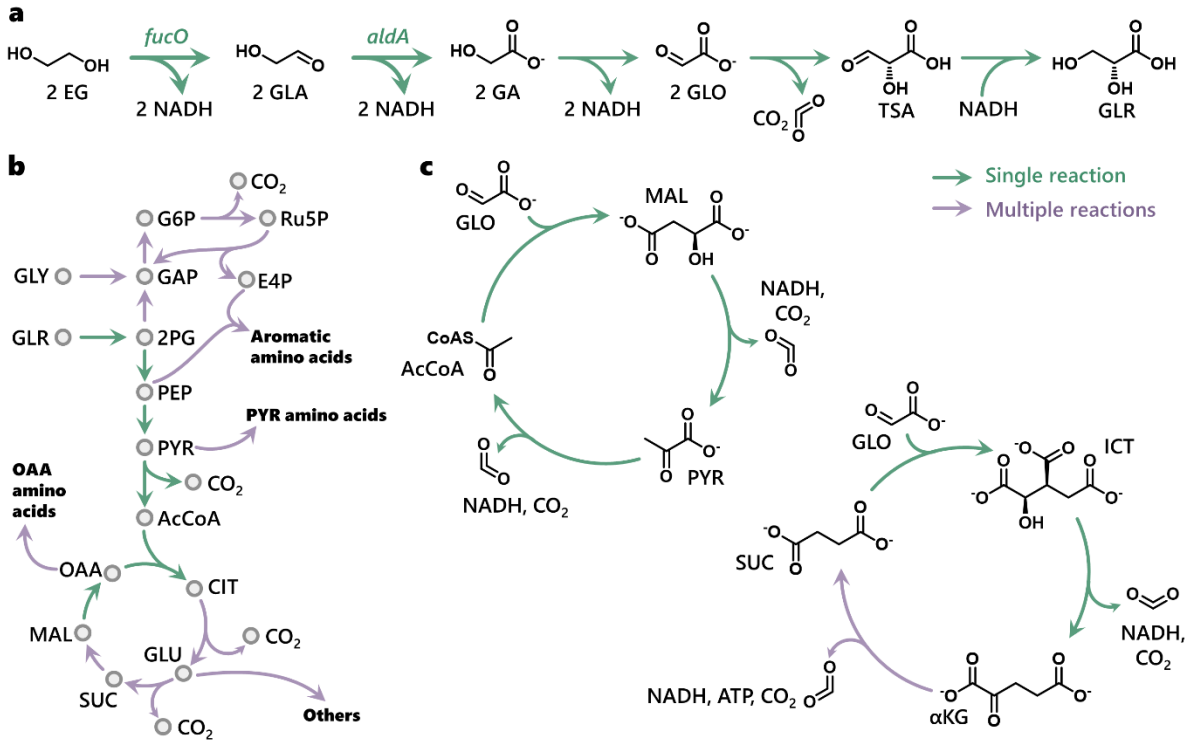


Fig. 1 Metabolic pathway for ethylene glycol (EG) utilization in *E. coli*. (**a**) The pathway for converting EG into glycerate (GLR). GLA: glycolaldehyde. GA: glycolate. GLO: glyoxylate. TSA: tartronate semialdehyde. GLY: glycerol. (**b**) The major reactions to convert glycerate into key biomass building blocks. PYR: pyruvate. OAA: oxaloacetate. (**c**) Condensation of glyoxylate with acetyl-CoA or succinate likely lead to full oxidation of one glyoxylate into two CO_2. The two pathways compete with the glycerate pathway for glyoxylate and could be used for energy generation. L-1,2-propanediol oxidoreductase (FucO) and aldehyde dehydrogenase A (AldA) were overexpressed in this study.

Table 1 Strains constructed in this study.

Strain	Parent strain	Plasmid in the strain
EG01	**BW25113**	pEG01 (P_BAD-*fucO-aldA*-pMB1-spect, Addgene ID: 154136)
EG02	**BW25113**	pEG02 (P_thrC3-*fucO-aldA*-p15A-cam, Addgene ID: 154137)
EG03	**BW25113**	pEG03 (P_gyrA-*fucO-aldA*-p15A-cam, Addgene ID: 154138)
EG03'	**BW25113**	pEG03' (P_gyrA-*fucO-aldA*-pMB1-spect, Addgene ID: 162578)
EG04	**MG1655 DE3**	pEG03 (P_gyrA-*fucO-aldA*-p15A-cam, Addgene ID: 154138)

pMB1 and p15A are the high copy number and medium copy number replication origins respectively; spect and cam are spectinomycin and chloramphenicol resistance cassettes respectively.

2. Results and discussion

2.1. Effect of initial glycerol concentration on EG utilization

We tested the ability of the wildtype *E. coli* K-12 and B strains to use EG or its degradation intermediate glycolate as sole source of carbon and energy. Neither strains showed any cell growth when EG was the sole carbon source, even after they were incubated for 72 h, while both strains could utilize glycolate as a sole source of carbon and energy (**Fig. S1**).

A recent study found that an *E. coli* utilized ethylene glycol when *fucO* and *aldA* were overexpressed and 4 g/L glycerol was added [17]. Similar to that study, we overexpressed *fucO* and *aldA* in *E. coli* **BW25113** (a K-12 strain) under the control of the P_BAD promoter that can be induced by L-arabinose. The obtained strain was named as **EG01**. Although the L-arabinose degradation operon (araBAD) was deleted in **BW25113**, the strain could still grow to cell density of 0.6 (optical density at 600 nm [OD_{600}]) within 72 h by using 1 g/L L-arabinose in a chemically defined medium (**Fig. 2**, **Fig. S2**). The medium also contained 5 g/L EG, but **BW25113** did not consume it. **EG01** consumed 6% of EG and reached cell density of 0.8 (OD_{600}) in the same medium (**Fig. 2**, **Fig. S2**). We only detected glycolate in the culture of **EG01** but not in

that of **BW25113** (**Fig. S2**), supporting that the consumed EG was assimilated via glycolate.

Glycerol was used in the previous study [17] to improve EG utilization because it did not incur catabolite repression, unlike glucose. In our study, adding 0.1 g/L of glycerol doubled final cell density of **EG01** (at 72 h) and enabled the cells to consume 17% of the added EG (**Fig. 2**). Adding the same amount of glycerol also improved final cell density and EG utilization of **BW25113**, but to a much less extent. Besides as a carbon source, glycerol may have induced expression of transporters and/or other enzymes involved in polyol degradation (both glycerol and EG are polyol), which facilitated EG utilization only in **EG01**. The wildtype strain (**BW25113**) may not benefit from this mechanism as its EG utilization could be limited by low expression levels of *fucO* and *aldA*. Increasing the glycerol concentration to 0.5 g/L improved the final cell density of both **EG01** and **BW25113**, but did not substantially enhance their EG utilization. After 72 h, only 20% of the EG was consumed by **EG01**, suggesting that its EG assimilation may be limited by FucO and AldA in the presence of glycerol. This challenge may be overcome by using better promoters to drive the expression of *fucO* and *aldA*.

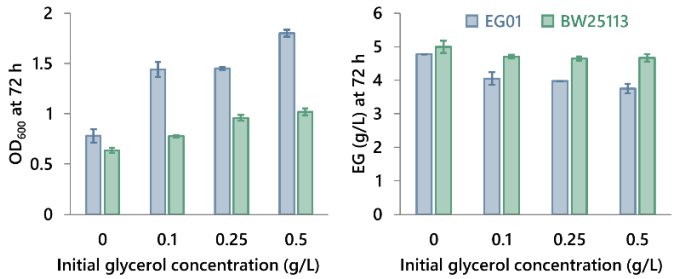


Fig. 2 Effects of initial glycerol concentration on final cell density and EG utilization of **EG01** and **BW25113**. Adding 0.1 g/L glycerol substantially improved final cell density and EG utilization of **EG01**. The cells were cultured in a chemically defined medium (5 mL of culture in 50 mL culture tube) containing 1 g/L L-arabinose and 5 g/L EG as carbon source. The culture temperature was 30 °C. Error bar indicates standard error (n=3).

2.2. Effects of promoter, growth temperature, and host strain on EG utilization

In the experiments described above, cells were grown at 30 °C. We tested a higher temperature (37 °C) but found that almost no EG was utilized under the condition (**Fig. S3**). As a result, 30 °C was used in the rest of this study. Two new expression vectors with an auto-inducible promoter P_thrC3 [18] and a strong constitutive promoter P_gyrA respectively were subsequently constructed and introduced into **BW25113** separately, creating strains **EG02** and **EG03** respectively. P_gyrA (**EG03**) substantially outperformed the P_BAD (**EG01**) in both cell growth and EG utilization, and was used in the rest of this study (**Fig. 3a**). The cell density of **EG03** was three times that of **EG01**, and **EG03** utilized close to 80% of the EG in the medium (**Fig. 3a**). When P_BAD was not used, no L-arabinose was added, which not only removed the confounding factor that some cells utilized L-arabinose as a carbon source, but also avoided catabolite repression caused by L-arabinose, if any. P_thrC3 was found to be autonomously upregulated when cells grew in media containing conventional carbon substrates (e.g. glucose and glycerol) [18]. It, however, failed to efficiently drive EG utilization (**Fig. 3a**) possibly because the promoter was not active when EG was the carbon source. P_BAD was used in a high copy number plasmid with a spectinomycin resistance gene (pEG01, **Table 1**), while P_gyrA was used in a medium copy number plasmid with a chloramphenicol resistance gene (pEG03, **Table 1**). To test if the better performance observed with pEG03 (compared with pEG01) was partly due to the difference in plasmid copy number and/or antibiotic resistance, we changed the replication origin and antibiotic resistance of pEG03 to be the same as pEG01, and introduced the new plasmid (pEG03') into **BW25113**, creating strain **EG03'**. This strain had almost the same performance as **EG03** in terms of EG utilization rate and growth rate (**Fig. S4**), invalidating the hypothesis that plasmid replication and antibiotic

resistance had a substantial effect on the EG utilization. In the rest of this study, we chose to use pEG03 instead of pEG03', because pEG03 has lower copy number, which theoretically incur less maintenance burden to the host cell.

We further introduced the pEG03 into *E. coli* **MG1655 DE3** (creating **EG04**) and found **EG04** slightly outperformed **EG03** in both cell growth and EG utilization (**Fig. 3b**). We, therefore, used **EG04** as our best strain in the rest of this study.

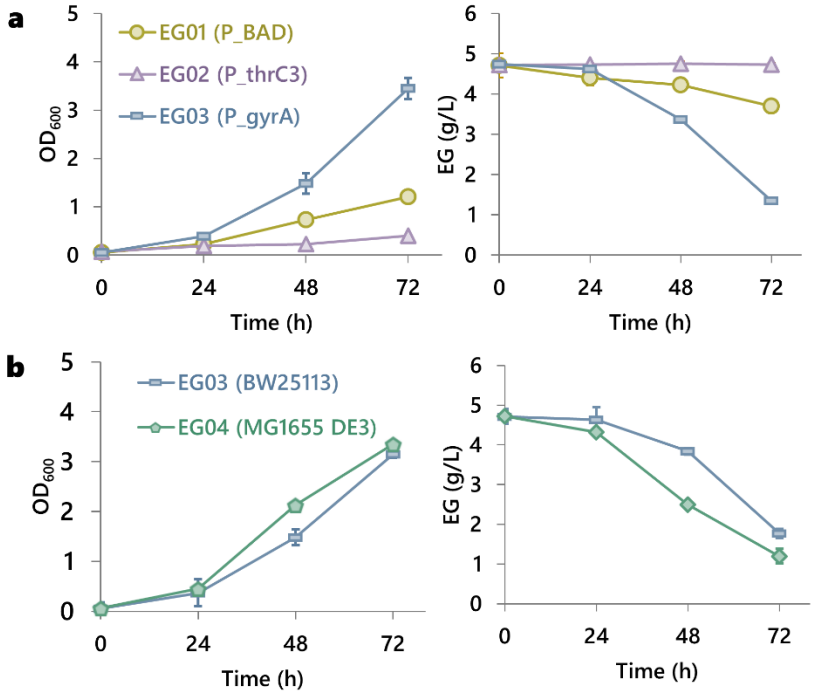


Fig. 3 Effect of (**a**) promoter and (**b**) host strain on cell growth and EG utilization. **EG01**, **EG02** and **EG03** drove expression of *fucO* and *aldA* by using P_BAD, P_thrC3, and P_gyrA respectively. They were all derived from **BW25113**. Their full genotypes are in **Table 1**. **EG03** and **EG04** both drove expression of *fucO* and *aldA* by using P_gyrA. **EG04** was derived from **MG1655 DE3**. All the cells were cultured in a chemically defined medium (5 mL of culture in 50 mL culture tube) containing 5 g/L EG and 0.1 g/L glycerol. In the medium of **EG01**, there was also 1 g/L L-arabinose. The culture temperature was 30 °C. Error bar indicates standard error (n=3). Some error bars are smaller than the markers in the plots.

2.3. Supplementation of amino acids

Similar to **EG01**, **EG04** could not utilize EG efficiently without the addition of glycerol (**Fig. S5**). We hypothesized that the synthesis of some essential building blocks such

as amino acids could limit the cell growth of **EG04**. Amino acids are building blocks to synthesize proteins, which are approximately half of biomass by weight, and supplementation of amino acids were recently found to substantially improve *E. coli* growth on ethanol, another non-conventional C2 substrate [19]. We supplemented 1 g/L (final concentration) of Complete Supplement Mixture (CSM, a mixture of amino acids and nucleobases) instead of 0.1 g/L glycerol and found that it indeed substantially improved cell growth and EG utilization. The final cell density reached 4.5 (OD_{600}) when 1 g/L CSM was supplemented, which was 27% higher than that achieved with addition of 0.1 g/L glycerol (**Fig. 4a**). EG (5 g/L) was almost completely consumed by the cells with the CSM supplementation (**Fig. 4a**). If no EG was included in the medium (with only 1 g/L CSM as carbon source), the cell density was much lower (OD_{600} = 1, **Fig. 4a**).

Reverse Transcription-Quantitative PCR (RT-qPCR) analysis found that the expression level of *fucO* and *aldA* in the cells grown with the amino acid supplementation were 250% and 70%, respectively, higher than those with the glycerol supplementation (**Fig. 4a**). The results suggested that EG assimilation may still be limited by the activity of FucO or AldA under the glycerol supplementation condition.

To understand which components of CSM were primarily responsible for the improvement of EG utilization, we first removed the nucleobases and found that adding the amino acids in CSM achieved the same effect as CSM (**Fig. 4a**). We then divided the amino acids in CSM into four families and supplemented each family of amino acids to understand their effect on cell growth and EG utilization. The

concentration of individual amino acid in each group was the same to that in CSM. The four families are the aspartate family, the aromatic family, the pyruvate family, and others (histidine and arginine) according to their biosynthetic pathways (**Fig. 1b**).

Although adding the aromatic family substantially improved the cell density and EG utilization, the improvement was much less pronounced than adding all the amino acids in CSM (**Fig. 4b**). Adding the aspartate family and others (histidine and arginine) also had positive effect on the EG utilization. We then tested all the combinations of the amino acid families, but none of them can reproduce the effect of adding all the amino acids in CSM (**Fig. 4b**). This observation was different from a similar experiment we did in another study on improving ethanol utilization in *E. coli*, in which adding the aspartate family of amino acids achieved the same results as supplementing all the amino acids in CSM [19]. This difference suggests that the amino acids limiting cell growth are not the same when different non-conventional carbon substrates are used, which is also expected because ethanol and EG use very different points to enter the central metabolism – ethanol and EG use acetyl-CoA and glyoxylate respectively. We continued to use CSM in the rest of the study because we did not manage to substantially reduce the list of supplementation components as described above and it was convenient to use CSM (it is a commercial, ready-to-use product).

We note that amino acid supplementation adds to fermentation cost at industrial scale. A future research direction is to engineer a strain that is able to utilize EG efficiently without amino acid supplementation.

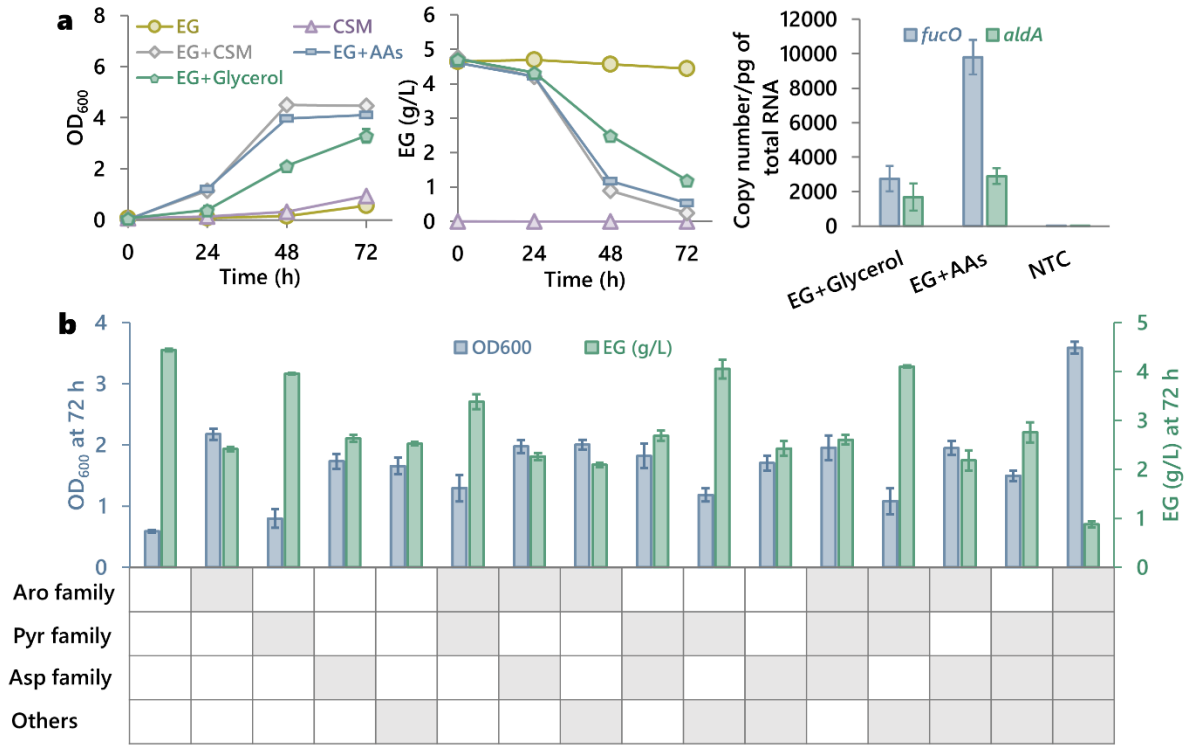


Fig. 4 (**a**) Effects of amino acids on cell growth, EG utilization, and expression levels of *fucO* and *aldA*. CSM: complete supplement mixture, including amino acids and nucleobases (1 g/L, composition described in Materials and Methods). AAs: the amino acids in CSM (concentration of each amino acid was the same to that in CSM). Gene transcription was quantified by using RT-qPCR. The data were normalized by using the amount of total RNA. No template control (NTC) was included to confirm that there was no contamination. In all the experiments, **EG04** was grown in a chemically defined medium (5 mL of culture in 50 mL culture tube). The growth temperature was 30 °C. No glycerol was added when amino acids were added. The glycerol concentration was 0.1 g/L when glycerol was added. The EG concentration was 5 g/L when EG was added. Fresh cell cultures harvested at the exponential phase were used for the RT-qPCR analysis. (**b**) The amino acids were further broken down into four families, but we did not manage to find a simpler combination of the families of the amino acids to replace CSM. All the amino acid families were required to reproduce the effect of the amino acids in CSM. ARO: aromatic. PYR: pyruvate. ASP: aspartate. Error bar indicates standard error (n=3). Some error bars are smaller than the markers in the scatter plots.

2.4. Further medium optimization for improving EG utilization

We further tested a higher concentration of EG (20 g/L) supplemented with 1 g/L CSM. The cell density obtained with 20 g/L EG was similar to what was achieved with 5 g/L EG and 51% of EG was utilized within 72 h (**Fig. 5a, Fig. S6**). We did not detect any of the following fermentation by-products in both conditions: acetate, lactate, ethanol, glycerol, succinate, formate and glycolate (the detection limits were approximately 0.1 g/L). We hypothesized that certain amino acids were limiting cell growth when the

concentration of the carbon source was increased from 5 g/L to 20 g/L (the concentration of CSM remained to be 1 g/L). To test this hypothesis, we tested different ratios of EG to CSM. The results indicated that the amino acids were not limiting when 5 g/L of EG was fed with 1 g/L CSM (**Fig. 5a**). However, when 20 g/L of EG was fed, EG utilization was improved substantially when the concentration of CSM was increased from 1 g/L to 2 g/L (**Fig. 5a**). We further tested supplementation of different nitrogen sources or compounds involved in nitrogen assimilation such as glutamate and α-ketoglutarate. We also increased the concentration of inorganic nitrogen source (NH_4Cl) in the medium, but cell growth and EG utilization were not further improved (**Fig. S7**). We supplemented 2 g/L of CSM in the rest of this study.

Antibiotics are commonly used to create selection pressure to maintain plasmid in microbes. In all the experiments above, 25 mg/L chloramphenicol or 50 mg/L spectinomycin was used. We hypothesized that the antibiotic might not be needed in this project because cells would not grow if they lose the plasmid because the plasmid carries the genes enabling growth on the major carbon source. A subsequent experiment indeed confirmed that not adding antibiotics did not negatively affect cell growth and EG utilization. In fact, the cells had a much shorter lag phase when the antibiotic was not added (**Fig. 5b**), possibly because the antibiotic (25 mg/L chloramphenicol) had inhibitory effect on cell growth in the early growth phase. Avoiding use of antibiotic in industrial scale fermentation reduces fermentation cost and avoids the risk of leaking antibiotic to the environment, which may cause development of microbial pathogens that are resistant to multiple antibiotic drugs.

To evaluate the maximum amount of EG that can be consumed by **EG04** in this study, we tested higher initial concentrations of EG with supplementation of 2 g/L of CSM without antibiotic addition. When 20 g/L EG was used, all EG was metabolized completely in 72 h (**Fig. 5c**). When the initial EG concentration was increased to 50 g/L, EG was metabolized at a similar rate till 48 h. The end of EG consumption after 48h in 50 g/L EG could be attributed to the inefficiency in synthesis of certain amino acids. Adding another dose of CSM at 48 h may enable further EG utilization. The strain only consumed a small quantity of EG when 100 g/L of EG was fed (**Fig. 5c**), indicating that a high concentration of EG was inhibitory to *E. coli*'s growth. With a fed-batch operation in a pH- and dissolved oxygen-controlled bioreactor and an optimized substrate feeding algorithm, utilizing a larger amount of EG at a higher rate should be expected.

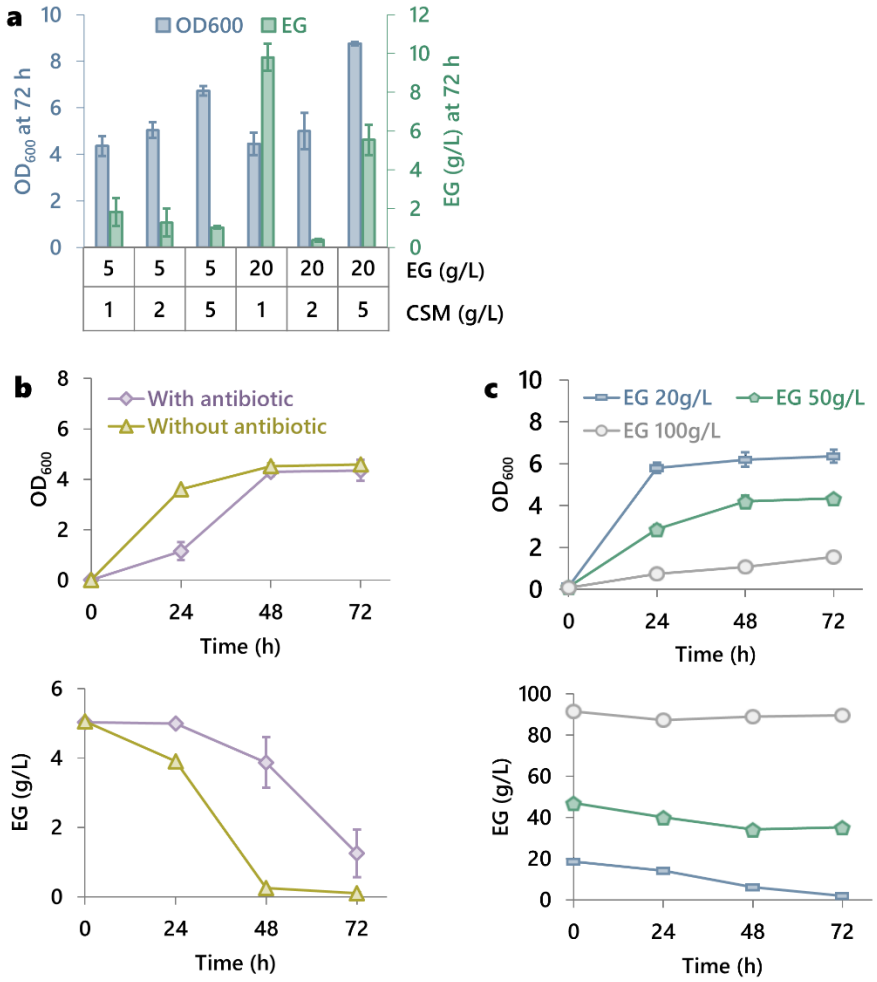


Fig. 5 Further optimization of growth medium and culture conditions for utilizing a larger quantity of EG. (**a**) Higher concentration of CSM was needed when EG concentration was increased to 20 g/L. (**b**) The cells had a shorter lag phase when the antibiotic (25 mg/L chloramphenicol) was not added. (**c**) Cells did not completely utilize EG when the concentration was more than 20 g/L. Further process optimization will be needed to utilize a larger quantity of EG at a faster rate. In all the experiments, **EG04** was grown in a chemically defined medium (20 mL of culture in 125 mL shake flask). The growth temperature was 30 °C. Error bar indicates standard error (n=3). Some error bars are smaller than the markers in the scatter plots.

3. Conclusions

This study systematically optimized gene expression and medium composition for improving EG utilization of *E. coli*. We found that a constitutive promoter worked substantially better than inducible promoters in driving expression of two key enzymes in EG assimilation, and that addition of a low concentration of glycerol (0.1 g/L) or supplementation of amino acids was critical in enabling efficient EG utilization. It was also demonstrated that antibiotic was not needed to maintain the plasmid expressing

the EG utilization genes. The strain developed and the growth conditions identified in this study should be useful to future studies that aimed to utilize EG as a fermentation substrate. With further development, raw EG derived from the plastic and cellulosic waste streams may be directly used by microbes to produce value-added chemicals.

4. Materials and methods

4.1. Materials

The chemicals used in this work were purchased from Sigma-Aldrich unless otherwise stated. Media were purchased from Biomed Diagnostic PTE Ltd. The CSM powder was purchased from Sunrise Science Products.

4.2. Plasmids and strains

The DNA oligos used in this study were synthesized by Integrated DNA Technologies (IDT) (**Table S1**). The genes of interest were amplified from *E. coli* K-12 strain. Q5® High-Fidelity 2×Master Mix (New England Biolabs) was used in all the PCR reactions except qPCR reactions, in which Taq polymerase was used (Thermo). All the plasmids were constructed according to the GT standard [20]. The list of plasmids is summarized in **Table S1**. Plasmid sequences and annotations can be found by using Addgene IDs (**Table 1**) on www.addgene.org. DNA assembly was done by using CLIVA [21] and the assembled DNA molecules were introduced into *E. coli* DH5alpha competent cells (New England Biolabs) using the standard heat-shock method. The transformants were isolated on LB agar plates containing appropriate antibiotics. The sequences of all the constructed plasmids were verified using Sanger sequencing (Service provider: Bio Basic Asia Pacific Pte Ltd, Singapore).

E. coli **BW25113** was obtained from the Coli Genetic Stock Center (Yale University) [22]. The constructed plasmids were introduced into *E. coli* **BW25113 or MG1655 DE3** by using the standard heat-shock method. The constructed strains are summarized in **Table 1**.

4.3. Growth media

We used a modified M9 medium which contained 6.78 g/L disodium phosphate (Na_2HPO_4), 3 g/L monopotassium phosphate (KH_2PO_4), 1 g/L ammonium chloride (NH_4Cl), 0.5 g/L sodium chloride (NaCl) and 0.17% (V/V) K3 master mix. The K3 master mix was prepared by mixing 2.5 mL of 0.1 M ferric citrate solution (autoclaved), 1 mL of 4.5 g/L thiamine solution, 3 mL of 4 mM Na_2MoO_4 (autoclaved), 1mL of 1 M $MgSO_4$ solution (autoclaved) and 1mL of 1000 X K3 trace elements stock. The 1000 X K3 trace elements stock contained 5 g/L $CaCl_2 \cdot 2H_2O$, 1.6 g/L $MnCl_2 \cdot 4H_2O$, 0.38 g/L $CuCl_2 \cdot 2H_2O$, 0.5 g/L $CoCl_2 \cdot 6H_2O$, 0.94 g/L $ZnCl_2$, 0.03 g/L H_3BO_3, 0.4 g/L $Na_2EDTA \cdot 2H_2O$. The pH of the medium was adjusted to 7. When antibiotic was used, 25 mg/L chloramphenicol or 50 mg/L spectinomycin was used. The antibiotic resistance of each strain is described in **Table 1**. In some experiments, antibiotic was not added as specified in the relevant sections. Different concentrations of EG was added as carbon source. Different concentrations of glycerol or CSM/amino acids were supplemented. The CSM was composed of adenine hemisulfate (10 mg/L), L-arginine (50 mg/L), L-aspartic acid (80 mg/L), L-histidine hydrochloride monohydrate (20 mg/L), L-isoleucine (50 mg/L), L-leucine (100 mg/L), L-lysine hydrochloride (50 mg/L), L-methionine (20 mg/L), L-phenylalanine (50 mg/L), L-threonine (100 mg/L), L-tryptophan (50 mg/L), L-tyrosine (50 mg/L), L-valine (140 mg/L) and uracil (20 mg/L). The aspartate family was composed of L-aspartate, L-lysine, L-threonine, L-

methionine and L-isoleucine; the aromatic family was composed of L-tryptophan, L-tyrosine and L-phenylalanine; the pyruvate family was composed of L-leucine and L-valine; others contained L-histidine and L-arginine; the concentration of individual amino acid in each group was the same to that in CSM.

4.4. Culture conditions

A single colony was inoculated into LB with 25 mg/L chloramphenicol or 50 mg/L spectinomycin and incubated at 37 °C/250 rpm overnight. The overnight culture was centrifuged at 4,000 g for 10 mins and washed two times with ultrapure water. The cell pellet was inoculated into the modified M9 medium in 50 mL tubes or shake flasks (as specified in the figure captions) with a seeding density of 0.1 (OD_{600}) with or without antibiotic addition. L-arabinose was added (1 g/L final concentration), when **EG01** was used. The culture was incubated at indicated temperatures with shaking at 250 rpm for 72 h.

4.5. RT-qPCR

Cells were harvested during their exponential growth phase for RNA extraction. A fixed volume of cell culture were used (1 mL). Total RNA was extracted using the Thermo Scientific GeneJET RNA purification kit, K0731 following the manufacturer's protocol.

The RNA concentration was measured using Nanodrop spectrophotometer. DNase treatment was done to remove gDNA from RNA samples (Promega, M6101). Same amount of RNA (75 ng) was used for all the samples for cDNA synthesis. Reverse transcription was carried out using ImProm-II reverse transcriptase (Promega, A3802). The reaction volume was 20 µL. A no reverse transcriptase control (-RT) was used to

assess genomic DNA contamination for each sample, and no genomic DNA contamination was found (data not shown). 5 µl of cDNA product (without purification) was used in a 30 µL qPCR reaction, which contained reaction buffer (Mirxes), Taq polymerase (Thermo), and primers. The instrument used was Bio-Rad CFX96. No template control (NTC) was prepared by replacing cDNA sample with ultrapure water. The copy number of the *fucO* and *aldA* transcripts were quantified by using the standard curves constructed by using a linearized plasmid that contained *fucO* and *aldA*. The results were normalized by using the amount of total RNA.

4.6. Quantification and analysis of metabolites

The optical density of the cultures was measured at 600 nm (OD_{600}) at indicated time points using a microplate reader (Tecan infinite M200). The absorbance data was converted into standard OD_{600} units using a standard curve. To measure the concentrations of EG and other fermentation by-products such as acetate, lactate, ethanol, glycerol, succinate, formate and glycolate, 0.2 mL of cell culture was collected at the indicated time points. The culture was centrifuged for 5 mins at 12,000 g, and the supernatant was filtered using a nylon syringe filter with a pore size of 0.22 µm and a diameter of 47 mm (IT Technologies Pte Ltd). 5 µL of the obtained filtered supernatant was injected into a high-performance liquid chromatography (HPLC) instrument that used an Aminex column (HPX-87H, 300X7.8 mm, Bio-Rad). 5 mM H_2SO_4 was used as mobile phase with a flow rate of 0.7 mL/min. The compounds were detected using a refractive index detector (RID). A calibration curve was obtained for each compound. Products were identified based on their retention time.

Data availability

All the data used to support the findings of this study are included in this paper and its supplementary information file.

Acknowledgement

This work was financially supported by research grants from Singapore Millennium Foundation (grant identifier: R-279-000-516-592) and from Singapore Ministry of Education (a Tier-2 grant; grant identifier: R-279-000-594-112). Smaranika Panda was supported by Singapore Ministry of Education through a PhD scholarship.

Author contributions

K.Z. and S.P. conceived the study and designed the experiments. S.P. performed the experiments and collected data. K.Z. and S.P. analyzed the data. V.F. and J.Z. were involved in analysis and discussions during the study. H.L. contributed to plasmid construction and medium design in the study. S.P. and K.Z. wrote the manuscript. All authors have read and approved the final version of this manuscript. The authors declare no conflict of interest.

Supporting information

Fig. S1-7 and **Table S1** are available in the Supplementary Information file.

References

[1] E.W. Carney, S.A. Stice, Ethylene Glycol, Reproductive and Developmental Toxicology, Elsevier2017, pp. 797-809.
[2] Z. Chen, J. Huang, Y. Wu, D. Liu, Metabolic engineering of Corynebacterium glutamicum for the de novo production of ethylene glycol from glucose, Metabolic engineering, 33 (2016) 12-18.
[3] C.A. Staples, J.B. Williams, G.R. Craig, K.M. Roberts, Fate, effects and potential environmental risks of ethylene glycol: a review, Chemosphere, 43 (2001) 377-383.

[4] H.K. Webb, J. Arnott, R.J. Crawford, E.P. Ivanova, Plastic degradation and its environmental implications with special reference to poly (ethylene terephthalate), Polymers, 5 (2013) 1-18.
[5] H. Kobayashi, A. Fukuoka, Chapter 2-Current Catalytic Processes for Biomass Conversion, New and Future Developments in Catalysis; Suib, SL, Ed.; Elsevier: Amsterdam, The Netherlands, (2013) 29-52.
[6] N. Yan, C. Zhao, P.J. Dyson, C. Wang, L.t. Liu, Y. Kou, Selective degradation of wood lignin over noble-metal catalysts in a two-step process, ChemSusChem: Chemistry & Sustainability Energy & Materials, 1 (2008) 626-629.
[7] S. Yu, X. Cao, S. Liu, L. Li, Q. Wu, Production of ethylene glycol from direct catalytic conversion of cellulose over a binary catalyst of metal-loaded modified SBA-15 and phosphotungstic acid, RSC advances, 8 (2018) 24857-24865.
[8] A. Wang, T. Zhang, One-pot conversion of cellulose to ethylene glycol with multifunctional tungsten-based catalysts, Accounts of chemical research, 46 (2013) 1377-1386.
[9] Z. Zhang, I.M. O'Hara, D.W. Rackemann, W.O. Doherty, Low temperature pretreatment of sugarcane bagasse at atmospheric pressure using mixtures of ethylene carbonate and ethylene glycol, Green Chemistry, 15 (2013) 255-264.
[10] N. Ji, T. Zhang, M. Zheng, A. Wang, H. Wang, X. Wang, Y. Shu, A.L. Stottlemyer, J.G. Chen, Catalytic conversion of cellulose into ethylene glycol over supported carbide catalysts, Catalysis Today, 147 (2009) 77-85.
[11] O. Rosales-Calderon, V. Arantes, A review on commercial-scale high-value products that can be produced alongside cellulosic ethanol, Biotechnology for biofuels, 12 (2019) 240.
[12] S. Dobson, W.H. Organization, Ethylene glycol: environmental aspects, World health organization2000.
[13] M.A. Franden, L.N. Jayakody, W.-J. Li, N.J. Wagner, N.S. Cleveland, W.E. Michener, B. Hauer, L.M. Blank, N. Wierckx, J. Klebensberger, Engineering Pseudomonas putida KT2440 for efficient ethylene glycol utilization, Metabolic engineering, 48 (2018) 197-207.
[14] S. Pontrelli, T.-Y. Chiu, E.I. Lan, F.Y.-H. Chen, P. Chang, J.C. Liao, Escherichia coli as a host for metabolic engineering, Metabolic engineering, 50 (2018) 16-46.
[15] Y. Liu, Z.J. Low, X. Ma, H. Liang, A.J. Sinskey, G. Stephanopoulos, K. Zhou, Using biopolymer bodies for encapsulation of hydrophobic products in bacterium, Metab Eng, 61 (2020) 206-214.
[16] A. Boronat, E. Caballero, J. Aguilar, Experimental evolution of a metabolic pathway for ethylene glycol utilization by Escherichia coli, Journal of bacteriology, 153 (1983) 134-139.
[17] B. Szappanos, J. Fritzemeier, B. Csörgő, V. Lázár, X. Lu, G. Fekete, B. Bálint, R. Herczeg, I. Nagy, R.A. Notebaart, Adaptive evolution of complex innovations through stepwise metabolic niche expansion, Nature communications, 7 (2016) 1-10.
[18] O. Anilionyte, H. Liang, X. Ma, L. Yang, K. Zhou, Short, auto-inducible promoters for well-controlled protein expression in Escherichia coli, Applied microbiology and biotechnology, 102 (2018) 7007-7015.
[19] H. Liang, X. Ma, W. Ning, Y. Liu, A.J. Sinskey, G. Stephanopoulos, K. Zhou, Constructing an ethanol utilization pathway in Escherichia coli to produce acetyl-CoA derived compounds, Metab Eng, (2020).
[20] X. Ma, H. Liang, X. Cui, Y. Liu, H. Lu, W. Ning, N.Y. Poon, B. Ho, K. Zhou, A standard for near-scarless plasmid construction using reusable DNA parts, Nature communications, 10 (2019) 1-12.

[21] R. Zou, K. Zhou, G. Stephanopoulos, H.P. Too, Combinatorial engineering of 1-deoxy-D-xylulose 5-phosphate pathway using cross-lapping in vitro assembly (CLIVA) method, PloS one, 8 (2013).
[22] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. Datsenko, Tomita 565 M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-566 gene knockout mutants: The Keio collection, Mol. Syst. Biol, 2 (2006) 570.

Supplementary file

Improving ethylene glycol utilization in *Escherichia coli* fermentation

Smaranika Panda[1], Vincent Fung Kin Yuen[1], Jie Fu J Zhou[1], Hong Liang[1], Kang Zhou[1 *]

[1]*Department of Chemical and Biomolecular Engineering, National University of Singapore*

**Corresponding author:*

Kang Zhou, Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore, 117585. Email: kang.zhou@nus.edu.sg

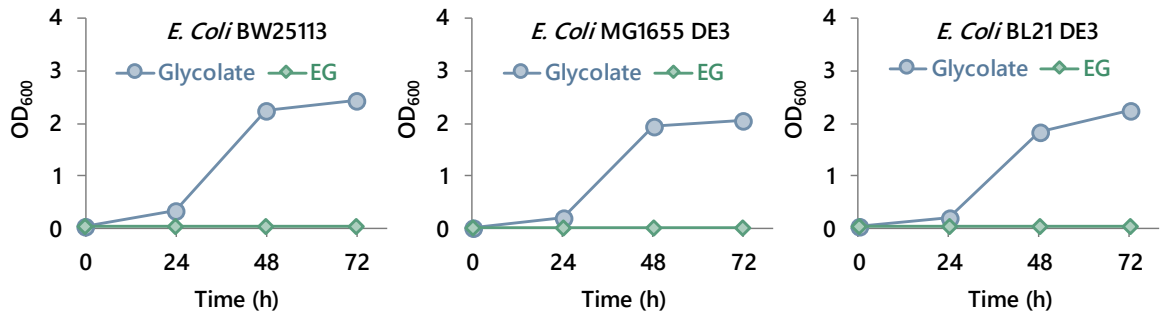


Fig. S1 Final cell density of wildtype *E. coli* K-12 strains (**BW25113** and **MG1655 DE3**), and B-strain (**BL21 DE3**) when grown on EG and glycolate separately. Both the strains did not show any cell growth when EG was the sole carbon source, while they could utilize glycolate as the sole source of carbon and energy. The cells were cultured in a chemically defined medium (5 mL culture volume in a 50 mL culture tube) containing 5 g/L EG or 5 g/L glycolate as carbon source. The culture temperature was 30 °C. Error bar indicates standard error (n=3).

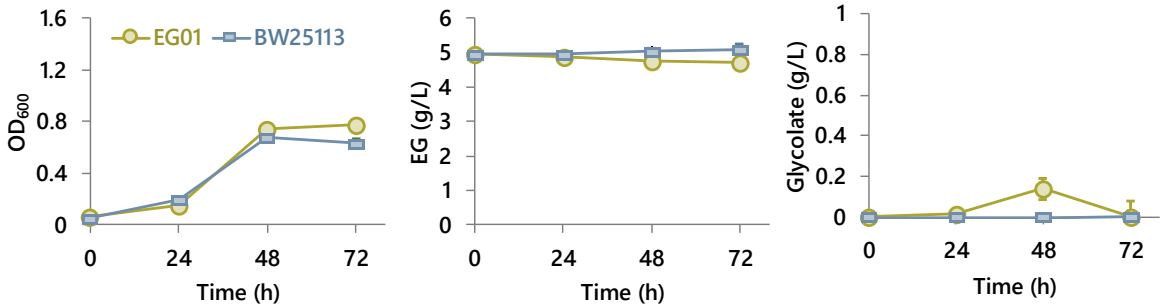
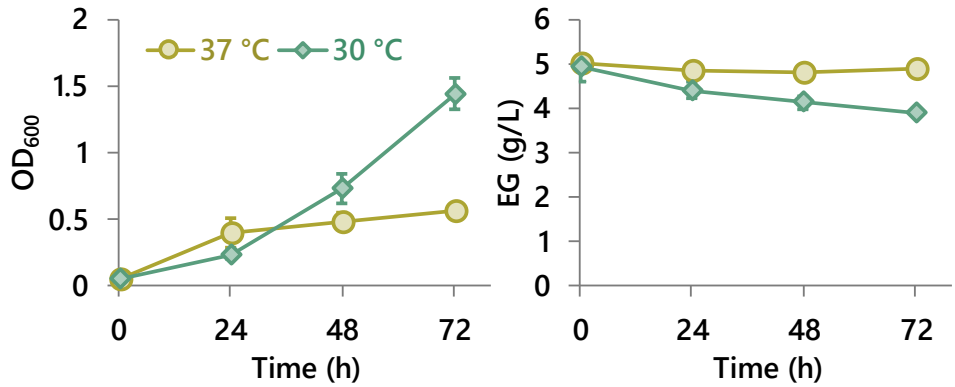


Fig. S2 Cell growth, EG utilization and glycolate production in **EG01** and **BW25113**. **BW25113** did not consume EG. All the cells were cultured in a chemically defined medium (5 mL culture volume in a 50 mL culture tube) containing 5 g/L EG and 1 g/L L-arabinose. The culture temperature was 30 °C. Error bar indicates standard error (n=3).



564

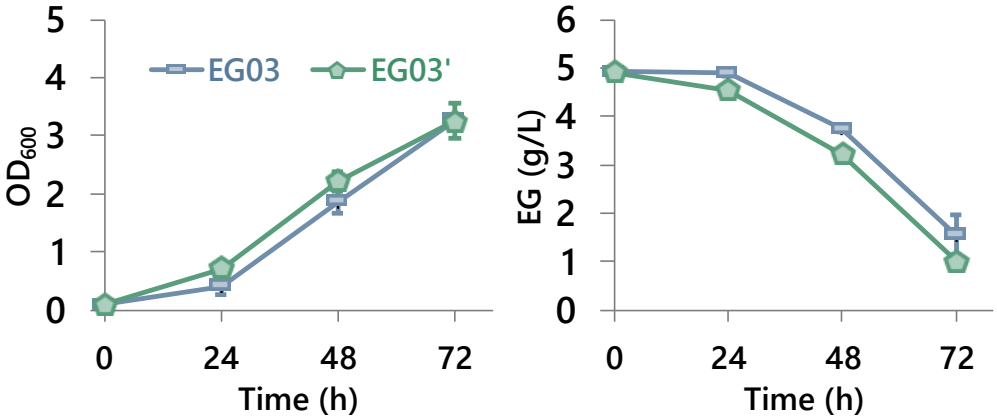
565

566 **Fig. S3** Effect of growth temperature on cell growth and EG utilization. Cell growth at 30 °C was better
567 than at 37 °C and almost no EG was utilized at 37 °C. **EG01** was cultured in a chemically defined medium
568 (5 mL culture volume in a 50 mL culture tube) containing 5 g/L EG, 0.1 g/L glycerol and 1 g/L L-arabinose.
569 The culture temperature was either 30 °C or 37 °C. Error bar indicates standard error (n=3).

570

571

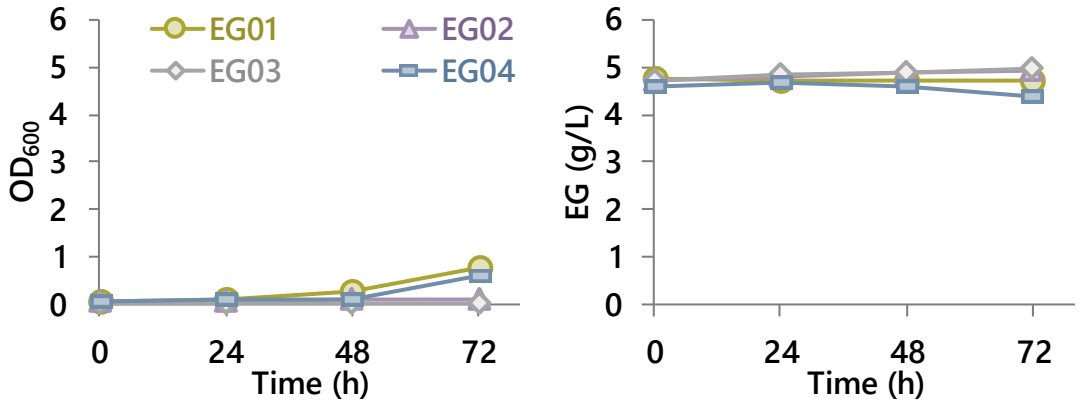
572



573

574 **Fig. S4** Effect of plasmid replication and antibiotic resistance on cell growth and EG utilization. **EG03**
575 drove the expression of *fucO* and *aldA* by using P_gyrA in a medium copy number plasmid with a
576 chloramphenicol resistance gene. **EG03'** drove the expression of *fucO* and *aldA* by using P_gyrA in a
577 high copy number plasmid with a spectinomycin resistance gene. Their full genotype is in **Table 1**. EG03'
578 and EG03 both had almost similar performance, invalidating the hypothesis that plasmid replication and
579 antibiotic resistance had a substantial effect on the EG utilization. All the cells were cultured in a
580 chemically defined medium (5 mL of culture in 50 mL culture tube) containing 5 g/L EG and 0.1 g/L
581 glycerol. The culture temperature was 30 °C. Error bar indicates standard error (n=3).

582



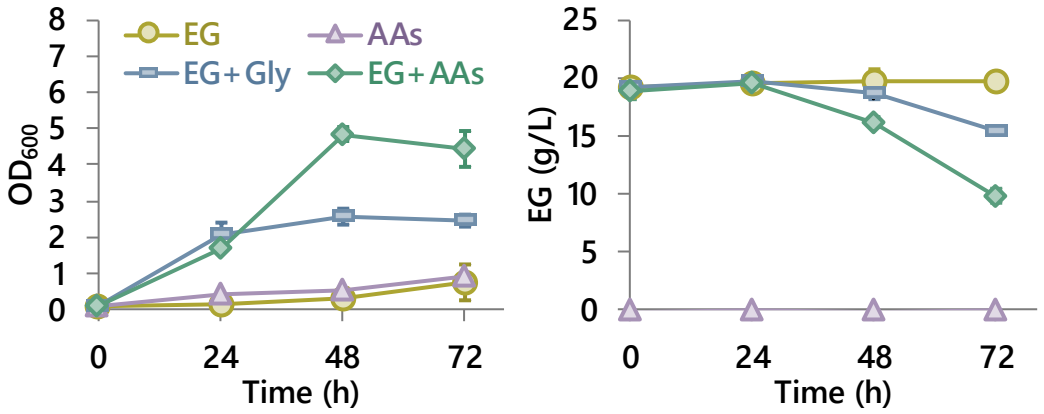
583

584

585 **Fig. S5** Cell growth and EG utilization of **EG01**, **EG02**, **EG03** and **EG04** without any supplementation of
586 glycerol or amino acids. Similar to **EG01**, other engineered cells could not utilize EG efficiently without
587 the addition of glycerol. All the cells were cultured in a chemically defined medium (5 mL culture volume
588 in a 50 mL culture tube) containing 5 g/L EG. In the medium of **EG01**, there was also 1 g/L L-arabinose.
589 The culture temperature was 30 °C. Error bar indicates standard error (n=3).

590

591



592

593

594 **Fig. S6** Final cell density and EG utilization of **EG04** when grown on a higher initial concentration of EG
595 (20 g/L). In all the experiments, **EG04** was grown in a chemically defined medium (20 mL culture volume
596 in a 125 mL shake flask). AAs: the amino acids in CSM (concentration of each amino acid was the same
597 to that in CSM). No glycerol was added when amino acids were added. The glycerol concentration was
598 0.1 g/L when glycerol was added. The concentration of amino acids was 1 g/L (composition described in
599 Materials and Methods) when amino acids were added. The EG concentration was 20 g/L when EG was
600 added. The growth temperature was 30 ºC. Error bar indicates standard error (n=3).

601

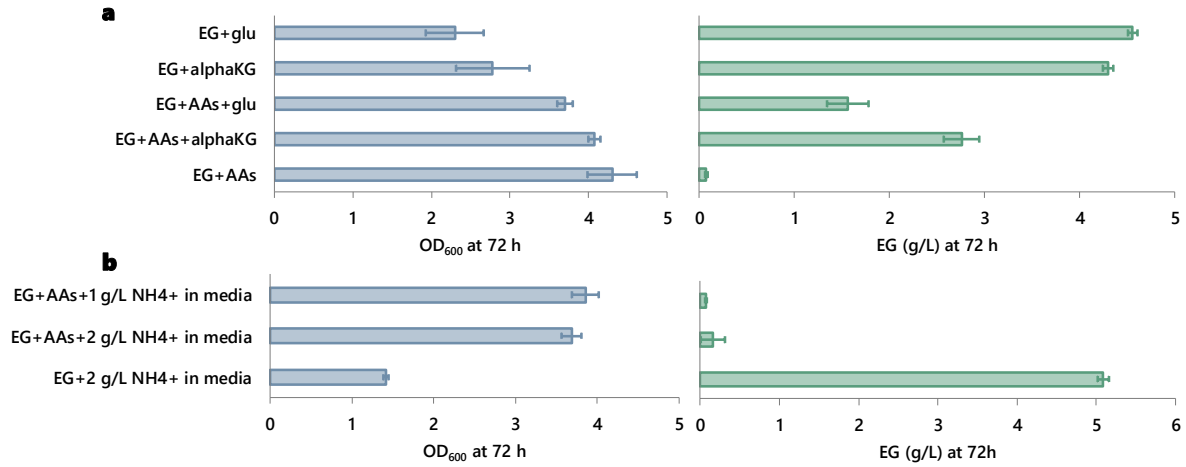


Fig. S7 (**a**) Effect of supplementation of different nitrogen sources or compounds involved in nitrogen assimilation such as alpha-ketoglutarate and glutamate on final cell density and EG utilization of **EG04**. (**b**) Effect of increasing the concentration of inorganic nitrogen source (NH_4Cl) in the medium on final cell density and EG utilization. Cell growth and EG utilization were not improved with supplementation of nitrogen sources. All the cells were cultured in a chemically defined medium (5 mL culture volume in a 50 mL culture tube) containing 5 g/L EG and 1 g/L AAs/glutamate/alpha KG. AAs: the amino acids in CSM (concentration of each amino acid was the same to that in CSM. Glu: Glutamate. The growth temperature was 30 ºC. Error bar indicates standard error (n=3).

Table S1 Plasmids constructed in this study.

Plasmids constructed	Genes	Sequences of oligos
pEG01 (P_BAD-*fucO*-*aldA*-pMB1-spect)	*fucO*	G-*fucO* F: G*gctaacagaatgattctgaacgaaacg *fucO*-T R: A*ccaggcggtatggtaaagct
	aldA	
pEG02 (P_thrC3-*fucO*-*aldA*-p15A-cam)	*fucO*	
	aldA	
pEG03 (P_gyrA-*fucO*-*aldA*-p15A-cam)	*fucO*	
	aldA	
pEG03' (P_gyrA-*fucO*-*aldA*-pMB1-spec)	*fucO*	
	aldA	

pMB1 and p15A are the high copy number and medium copy number replication origins respectively; spect and cam are spectinomycin and chloramphenicol resistance cassettes respectively.