Ethylene glycol and glycolic acid production by wild-type *Escherichia coli*

Xiyang Lu[1], YaoYao[1],Yang Yang[1], Zhongxi Zhang[1], Jinjie Gu[1,4], Ljiljana Mojovic[3], Zorica Knezevic-Jugovic[3], Frank Baganz[2], Gary Lye[2], Jiping Shi[1,4]*, Jian Hao[1,2]*

1. Lab of Biorefinery, Shanghai Advanced Research Institute, Chinese Academy of Sciences, No. 99 Haike Road, Pudong, Shanghai, 201210, PR China

2. Department of Biochemical Engineering, University College London, Gordon Street, London WC1H 0AH, UK

3. Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, 11000 Belgrade, Serbia

4. School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, PR China

*Corresponding author.

Email: haoj@sari.ac.cn

shijp@sari.ac.cn

Tel.: +86 21 20325163

Running title: EG and GA production by *E. coli*

Highlights

1. EG and GA were produced from xylonic acid by native *E. coli*

2. 10.3 g/L of glycolic acid was produced by *E. coli* Δ*yqhD*+*aldA*

3. *E. coli* Δ*aldA*+*yqhD* produced 8.0 g/L of ethylene glycol

Abstract

Ethylene glycol and glycolic acid are bulk chemicals with a broad range of applications. The ethylene glycol and glycolic acid biosynthesis pathway has been constructed in microorganisms and used as a biological route for their production. Unlike the methods that use xylose or glucose as carbon sources, xylonic acid was used as a carbon source to produce ethylene glycol and glycolic acid in this paper. Amounts of 4.2 g/L of ethylene glycol and 0.7 g/L of glycolic acid were produced by a wild-type *E. coli* W3110 within 10 hours of cultivation with a substrate conversion ratio of 0.5 mol/mol. Furthermore, *E. coli* strains that produce solely ethylene glycol or glycolic acid were constructed. 10.3 g/L of glycolic acid was produced by *E. coli* Δ*yqhD*+*aldA*, and the achieved conversion ratio was 0.56 mol/mol. Similarly, the *E. coli* Δ*aldA*+*yqhD* produced 8.0 g/L of ethylene glycol with a conversion ratio of 0.71 mol/mol. Ethylene glycol and glycolic acid production by *E. coli* on xylonic acid as a carbon source provides new information on the biosynthesis pathway of these products and opens a novel way of biomass utilization.

Key words: ethylene glycol; glycolic acid; xylonic acid; xylose; *E. coli*

1 Introduction

Ethylene glycol is the simplest diol. It has broad applications, such as antifreeze and coolant in automobiles, de-icing fluid for windshields and aircraft, desiccant for natural gas production, and a precursor for the manufacture of polyesters and resins [1]. Traditionally, ethylene glycol is produced from ethylene through a chemical route. Ethylene is a bulk chemical of the petroleum industry. The utilization of renewable resources as feedstock for ethylene glycol production has recently become a hot research point. After pretreatment, the hemicellulose from corn stalk can be converted to ethylene glycol and 1,2-propylene glycol catalysts under hydrothermal conditions and a hydrogen atmosphere [2]. With Pt-modified Ni catalyst, glycerol conversion to ethylene glycol and CH_4 by hydrocracking had been performed [3].

Beside chemical routes, biological routes for ethylene glycol production have appeared with the development of synthetic biology. Wook-Jin Chung's group from Myongji University first reported the biological route for ethylene glycol synthesis with an engineered *Escherichia coli*. A heterologous D-xylose dehydrogenase, originally from *Caulobacter crescentus,* was over-expressed in *E. coli*. D-xylose dehydrogenase catalysed xylonic acid formation from xylose. Xylonic acid in the cell was then converted to 2-dehydro-3-deoxy-D-pentonate, the later was cleaved to form pyruvate and glycolaldehyde. This pathway was named Dahms pathway. Ethylene glycol and glycolic acid were obtained from glycolaldehyde by reduction and oxidation reaction (Figure 1) [4]. Besides the Dahms pathway, a pathway through serine as an intermediate was developed with *Corynebacterium glutamicum* and *E. coli* as the host cell [5, 6].

Glycolic acid is the smallest two-carbon α-hydroxy acid with both alcohol and acid groups. It has been used in the textile industry as a dyeing and tanning agent, as a flavour and preservative in the food industry, and as a skincare agent in the pharmaceutical industry. Currently, glycolic acid in the market is produced chemically from petrochemical resources mainly in a process where formaldehyde is carbonylated by synthesis gas or treated with carbon monoxide and water [7]. Several biological routes of glycolic acid production have been reported. Glycolic acid production from ethylene glycol or glycolonitrile was a biotransformation process. Glycolic acid was obtained from ethylene glycol in the biotransformation catalyzed by the resting cells of *Gluconobacter oxydans* [8]. An *Alcaligenes sp* strain can transform glycolonitrile to glycolic acid with high efficiency [9]. Besides biotransformation, glycolic acid can be synthesized through Dahms pathway or glyoxylate pathway using glucose, xylose, or other sugar as the feedstock. Using glucose as feedstock, glycolic acid can be synthesized by *Saccharomyces*

cerevisiae or *Kluyveromyces lactis* [10]. Glycolic acid can also be synthesized by engineered *Corynebacterium glutamicum* and *E. coli* [11, 12]. A recent report from Wook-Jin Chung's group shows an increase of glycolic acid production by *E. coli* with overexpression of a membrane-bound pyridine nucleotide transhydrogenase, the substrate used was xylose [13]. Followed the Wook-Jin Chung's work, an engineered *E. coli* with deleted the acetic acid biosynthesis pathway produced a high level of glycolic acid on xylose as substrate [14].

Xylose has been used as a substrate for ethylene glycol synthesis by engineered *Escherichia coli*, and xylonic acid is an intermediate of the process [4]. Many microorganisms, including *Acetobacter sp*., *Pseudomonas aeruginosa*, *Pseudomonas fragi*, and *Gluconobacter oxydans*, have the ability to convert xylose to xylonic acid [15]. In the previous research, we developed a technology for the production of xylonic acid by *Klebsiella pneumoniae*, and this technology was highly efficient [16]. Xylonic acid can be used as a carbon source for ethylene glycol and glycolic acid production by wild type *Enterobacter cloacae*. Using this technology, high substrate conversion ratio and high productivity were achieved. [17]. *E. coli* was commonly used host for the construction of ethylene glycol producing strains. Here, ethylene glycol and glycolic acid production from xylonic acid by wild type and engineered *E. coli* was investigated in detail.

2 Materials and methods

2.1 Strains, plasmids, and primers

Bacterial strains and plasmids used in this study are listed in Table 1. Primers used for PCR are listed in Table 2.

2.2 Xylonic acid preparation.

Xylonic acid (Ammonium salt) was produced from xylose by *K. pneumoniae*, as described previously [16]. The sterilized fermentation broth was centrifuged to eliminate cells and other insoluble impurities. 1% of activated carbon was added to the supernatant and paper filtrated. The discolored liquid was concentrated to 700 g/L with a rotary evaporator at 70 °C. The xylonic acid crystals were formed after keeping the liquid at room temperature for 1 week. This xylonic acid obtained was used in the following experiments.

2.3 Construction of mutants of *E. coli*

Mutants of *E. coli* were grown in Luria–Bertani (LB) medium at 37 °C. The antibiotics used in the selective medium were kanamycin (50 μg/mL), apramycin (50 μg/mL) and streptomycin (25 μg/mL).

Construction of mutants of *E. coli* was generated by the method of Red recombinase associated gene replacement [19]. *E. coli* ΔaldA construction was described in detail as an example.

Right fragment of *aldA* was amplified from genome of *E. coli* W3110 using the primer pair aldA-F and loop-aldA-FRT-F. Left fragment of *aldA* was amplified in the same way using the primer pair aldA-R and loop-aldA-FRT-R. Apramycin resistance gene *aac(3)IV* was amplified with plasmid pIJ773 as a template using the primer pair aldA-FRT-F/aldA -FRT-R. Right and left fragments of *aldA* and apramycin resistance gene fragment were mixed together with the molecular ratio of 1: 1: 1 and linked together using the kit of PrimeSTAR ® (TaKaRa Code: DR040A). The linked linear DNA (2μg/mL) was transformed into electrocompetent cells of *E. coli* W3110, which already hosted the plasmid pDK6-red. Homologous recombination between the linear DNA and the chromosome was facilitated by Red recombinase and led to *aldA* deletion in the *E. coli*. The colonies grown on the selective medium plates were confirmed by PCR, and primer pair aldA-check-F and aldA-check-R were used.

2.4 Construction of *E. coli* /aldA and other protein over-expression strains

The ORF of *aldA* in *E. coli* W3110 was amplified using the primer pair OE-aldA-F and OE-aldA-F, which both contain 21 nt homologous sequence of the pRNA. pRNA is an expression plasmid that contains a Phya promoter for continuous expression of the target protein. Phya promoter was a hydrogenase promoter of *E. coli*. The ORF of *aldA* was ligated into pRNA to generate pRNA-aldA. pRNA-aldA was transformed into *E. coli* W3110 to generate *E. coli* +aldA. *E. coli* +yqhD was constructed using the same approach.

2.5 Physiological characterizations of strains

The wild-type and the constructed strains were inoculated in 250 ml flasks containing 50 ml fermentation medium or M9 medium (either 5 g/L of xylose or 5 g/L of xylonic acid used as carbon sources) and incubated on a rotary shaker at 37 °C and 120 rpm for 1 day. All experiments were done in triplicate, and data are expressed as the mean ± standard error. Chemical compounds in the broth were quantified by a Shimadzu 20AVP high performance liquid chromatograph system (HPLC) equipped with a RID-20A refractive index detector and a SPD-M20A photodiode array detector. An Aminex HPX-87H column (300×7.8 mm) (Bio-Rad, USA) was used and the column temperature was set at 65 ºC. The mobile phase was 0.005 mol/L H_2SO_4 solution with a flow rate of 0.8 ml/min.

2.6 Culture conditions

Stirred tank bioreactors were used to optimize the culture parameters. For the seed culture, 250-mL flasks containing 50 mL of LB medium were incubated on a rotary shaker at 37 °C and 200 rpm overnight. The

seed culture was inoculated into a 5-L bioreactor (BIOSTAT-A plus Sartorius) with a working volume of 3 L and air flow rate of 2 L/min. Culture pH and the stirring rate were optimized individually. Experiments with different culture pH were done once at the time. Experiments with different stirring rates were done in triplicate, and data were expressed as the mean ± standard error.

The fermentation medium contained: xylonic acid 30 g/L, corn steep liquor 4 g/L, $(NH_4)_2SO_4$ 5 g/L, KCl 0.4 g/L, and $MgSO_4$ 0.1 g/L.

Fed batch cultures were performed under optimal conditions, with culture pH=6.5, culture temperature 37 °C and agitation rate of 600 rpm. Experiments were done in triplicate, and data were expressed as the mean ± standard error.

3 Results

3.1 *E. coli* strain selection.

W3110 and BW25113 are commonly used *E. coli* strains in biotechnology and both have been used for the construction of ethylene glycol producing strains. The wild-type strains were selected to test the ability for ethylene glycol and glycolic acid production using xylonic acid as the feedstock. The cells were cultured in shake flasks with fermentation medium for one day, and the metabolic products obtained in the fermentation broth are shown in Table 3.

Both wild-types of *E.coli* W3110 and BW25113 could synthesize ethylene glycol and glycolic acid using xylonic acid as the substrate. *E. coli* W3110 produced higher amount of ethylene glycol and glycolic acid, thus this strain was selected for further investigation.

3.2 Role of AldA and YqhD in ethylene glycol and glycolic acid synthesis

A pathway for biosynthesis of ethylene glycol and glycolic acid that uses xylose as the sole carbon source had been constructed. In this pathway, the genes for lactaldehyde dehydrogenase and aldehyde reductase, encoded as *aldA* and *yqhD,* were found responsible for the synthesis of glycolic acid and ethylene glycol from glycolaldehyde [4]. The two genes were individually knocked out in *E. coli* W3110 to obtain *E. coli* ΔaldA and *E. coli* ΔyqhD, respectively. A strain with double gen deletion - *E. coli* ΔaldA-ΔyqhD was also constructed. Thus constructed strains were further used as producers of ethylene glycol and glycolic acid in shake flasks, and the results are shown in Figure 2. Wild type *E. coli* W3110 was used as a control. Minimal M9 medium with xylonic acid as the sole carbon source was used.

As shown in Figure 2, the xylonic acid can be used as the sole carbon source for the growth of wild type *E. coli*. It was exhausted after 10 hours of cultivation, and 0.5 g/L of ethylene and 0.8 g/L of glycolic acid were produced. The growth rate of *E. coli* ΔyqhD was similar to the wild type strain when compared

to other engineered strains. However, the differences in xylonic acid consumption rates were significant, and 1.6 g/L of xylonic acid remained in the fermentation broth after 12 hours of cultivation of *E. coli* ΔyqhD. The ethylene glycol synthesis ceased in *E. coli* ΔyqhD, while glycolic acid synthesis also decreased compared to the wild-type strain. In contrast, *E. coli* ΔaldA could not grow at the basic medium with xylonic acid as the sole carbon source. *E. coli* ΔaldA-ΔyqhD also could not grow at this basic medium.

To further investigate the roles of *yqhD* and *aldA* on the growth of *E. coli*, the strains were cultured in shake flasks with the basic medium using xylose as the sole carbon source, and the results are shown in Figure 3.

All the strains were able to growth on xylose as the sole carbon source. The main products of these strains were lactic acid and acetic acid. The growth rate and xylose consumption rate of *E. coli* ΔyqhD were almost the same as that of the wild type strain. The xylose consumption rates of *E. coli* ΔaldA and *E. coli* ΔaldA-ΔyqhD had an 8 h delay compared to that of the wild type strain. However, the final cell densities of *E. coli* ΔaldA and *E. coli* ΔaldA-ΔyqhD reached 3.8 and 3.2, which were higher than that of wild type and *E. coli* ΔyqhD (2.0 and 1.8).

3.3 Ethylene glycol and glycolic acid production by strains with overexpression of *aldA*

AldA is a key enzyme in glycolic acid synthesis. The open reading frame (ORF) of *aldA* was ligated to overexpression plasmid and transformed into wild type *E. coli* W3110 and *E. coli* ΔyqhD to generate *E. coli* +aldA and *E. coli* ΔyqhD+aldA. These strains were cultured in shake flasks with fermentation medium and the results are shown in Figure 4.

After 10 hours of cultivation, the wild-type strain consumed 6 g/L of xylonic acid and synthesized 0.8 g/L of ethylene glycol and 0.4 g/L of glycolic acid. Furthermore, xylonic acid was exhausted by *E. coli* +aldA after 10 hours of cultivation, with 2.3 g/L of glycolic acid synthesized. Amount of 0.03 g/L of ethylene glycol was detected after 2 hours of cultivation; however, it was reused by the cells. Xylonic acid consumption by *E. coli* ΔyqhD was weaker than that of the wild-type strain, and only 0.9 g/L of glycolic acid was synthesized after 18 hours of cultivation. As expected, no ethylene glycol was synthesized by *E. coli* ΔyqhD. The amount of glycolic acid synthesized by *E. coli*ΔyqhD+aldA was higher than that of *E. coli* ΔyqhD but lower than amount produced by *E. coli* +aldA. Xylonic acid was not exhausted in the culture of *E. coli* ΔyqhD and *E. coli*ΔyqhD+aldA. The cell growths coincided with the xylonic acid consumption, with highest cell density achieved by *E. coli*+aldA. The largest conversion ratio of xylonic acid to glycolic acid of 0.41 mol/mol was obtained by *E. coli* ΔyqhD+aldA, while *E.*

coli+aldD accomplished lower conversion ratio (0.2 mol/mol). Thus *E. coli* ΔyqhD+aldA was selected for the following experiments in bioreactor.

3.4 Ethylene glycol and glycolic acid production by strains with overexpression of *yqhD*

Similarly to aldA, yqhD is a key enzyme of ethylene glycol synthesis. The *yqhD* gene was overexpressed in the wild type *E. coli* W3110 and *E. coli* ΔaldA to generate *E. coli* +yqhD and *E. coli* ΔaldA+yqhD. These engineered strains were cultured in shake flasks with fermentation medium and the results are shown in Figure 5.

Xylonic acid consumption by *E. coli* ΔaldA, *E. coli* +yqhD, and *E. coli* ΔaldA+yqhD were all slower than that of the wild-type strain. The cell growth coincided with xylonic acid consumption, and the wild-type strain grew quickly during the beginning 4 hours of cultivation. Ethylene glycol synthesis by *E. coli* +yqhD was the highest among all tested strains, as well as the production of glycolic acid. In contrast, the production of glycolic acid was blocked in *E. coli* ΔaldA and *E. coli* ΔaldA+yqhD. However, the production of ethylene glycol by these two strains was not increased.

3.5 Culture parameters optimization

Based on the shake flasks data *E. coli* W3110 was further cultured in 5L stirred tank bioreactor to improve ethylene glycol and glycolic acid production. The culture pH was controlled at 6.0, 6.5, 7.0 and 7.5, respectively while the stirring rate was maintained constant at 500 rpm. The cell growth and concentrations of metabolic products are presented in Figure 6.

The cell growth was very weak at pH 5.5, and the xylonic acid was not consumed by the cells. The final cell density and the xylonic acid consumption at pH 7.5 were also very low. The cells grew rather well in the culture media at pH range 6.0-7.0. The fastest xylonic acid consumption rate was achieved at culture pH 6.5 and ethylene glycol and glycolic acid productivity was the highest accordingly. Thus, pH 6.5 was selected as an optimal culture pH.

Oxygen supplementation is a key parameter for cell growth and product synthesis. Appropriate mixing conditions are enabling good mass transfer of microbial substrates and products as well as an adequate oxygen transfer rate. The stirring rate was varied from 200, 400, 600 up to 800 rpm to provide micro-aerobic conditions at the lowest rate, up to fully aerobic conditions at the highest rate, while the culture pH was kept constant at 6.5. Fermentation results of *E. coli* W3110 at different stirring rates are presented in Figure 7.

The cell growth rate had a positive correlation with the stirring rate. The highest cell densities were achieved at stirring rates of 600 rpm and 800 rpm, while the lowest cell density (OD 5.6) was obtained

at the stirring rate of 200 rpm. The trend of xylonic acid consumption was similar to that of the cell growth, the cells grown at 600 rpm exhibited the fastest xylonic acid consumption rate, and those grown at 200 rpm had the lowest xylonic acid consumption rate. Ethylene glycol and glycolic acid production increased with an increase of stirring rate from 200 to 600 rpm. However, at the lowest stirring rate of 200 rpm, the products were not synthesized. Xylonic acid consumption and ethylene glycol and glycolic acid production were rather close at 600 rpm and 800 rpm. Therefore, 600 rpm was selected as an optimal stirring rate. Under the optimal conditions, 4.2 g/L of ethylene glycol and 0.7 g/L of glycolic acid were produced by the wild type strain with 23 g/L of xylonic acid consumed. The total substrate conversion ratio for ethylene glycol and glycolic acid synthesis from xylonic acid was 0.50 mol/mol.

3.6 Individual production of ethylene glycol or glycolic acid by engineered strains in batch fermentation

The *e*ngineered strains *E. coli* ΔaldA and *E. coli* ΔaldA+yqhD solely synthesized ethylene glycol. On the other hand, *E. coli* ΔyqhD and *E. coli* ΔyqhD+aldA solely synthesized glycolic acid. These strains were cultured in bioreactors under the optimal conditions and the results are shown in Figure 8.

After 6 hours in lag phase, *E. coli* ΔaldA transitioned into the exponential phase and xylonic acid was quickly consumed by the cells. The substrate - xylonic acid was exhausted after 12 hours of cultivation and 5.3 g/L of ethylene glycol was synthesized. The lag phase was extended to 11 hours in the culture of *E. coli* ΔaldA+yqhD. Xylonic acid was exhausted after 22 hours of cultivation by this strain, and 8.0 g/L of ethylene glycol was produced. The conversion ratio of ethylene glycol synthesis from xylonic acid was 0.71 mol/mol.

E. coli ΔyqhD cell grew slowly during the first 9 hours of cultivation and then transitioned into the exponential phase. Xylonic acid consumption by the cell started at 12 hours of cultivation, which was 3 hours later. The substrate consumption stopped after 21 hours of cultivation, and 4.3 g/L of glycolic acid was produced. The cell growth and xylonic acid consumption were delayed up to 12 and 15 hours of cultivation in the culture of *E. coli* ΔyqhD+aldA. Xylonic acid was exhausted after 27 hours and 10.3 g/L of glycolic acid was produced. Achieved conversion ratio of glycolic acid synthesis from xylonic acid was 0.56 mol/mol.

4 Discussion

4.1 Ethylene glycol and glycolic acid synthesis by *E. coli* on xylonic acid as a carbon source.

In the first report of ethylene glycol production by biological route, *E. coli* was used as the producer on

the xylose as a carbon source [4]. There are a few reports of ethylene glycol synthesis by *E. coli* using other pentoses as carbon sources [20, 21]. All of these studies report using genetically modified *E. coli* strains that block the D-xylose metabolism through the pentose phosphate pathway. In wild-type *E. coli*, the xylose was metabolized through the pentose phosphate pathway where the final products were lactic acid and acetic acid (Figure 3).

In our previous research, xylonic acid production by *K. pneumoniae* was developed, and this process had a high conversion ratio and productivity. Glucose dehydrogenase was identified as the enzyme that catalyzes the reaction in *K. pneumonia*. It is located in the inner membrane of the periplasmic space and uses pyrroloquinoline quinine (PQQ) as a cofactor [16]. The xylose obtained from the hydrolysis of bamboo can also be used as a feedstock for xylonic acid production [22]. In the present study, ethylene glycol and glycolic acid could be synthesized by *E. coli* W3110 and BW25113 on xylonic acid as a carbon source. (Table 3). It is confirmed that the xylonic acid can be transferred into *E. coli* cell and catabolized through Dahms pathway for ethylene glycol and glycolic acid production. This is a novel route of xylose utilization and it prevents the xylose catabolism through the pentose phosphate pathway.

4.2 Role of AldA and YqhD in xylose and xylonic acid catabolism

aldA and *yqhD* have been pointed responsible for glycolic acid and ethylene glycol synthesis from glycolaldehyde in many reports exploring biological routes of ethylene glycol production [4]. The *aldA* encodes lactaldehyde dehydrogenase. This enzyme catalyzes the conversion of lactaldehyde to lactic acid, and also the conversion of glycolaldehyde to glycolic acid. The enzyme is a key enzyme for catabolism of 1,2-propanediol and ethylene glycol in *E. coli* [23]. *E. coli* ΔaldA was grown with xylose as the sole carbon source but without xylonic acid (Figure 2, 3). The results indicated that the critical reaction for glycolic acid formation from glycolaldehyde is utilization of xylonic acid, since there was not *aldA* isoenzyme in *E. coli*. The amount of lactic acid produced by *E. coli* ΔaldA and *E. coli* ΔaldA-ΔyqhD was increased, compared with the wild-type strain (Figure 3). This indicated that the lactic acid synthesized in the processes was not from lactaldehyde. It has been recorded that the substrate of AldA was not very strict, and AldA appeared to function as a detoxifying enzyme converting various aldehydes into their corresponding carboxylic acids [24]. The delay of xylose utilization by *aldA* inactivated strains might be due to the fact that the detoxifying ability of these strains was weak, as inactivity of *aldA* was not directly related to enzymes responsible for xylose catabolism. The increase of final cell density in the culture of *E. coli* ΔaldA and *E. coli* ΔaldA-ΔyqhD shown in Figure 3, indicates that the function of AldA in the cell was not straightforward; it has also other important functions in cell metabolism.

The *yqhD* is encoding an aldehyde reductase, also called alcohol dehydrogenase. This enzyme use NADPH as the cofactor and its substrates are very broad in range. The YqhD has been used to catalyze the reaction of 1,3-propanediol formation from 3-hydroxypropionaldehyde in an engineered *E. coli* strain [25]. This enzyme also catalyzes isobutanol formation from isobutyraldehyde in *K. pneumoniae* [26]. Inactivity of *yqhD* lead to losing the cell ability to synthesize ethylene glycol in M9 medium and therefore a very low level of ethylene glycol was present in the fermentation broth (Figure 2, 8). Thus, YqhD is the main enzyme that catalyzes the ethylene glycol formation in *E. coli*. This is in agreement with the report of Alkim et. al, that the ethylene glycol production by *yqhD* knockout strain was decreased to 30% of the wild type strain [21]. However, some reports pointed out that *fuco* encoded lactaldehyde reductase was responsible for ethylene glycol synthesis from glycolaldehyde [20]. In a report on glycolic acid production on glucose as the substrate, glycolic acid was synthesized from glyoxylate via glyoxylate cycle [12]. The glyoxylate was synthesized from pyruvate. In this report, the pyruvate obtained from 2-dehydro-3-deoxy-D-pentonate was all consumed by the cell. If the pyruvate can be converted to glycolic acid by overexpressing glyoxylate pathway enzymes, the total conversion ratio of glycolic acid from xylonic acid would increase.

4.3 Effect of the culture pH and oxygen supplementation on ethylene glycol and glycolic acid synthesis

Xylonic acid is a kind of organic acid, and ammonium xylonate was used in this study With xylonic acid consumption, the pH values of the culture broth will increase. The effect of buffer in the broth of the flasks culture was very weak, so xylonic acid had not been exhausted in most of the flask cultures (Figure 4, 5).

The culture pH in bioreactor experiments was kept constant by automatically feeding HCl. When the culture pH value was maintained at 7.5, the growth of cell was very weak and the xylonic acid was not consumed. At the culture pH 5.5, the cell grew to a high density. However, xylonic acid was not consumed (Figure 6). It indicated that the enzymes responsible for xylonic acid catabolism were not expressed or were not active under these culture conditions.

Ethylene glycol and glycolic acid synthesis from glycolaldehyde were obtained by reduction and oxidization reactions, respectively. Therefore, the levels of both ethylene glycol and glycolic acid production increased under the conditions of high oxygen supplementation. In a report of ethylene glycol production from xylose via xylulose-1P as an intermediate by *E. coli*, the ethylene glycol production achieved under fully aerobic conditions was higher than that under microaerobic conditions. However,

glycolic acid production was in a reversed relationship, a high level of glycolic acid was produced under microaerobic conditions [21]. Ethylene can be synthesized from glucose through serine as an intermediate; the fermentation was also facilitated under high aerobic conditions [6]. Similarly, glycolic acid production from xylose through Dahms pathway by *E. coli* was an aerobic process [14].

4.4 The metabolic pathway of xylonic acid catabolism and ethylene glycol and glycolic acid synthesis

Under lower oxygen supplementation, some xylonic acid was consumed without the production of ethylene glycol or glycolic acid (Figure 7). The conversion ratio of the produced ethylene glycol plus glycolic acid from xylonic acid by the wild type strain under the optimal culture conditions was 0.5 mol/mol (Figure 7). The conversion ratios of the individual productions of ethylene glycol or glycolic acid from xylonic acid in the culture of *E. coli* ΔyqhD+aldA and *E. coli* ΔaldA+yqhD were 0.56 and 0.71 mol/mol (Figure 8). All conversion ratios were less than 1, suggesting that a fraction of intermediates of the ethylene glycol and glycolic acid synthesis from xylonic acid might be converted to other products. This is different from ethylene glycol and glycolic acid production by wild-type *Enterobacter cloacae*, which has a total conversion ratio of nearly 1 mol/mol [17]. It has been mentioned that glycolic acid can be converted to glyoxylate into the glyoxylate cycle [20]. The reaction of conversion of glycolic acid to glyoxylate is reversible, and this reaction has been used to improve the glycolic acid production from glucose or xylose [14, 27]. However, this cannot explain the conversion ratio of ethylene glycol production from xylonic acid by strain *E. coli* ΔaldA+yqhD, in which no glycolic acid was produced. Thus, it was suspected that xylonic acid or intermediates of the Dahms pathway were catabolised through unknown pathways, or that some glycolic acid was synthesized by isoenzymes of AldA.

There are numerous reports on ethylene glycol production using xylose as a substrate and *E. coli* as the producer. The conversion ratios were in the ranges of 0.93-0.94 mol/mol, and the final titers obtained in bioreactors were in the ranges of 7.72-108.18 g/L [28]. A conversion ratio for glycolic acid production from xylose reached 1.24 mol/mol [20]. The conversion ratios of ethylene glycol or glycolic acid production from xylonic acid obtained in this research were lower than most of these previous reports. However, the unknown metabolic pathway that divided carbon flux and reduced the conversion of ethylene glycol production was not mentioned in these previous reports. This is interesting for further investigation, and it would be beneficial to reveal the unknown pathway of all biological routes of ethylene glycol and glycolic acid production.

E. coli is a model bacterium in biotechnology and it is commonly used to study various metabolic

mechanisms and possibilities. Prevent xylonic acid or intermediates of the Dahms pathway catabolized through other pathways would increase the conversion ratio. However, *E. coli* might not be the best workhorse for ethylene glycol and glycolic acid production. In our recent research, *Enterobacter cloacae* has shown higher efficiency in ethylene glycol and glycolic acid production on xylonic acid than the *E. coli* here [17]. The traditional way of biological utilization of biomass includes the hydrolysis of biomass to monosaccharides, which are further used as a carbon source by microorganisms for their growth and production of various products. The biomass pretreatment and hydrolysis is a high cost step and limits the whole process of biomass utilization. If the xylonic acid can be obtained directly from biomass, pretreatment and hydrolysis could be avoided. This would be a novel way of biomass utilization and have challenges different from the traditional way of biomass hydrolysis and pretreatment.

5 Conclusions

Here, an ethylene glycol and glycolic acid production method with a wild type *E. coli* as a producer was established. Furthermore, ethylene glycol solely producing strain was constructed by knockout the *aldA* and was combined with overexpression of *yqhD* to further increase the ethylene glycol level. Accordingly, the ethylene glycol synthesis was ceased by knockout of *yqhD* and this strain synthesized glycolic acid only. Similarly, overexpression of *aldA* enhanced the glycolic acid production. In this paper, combined or individual production technology of ethylene glycol and glycolic acid were provided. Analysis of the achieved conversion ratios obtained with engineered *E. coli* strains and its comparison with well known metabolic pathway indicated that the intermediates of ethylene glycol and glycolic acid biosynthesis pathway could be catabolised through still unknown routes.

Declaration of interest

The authors declare that they have no competing interests

Acknowledgement

This work was supported by National Key R&D Program of China (Grant No. 2017YFE0112700), Royal Society-Newton Advanced Fellowship (Grant No. NAF\R2\180721), Natural Science Foundation of Shanghai (Grant No. 19ZR1463600) and National Natural Science Foundation of China (Grant No. 21576279).

References

1 Yue, H., Zhao, Y., Ma, X. and Gong, J. (2012) Ethylene glycol: properties, synthesis, and

applications. Chem. Soc. Rev. 41, 4218-4244

2 Pang, J., Zheng, M., Wang, A. and Zhang, T. (2011) Catalytic Hydrogenation of Corn Stalk to Ethylene Glycol and 1,2-Propylene Glycol. Ind. Eng. Chem. Res. 50, 6601-6608

3 Ueda, N., Nakagawa, Y. and Tomishige, K. (2010) Conversion of Glycerol to Ethylene Glycol over Pt-modified Ni Catalyst. Chem. Lett. 39, 506-507

4 Liu, Ramos, Kristine, R. M., Valdehuesa, Kris, N. G., Nisola, Grace, M., Chung and Wook-Jin. (2013) Biosynthesis of ethylene glycol in *Escherichia coli*. Appl. Microbiol. Biotechnol. 97, 3409-3417

5 Chen, Z., Huang, J., Wu, Y. and Liu, D. (2016) Metabolic engineering of *Corynebacterium glutamicum* for the de novo production of ethylene glycol from glucose. Metab. Eng. 33, 12-18

6 Pereira, B., Zhang, H., De, M. M., Lim, C. G., Li, Z. J. and Stephanopoulos, G. (2016) Engineering a novel biosynthetic pathway in *Escherichia coli* for production of renewable ethylene glycol. Biotechnol. Bioeng. 113, 376-383

7 Salusjarvi, L., Havukainen, S., Koivistoinen, O. and Toivari, M. (2019) Biotechnological production of glycolic acid and ethylene glycol: current state and perspectives. Appl. Microbiol. Biotechnol. 103, 2525-2535

8 Wei, G., Yang, X., Gan, T., Zhou, W., Lin, J. and Wei, D. (2009) High cell density fermentation of *Gluconobacter oxydans* DSM 2003 for glycolic acid production. J. Ind. Microbiol. Biotechnol. 36, 1029-1034

9 He, Y., Xu, J., Su, J. and Zhou, L. (2010) Bioproduction of Glycolic Acid from Glycolonitrile with a New Bacterial Isolate of *Alcaligenes sp*. ECU0401. Appl. Biochem. Biotechnol. 160, 1428-1440

10 Koivistoinen, O. M., Kuivanen, J., Barth, D., Turkia, H., Pitkänen, J.-P., Penttilä, M. and Richard, P. (2013) Glycolic acid production in the engineered yeasts *Saccharomyces cerevisiae* and Kluyveromyces lactis. Microb. Cell Factories. 12, 1

11 Zahoor, A., Otten, A. and Wendisch, V. F. (2014) Metabolic engineering of *Corynebacterium glutamicum* for glycolate production. J. Biotechnol. 192, 366-375

12 Deng, Y., Ma, N., Zhu, K., Mao, Y., Wei, X. and Zhao, Y. (2018) Balancing the carbon flux distributions between the TCA cycle and glyoxylate shunt to produce glycolate at high yield and titer in *Escherichia coli*. Metab. Eng. 46, 28-34

13 Cabulong, R. B., Valdehuesa, K. N. G., Banares, A. B., Ramos, K. R. M., Nisola, G. M., Lee, W. and Chung, W. (2019) Improved cell growth and biosynthesis of glycolic acid by overexpression of membrane-bound pyridine nucleotide transhydrogenase. J. Ind. Microbiol. Biotechnol. 46, 159-169

14 Liu, M., Ding, Y., Xian, M. and Zhao, G. (2018) Metabolic engineering of a xylose pathway for biotechnological production of glycolate in *Escherichia coli*. Microb. Cell Factories. 17, 51

15 Buchert, J., Viikari, L., Linko, M. and Markkanen, P. (1986) Production of xylonic acid by *Pseudomonas fragi*. Biotechnol. Lett. 8, 541-546

16 Wang, C., Wei, D., Zhang, Z., Wang, D., Shi, J., Kim, C. H., Jiang, B., Han, Z. and Hao, J. (2016) Production of xylonic acid by *Klebsiella pneumoniae*. Appl. Microbiol. Biotechnol. 100, 10055-10063

17 Zhang, Z., Yang, Y., Wang, Y., Gu, J., Lu, X., Liao, X., Shi, J., Kim, C. H., Lye, G., Baganz, F. and Hao, J. (2020) Ethylene glycol and glycolic acid production from xylonic acid by *Enterobacter cloacae*. Microb. Cell Factories. 19, 89

18 Gust, B., Challis, G. L., Fowler, K., Kieser, T. and Chater, K. F. (2003) PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc. Natl. Acad. Sci. U.S.A. 100, 1541-1546

19 Wei, D., Wang, M., Shi, J. and Hao, J. (2012) Red recombinase assisted gene replacement in

Klebsiella pneumoniae. J. Ind. Microbiol. Biotechnol. 39, 1219-1226

20 Pereira, B., Li, Z.-J., De Mey, M., Lim, C. G., Zhang, H., Hoeltgen, C. and Stephanopoulos, G. (2016) Efficient utilization of pentoses for bioproduction of the renewable two-carbon compounds ethylene glycol and glycolate. Metab. Eng. 34, 80-87

21 Alkim, C., Cam, Y., Trichez, D., Auriol, C., Spina, L., Vax, A., Bartolo, F., Besse, P., François, J. M. and Walther, T. (2015) Optimization of ethylene glycol production from (D)-xylose via a synthetic pathway implemented in *Escherichia coli*. Microb. Cell Factories,14,1(2015-09-04). 14, 127

22 Wei, D., Gu, J., Zhang, Z., Wang, C., Wang, D., Kim, C. H., Jiang, B., Shi, J. and Hao, J. (2017) Production of Chemicals by *Klebsiella pneumoniae* Using Bamboo Hydrolysate as Feedstock. J. Vis. Exp. e55828

23 Caballero, E., Baldoma, L., Ros, J., Boronat, A. and Aguilar, J. (1983) Identification of lactaldehyde dehydrogenase and glycolaldehyde dehydrogenase as functions of the same protein in *Escherichia coli*. J. Biol. Chem. 258, 7788-7792

24 Rodriguezzavala, J. S., Allalihassani, A. and Weiner, H. (2006) Characterization of *E. coli* tetrameric aldehyde dehydrogenases with atypical properties compared to other aldehyde dehydrogenases. Protein Sci. 15, 1387-1396

25 Nakamura, C. E. and Whited, G. M. (2003) Metabolic engineering for the microbial production of 1,3-propanediol. Curr. Opin. Biotechnol. 14, 454-459

26 Gu, J., Zhou, J., Zhang, Z., Kim, C. H., Jiang, B., Shi, J. and Hao, J. (2017) Isobutanol and 2-ketoisovalerate production by *Klebsiella pneumoniae* via a native pathway. Metab. Eng. 43, 71-84

27 Deng, Y., Mao, Y. and Zhang, X. (2015) Metabolic engineering of *E. coli* for efficient production of glycolic acid from glucose. Biochem. Eng. J. 103, 256-262

28 Chae, T. U., Choi, S. Y., Ryu, J. Y. and Lee, S. Y. (2018) Production of ethylene glycol from xylose by metabolically engineered *Escherichia coli*. Aiche J. 64, 4193-4200

Table 1 Strains and plasmids

Strains or plasmids	Relevant genotype and description	Reference or source
E. coli DH5α	F+ supE44 ΔlacU169 (ΔlacZΔM15) hsdR17 recA1 endA1 gyrA96	Lab stock
E .coli W3110	F- λ-rph-1 INV (rrnD, rrnE)	Lab stock
E. coli BW25113	Δ(araD-araB) 567, ΔlacZ4787 (::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB) 568, hsdR 514	Lab stock
E. coli ΔaldA	*E. coli* W3110 ΔaldA, Apr^r	This work
E. coli ΔyqhD	*E. coli* W3110 ΔyqhD, Str^r	This work
E. coli ΔaldAΔyqhD	*E. coli* W3110 ΔaldA, Apr^r, ΔyqhD, Str^r	This work
E. coli +aldA	*E. coli* W3110, carries pRNA-aldA, Kan^r	This work
E. coli +yqhD	*E. coli* W3110, carries pRNA-yqhD, Kan^r	This work
E. coli ΔaldA+yqhD	*E. coli* W3110, ΔaldA, Apr^r, carries pRNA-yqhD, Kan^r	This work
E .coli ΔyqhD+aldA	*E. coli* W3110, ΔyqhD, Str^r, carries pRNA-aldA, Kan^r	This work
pIJ773	Apr^r , *aac(3)IV* with FRT sites 4334 bp	[18]
pIJ778	Str^r , *aadA* FRT sites, 4337 bp	[18]
pDK6-red	Kan^r, carries λ-Red genes (gam, bet, exo), 7.1 kbp	[19]
pRNA	Kan^r, carries, 3975bp	Lab stock
pRNA-yqhD	Kan^r, carries yqhD genes, 4679bp	This work
pRNA-aldA	Kan^r, carries aldA genes, 4934bp	This work

Table 2 Primers

Primer name	Sequence (5'-3')
aldA-FRT-F	GTTACCTGGCGTGGAGACGCATGGATTGATGTGGTAATGATTCCG GGGATCCGTCGACC
aldA-FRT-R	GTGAAAACAGGTTCGGTCACTGAACTGGTATCGATTCTCATGTAG GCTGGAGCTGCTTC
aldA-F	ACTACAACACTATCCGCACCAC
aldA-R	GCTTTTATACCTCCGCCGAGA
loop-aldA-FRT-F	GAAGCAGCTCCAGCCTACATGAGAATCGATACCAGTTCAGTGACC GAACCTGTTTTCAC
loop-aldA-FRT-R	GGTCGACGGATCCCCGGAATCATTACCACATCAATCCATGCGTCTC CACGCCAGGTAAC
aldA-check-F	GCCATAAATGTTATCGGACAGT
aldA-check-R	ACGGAAGATTCACTTATCGTTG
yqhD-F	ATCTGTTTGCCGAGAATACGC
yqhD-R	ATGCCTTTCCATGCTTCGAC

yqhD-FRT-F	ATTTTGTAGCATTTCTCCAGCACTCTGGAGAAATAGATGATTCCGGGGATCCGTCGACC
yqhD-FRT-R	GAACTTAAGTCTGGACGAAATGCCCGAAAACGAAAGTCATGTAGGCTGGAGCTGCTTC
loop-yqhD-FRT-F	GAAGCAGCTCCAGCCTACATGACTTTCGTTTTCGGGCATTTCGTCCAGACTTAAGTTC
loop-yqhD-FRT-R	GGTCGACGGATCCCCGGAATCATCTATTTCTCCAGAGTGCTGGAGAAATGCTACAAAAT
yqhD-check-F	CGATACGCTCATGTTGGCTT
yqhD-check-R	CAATTTCGCCGAGTTCGTCT
OE-aldA-F	GGAGGAGAGACGTGCCATATGATGTCAGTACCCGTTCAACATCC
OE-aldA-R	ATGAGCGGATACATACTCGAGTTAAGACTGTAAATAAACCACCTGGG
OE-yqhD-F	GGAGGAGAGACGTGCCATATGATGAACAACTTTAATCTGCACACCC
OE-yqhD-R	ATGAGCGGATACATACTCGAGTTAGCGGGCGGCTTCGTA
OE-check-F	TGTGCAAAAGTTTCACTACGC
OE-check-R	TTCTCACCGGATTCAGTCGTC

Table 3. Ethylene glycol and glycolic acid production by *E. coli* strains

Strains	Cell density (OD 600 nm)	Metabolic products (g/L)	
		Ethylene glycol	Glycolic acid
BW25113	5.1± 0.4	1.0± 0.2	0.2± 0.2
W3110	4.9± 0.6	2.2± 0.6	3.1± 0.7

Figure captions

Figure 1. Metabolic pathway of ethylene glycol and glycolic acid synthesis in *E. coli*

Figure 2. Ethylene glycol and glycolic acid synthesis by *E. coli* strains in M9 medium with xylonic acid as the sole carbon source.
Wt: *E. coli* W3110; ΔaldA: *E. coli* ΔaldA; ΔyqhD: *E. coli* ΔyqhD; ΔaldA-ΔyqhD: *E. coli* ΔaldA-ΔyqhD
Cultures were done in 250 ml flasks containing 50 ml medium and at 37 °C and 120 rpm. Data points are the average of n = 3; error bars represent standard error about the mean.

Figure 3. Cultivation of *E. coli* strains in M9 medium with xylose as the sole carbon source.
Wt: *E. coli* W3110; ΔaldA: *E. coli* ΔaldA; ΔyqhD: *E. coli* ΔyqhD; ΔaldA-ΔyqhD: *E. coli* ΔaldA-ΔyqhD
Cultures were done in 250 ml flasks containing 50 ml medium and at 37 °C and 120 rpm. Data points are the average of n = 3; error bars represent standard error about the mean.

Figure 4. Ethylene glycol and glycolic acid synthesis by *aldA* overexpression strains in fermentation medium.
Wt: *E. coli* W3110; Wt+aldA: *E. coli* +aldA; ΔyqhD: *E. coli* ΔyqhD; ΔyqhD+aldA: *E. coli* ΔyqhD+aldA
Cultures were done in 250 ml flasks containing 50 ml medium and at 37 °C and 120 rpm. Data points are the average of n = 3; error bars represent standard error about the mean.

Figure 5. Ethylene glycol and glycolic acid synthesis by overexpression of *yqhD E. coli* strains in the fermentation medium.
Wt: *E. coli* W3110; Wt+yqhD: *E. coli* +yqhD; ΔaldA: *E. coli* ΔaldA; ΔaldA+yqhD: *E. coli* ΔaldA+yqhD
Cultures were done in 250 ml flasks containing 50 ml medium and at 37 °C and 120 rpm. Data points are the average of n = 3; error bars represent standard error about the mean.

Figure 6. Ethylene glycol and glycolic acid synthesis by *E. coli* W3110 at different culture pH.
Cultures were performed in 5-L bioreactors with a working volume of 3 L, air flow rate of 2 L/min and stirring rate of 500 rpm.

Figure 7. Ethylene glycol and glycolic acid synthesis by *E. coli* W3110 at different stirring rates. Cultures were done in 5-L bioreactors with a working volume of 3 L and air flow rate of 2 L/min and culture pH 6.5.

Figure 8. Solely production of ethylene glycol or glycolic acid by engineered strains of *E. coli*. ΔaldA: *E. coli*ΔaldA; ΔaldA+yqhD: *E. coli*ΔaldA+yqhD; ΔyqhD: *E. coli* ΔyqhD; ΔyqhD+aldA: *E. coli* ΔyqhD+aldA
Cultures were done in 5-L bioreactors with a working volume of 3 L and air flow rate of 2 L/min with culture pH 6.5 and agitation rate of 600 rpm.

Figure 1

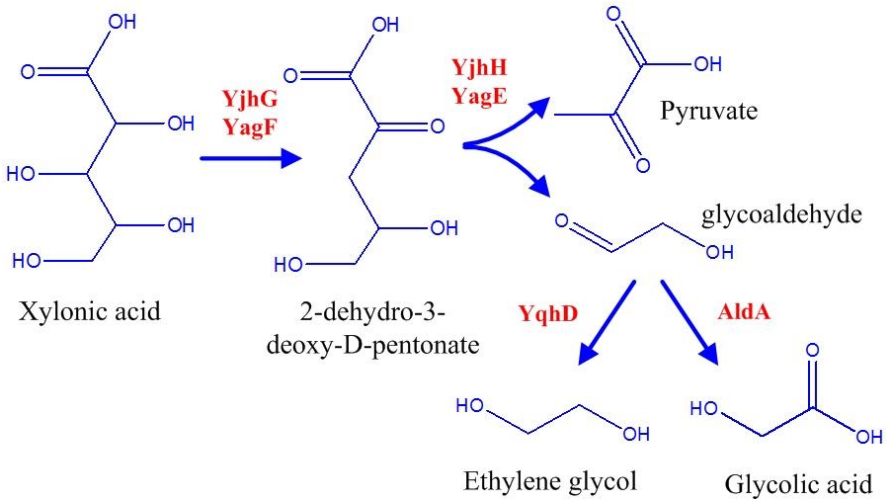

Figure 2

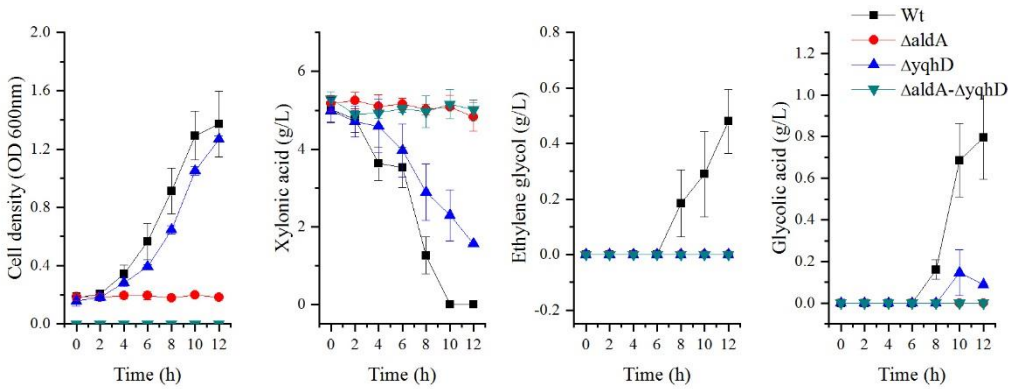

Figure 3

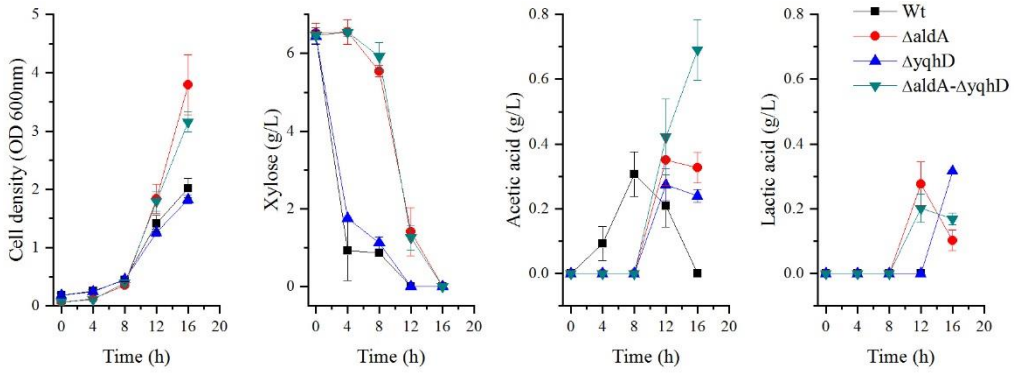

Figure 4

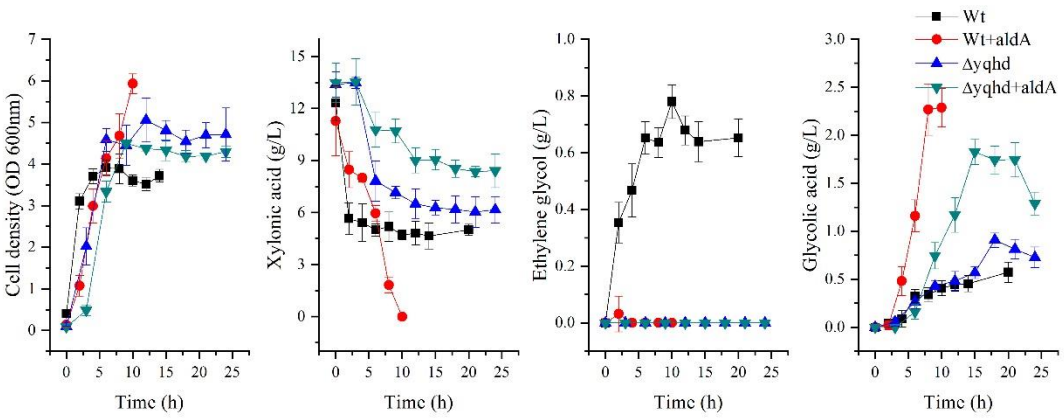

Figure 5

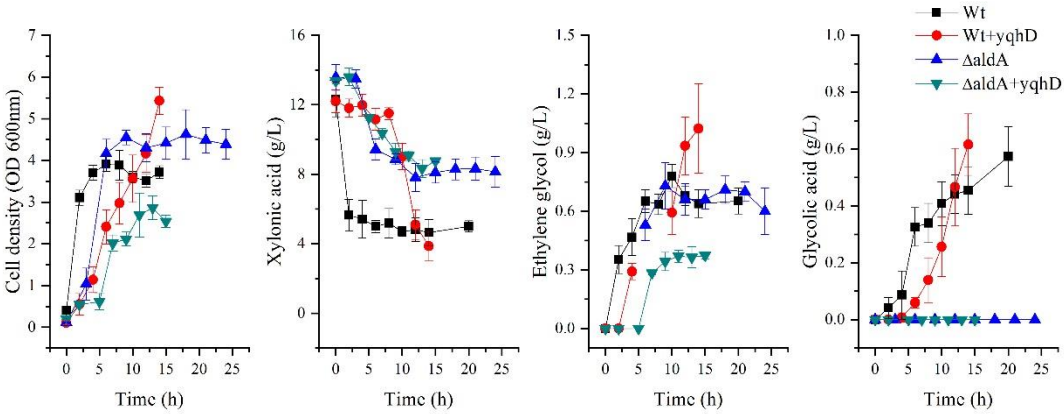

Figure 6

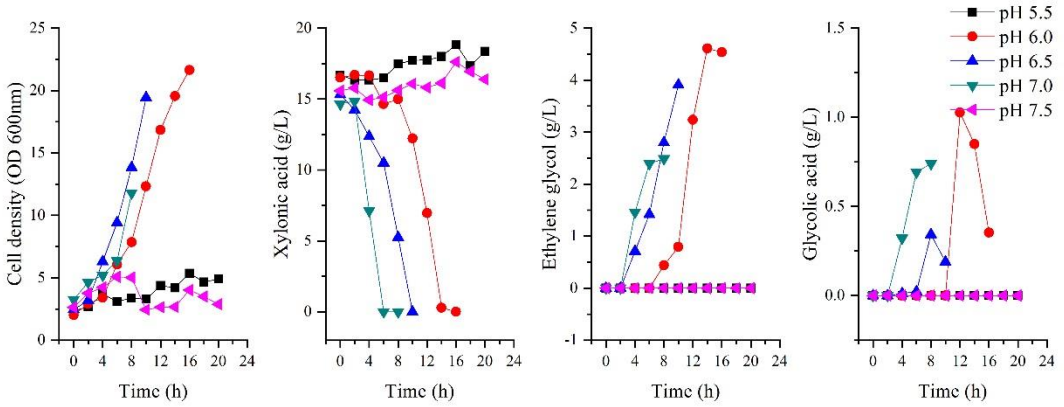

Figure 7

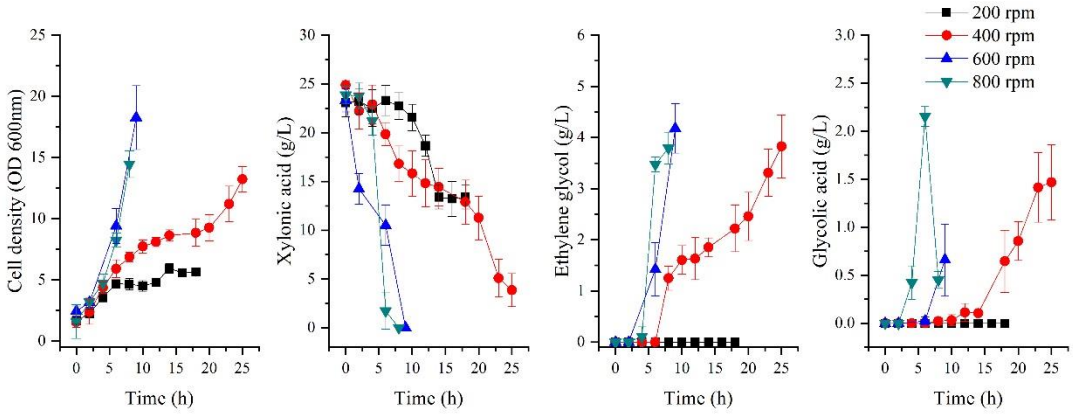

Figure 8