Reexamination of the physiological role of PykA in *Escherichia coli* revealed that it negatively regulates the intracellular ATP levels under anaerobic conditions.

**Running title: Physiological role of PykA in *E. coli***

Chunhua Zhao,\(^a\)^b Zhao Lin,\(^a\)^b Hongjun Dong,\(^a\) Yanping Zhang,\(^a\) Yin Li\(^a\)

CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, No. 1 West Beichen Road Chaoyang District, Beijing 100101, China; University of Chinese Academy of Sciences, Beijing 100049, China

Address correspondence to Yin Li, yli@im.ac.cn, or Hongjun Dong, redarmy305@gmail.com.
ABSTRACT

Pyruvate kinase is one of the three rate-limiting glycolytic enzymes that catalyzes the last step of glycolysis, conversion of phosphoenolpyruvate (PEP) into pyruvate which is associated with ATP generation. Two isozymes of pyruvate kinase PykF and PykA are identified in Escherichia coli. The PykF is considered important, whereas the PykA has less-defined role. Prior studies inactivated the pykA gene to increase the level of its substrate, PEP, and thereby increased the yield of end-products derived from PEP. We were surprised when we found a pykA::Tn5 mutant in a screen for increased yield of an end-product derived from pyruvate (n-butanol), suggesting the role of PykA needs to be reexamined. We show that the pykA mutant exhibited elevated intracellular ATP levels, increased biomass concentration, glucose consumption, and n-butanol production. We also discovered that the pykA mutant expresses higher levels of a presumed pyruvate transporter, YhjX, permitting the mutant to recapture and metabolize excreted pyruvate. Furthermore, we demonstrated the nucleotide diphosphate kinase activity of PykA leads to negative regulation on the intracellular ATP levels. Taken together, we propose that inactivation of pykA can be considered as a general strategy to enhance the production of pyruvate-derived metabolites under anaerobic conditions.
This study discovered that knocking out pykA significantly increased the intracellular ATP level and thus significantly increased the glucose consumption, biomass formation, and pyruvate-derived product formation under anaerobic conditions. pykA was considered encoding a dispensable pyruvate kinase, here we show that pykA negatively regulates the anaerobic glycolysis rate through regulating the energy distribution. Thus, knocking out pykA can be used as a general strategy to increase the productivity of pyruvate-derived fermentative products.
INTRODUCTION

Industrial fermentation aims to produce valuable products from cheap feedstocks by utilizing the diverse functions of microbes. Products such as antibiotics, amino acids, and vitamins, are mostly produced through aerobic fermentations, while products such as alcohols (ethanol, butanol, butanediol) are mainly produced through anaerobic fermentation (1-3). Organic acids such as lactic acid and succinic acid are also produced through anaerobic fermentation because higher yields could be obtained (4, 5).

During anaerobic fermentation, the reducing power will be mostly directed to product synthesis rather than being oxidized, resulting in a higher yield of target products. In addition, the low energy level due to absence of oxidative phosphorylation often leads to a higher glycolysis rate, resulting in a higher productivity (6). However, as the available ATP in anaerobic fermentation can only be generated from substrate-level phosphorylation, biomass concentration in anaerobic fermentation is usually much lower than that in aerobic fermentation (2, 7). Although a lower energy charge in anaerobic fermentation is beneficial for increasing the glycolysis rate (8-10), it has been shown that increasing ATP concentration can improve protein synthesis and increase biomass flux during anaerobic fermentation (11, 12). This suggests that factors affecting the intracellular ATP level may well be the targets for engineering so as to increase the efficiency of anaerobic fermentation processes.

There are three rate-limiting enzymes in the glycolysis, i.e. hexokinase, phosphofructokinase, and pyruvate kinase, among which only pyruvate kinase is associated with ATP generation (13). Pyruvate kinase catalyzes the last step of glycolysis, converting PEP into pyruvate. This step can
also be catalyzed by the glucose-specific transporter PtsG of the phosphotransferase system, but no free ATP will be generated (14). In *Escherichia coli*, there are two pyruvate kinase isoenzymes, PykF and PykA, encoded by *pykF* and *pykA* genes, respectively. PykF is characterized as an enzyme that is allosterically activated by fructose-1,6-bisphosphatase (FBP), whereas PykA is activated by AMP (15). In activity assays, the activity of PykF significantly surpasses that of PykA under aerobic conditions in the same reaction system using purified enzymes or crude extracts (15-17). It is generally believed that PykF contributes a greater extent to the activity of pyruvate kinase, while PykA contributes little (18).

However, in our previous study, inactivation of *pykA* in an engineered *n*-butanol producing *E. coli* strain was found to increase the glucose consumption rate and *n*-butanol production rate under anaerobic conditions (17). This surprising observation suggests that the physiological roles of PykA in glycolysis have been overlooked and should be reexamined. We therefore designed a series of experiments to better understand the physiological role of PykA in anaerobic fermentation.

**RESULTS**

**Construction of the basic *n*-butanol producing *E. coli* strain**

Genes involved in *n*-butanol synthetic pathway (*atoB*, *hbd*, *crt*, *ter*, *adhE2*, and *fdh*) were chosen according to the published research (2, 19). For construction of a genetically stable strain for *n*-butanol production that may have industrial application potential, we managed to integrate these genes into the chromosome of *E. coli*, which is different from all previous reported work. Six genes under control of the miniPtac promoter (20) were successfully integrated into wild type...
BW25113 to generate strain EB205, while genes involved in pathways for production of native fermentative byproducts (ethanol, lactate, acetate, and succinate) were disrupted simultaneously. In order to remove the native formate lysis pathway and improve the NADH supply, the hyc operon and hyp operon were deleted and a codon-optimized $fdh$ gene (specific sequence in supplemental material) were introduced in strain EB205, resulting in strain EB215. The mdh gene involved in fumarate pathway competing for NADH was further deleted to generate strain EB216. Thus, an $n$-butanol producing strain EB216, with all heterologously introduced genes integrated into the chromosome, was successfully constructed and used as a starting strain for this study. Tube fermentation results of strain EB205, EB215, and EB216 were shown in Table 2.

Finding new targets for improving $n$-butanol production using Tn5 transposon mutation

Tn5 transposon was introduced into strain EB216 with the aim to discover new gene targets that may contribute to improved $n$-butanol producers. 1196 mutants were screened in the first round among which 200 well-performed mutants were selected for the second round high performance liquid chromatography (HPLC) screening. After repeated tube fermentations, 3 mutants among 16 mutants screened in the second round showed significantly increased $n$-butanol production. The $n$-butanol titer of mutants EB216-Tn5-17#, EB216-Tn5-304#, and EB216-Tn5-599# increased 29%, 49%, and 57%, respectively (data not shown, (17)). Strain EB216-Tn5-599# was the most interesting mutant as next-generation sequencing (NGS) indicated the Tn5 transposon inserted into a gene $pykA$ (17) which is involved in glycolysis.

Inactivation of $pykA$ gene improved utilization of glucose and $n$-butanol production
Strain EB216-Tn5-599# described above produced 2.5 g/L n-butanol, which is 1.8-fold that of strain EB216 in tube fermentations (Table 3). To exclude any potential bias possibility of the introduced Tn5 transposon, the pykA open reading frame in strain EB216 was deleted, generating strain EB216-ΔpykA (also named as EB222). Table 3 shows that strain EB216-Tn5-599# and EB216-ΔpykA exhibited similar ability for glucose consumption and n-butanol production (P>0.05), which increased by over 30% and 70%, respectively, compared with that of strain EB216. This suggests that the pykA clean deletion mutant behaves similarly with the Tn5 insertion mutant, demonstrating that inactivation of pykA indeed improved glucose consumption and n-butanol production.

**Fermentation performance of the pykA inactivated mutant**

To better understand the effect of pykA inactivation on cellular metabolism, fermentation profiles of strain EB216 and EB216-ΔpykA were monitored. As shown in Fig. 1, cell growth (Fig. 1A), glucose consumption (Fig. 1B) and n-butanol production (Fig. 1C) of strain EB216-ΔpykA all significantly improved compared with that of the control strain. In addition, acetate and butyrate titers also increased as more glucose was consumed in strain EB216-ΔpykA. Among the profiles of other metabolites, the most significant difference exists in the production of pyruvate. Strain EB216 produced 1.36 g/L pyruvate at 12 h and the pyruvate concentration did not change significantly afterwards. Whereas, strain EB216-ΔpykA only produced 0.71 g/L of pyruvate at 12 h, followed by a decline of pyruvate concentration, suggesting reassimilation (Fig. 1I).

Glucose distribution analysis showed that the significantly increased (dark grey) glucose-to-n-butanol (68%) in strain EB216-ΔpykA can be ascribed to the significantly decreased (light grey) glucose-to-pyruvate (3%) (Table 4).
**yhjX gene is involved in pyruvate reassimilation**

One of the important characteristics of EB216-ΔpykA strain is its reutilization of the secreted pyruvate. Interestingly, from the transcriptomic analysis data (Table S1), we found the *yhjX* gene, which encodes a putative pyruvate transporter (21-23), was upregulated for 15-fold compared with the control strain EB216 at 48 h. Fig. 2A showed the *yhjX* expression strength is highly associated with the fermentation process (from 216 to 5703 of Fragment Per Kilo bases per Million reads (FPKM) value, which was 10−15 times higher than that in strain EB216 during the *n*-butanol producing period). To investigate whether *yhjX* is associated with pyruvate reassimilation, the *yhjX* gene was knocked out in the EB216 and EB216-ΔpykA strain, respectively. The results showed that the two *yhjX* mutant strains produced more pyruvate and less *n*-butanol (Fig. 2B). Hence, *yhjX* is an important target for further metabolic engineering.

**Reexamination and comparison of the roles of *pykA* and *pykF***

Fermentation and transcriptomic analyses showed that strain EB216-ΔpykA is a more active and robust strain (Fig. 1, Fig. S1 and Table S1). To further understand the role of *pykA* and its isoenzyme gene *pykF* under the circumstance of *n*-butanol production, we constructed a *pykF* mutant strain EB216-ΔpykF. Strain EB216-ΔpykA and EB216-ΔpykF were subjected to a series of biochemical analysis, with the aim to gain more insights into the physiological role of PykA and PykF in an *n*-butanol producing strain.

Pyruvate kinase assays of *E. coli* EB216 strain (Table 5) showed that the activity of PykF accounted for 93% of total pyruvate kinase activity, while that of PykA only accounted for 7%,
under aerobic conditions. This is consistent with previous recognition that PykF is the main pyruvate kinase in *E. coli* (24, 25). However, under anaerobic conditions, the activity of PykF accounted for 70% of total pyruvate kinase activity, while that of PykA accounted for 30%. This indicates that PykA plays a more important role under anaerobic conditions than under aerobic conditions. Moreover, it was observed that knocking out *pykF* would not affect the absolute activity of PykA, and vice versa, under both aerobic and anaerobic conditions. This suggests that the activity of PykF and PykA is independent of their counterparts.

We also noticed that literatures reported that PykA is capable of catalyzing four kinds of NDPs (ADP, UDP, GDP, CDP) to NTPs in the reaction of PEP to pyruvate, while PykF typically use ADP as substrate to form ATP (26-29). To study the selectivity of NDP by PykF or PykA, we did the pyruvate kinase assay and determined the kinetic parameters by replacing ADP with UDP, GDP, or CDP using the purified enzymes. The specificity constant (*k*$_{cat}$/*K*$_{m}$) of PykA and PykF on four NDPs was determined. Results showed that PykA is broadly active on four NDPs, while PykF is only active on ADP and UDP (Table 6). Strikingly, the *k*$_{cat}$/*K*$_{m}$ of PykA on ADP and UDP is about one order of magnitude higher than that of PykF, while the *k*$_{cat}$/*K*$_{m}$ of PykA on GDP is at the same magnitude but still higher than that of PykF on ADP and UDP. CDP is the last substrate choice according to the results. We therefore further analyzed the intracellular ATP concentration of strains EB216, EB216-ΔpykF and EB216-ΔpykA. The results showed that the intracellular ATP concentration of strain EB216-ΔpykA increased by 150% compared with strain EB216 (Table 7). This suggests the phosphorus groups that were used for generation of GTP and CTP by PykA might be redirected for generation of ATP and UTP by PykF, resulting in an
increased intracellular concentration of ATP and enabled strain EB216-ΔpykA grow to a higher cell density (Fig. 1A).

In addition, the NADH which is the reducing power for n-butanol formation was also determined. NADH assay showed that the intracellular NADH level of strain EB216-ΔpykF did not significantly differ from that of strain EB216 (P>0.05), while the intracellular NADH level of strain EB216-ΔpykA decreased by 21% (P<0.05) compared with that of strain EB216 (Table 7). This suggests an increased turn-over rate of NADH in strain EB216-ΔpykA, which indicates the NADH consuming n-butanol pathway in EB216-ΔpykA strain is more active.

DISCUSSION

Previously, knockout of pykA was used to weaken pyruvate formation from PEP for enhancing production of PEP-derived metabolites. One example was the production of 3-Deoxy-D-arabino-heptulosonic acid 7-phosphate (DAHP) from PEP. Knockout of pykF led to a 19% improved DAHP titer, while knockout of pykA led to a 46% improved DAHP titer, and double knockout of pykA and pykF led to a 235% improved DAHP titer (30). Another example was the production of succinate from PEP. Knockout of pykF led to a 31% improved succinate titer, while double knockout of pykA and pykF led to an 188% improved succinate titer (31). In this study, deletion of pykA was found to be associated with the significantly prolonged growth phase, and increased cell growth and production of pyruvate-derived n-butanol (Fig. 1A−C) under anaerobic conditions. This suggests that the role of PykA in glycolysis was overlooked and needs reexamination.
In aerobic glycolysis pathway of *E. coli*, the PykF is believed to be the major contributor of pyruvate kinase activity, while PykA contributes little (24, 25). Our enzymatic assay confirmed that PykA does contribute very little to total pyruvate kinase activity (7%) (Table 5). But it is interesting to notice that PykA contributes to 30% of total pyruvate kinase activity (Table 5) under anaerobic conditions. Previous studies also showed that the transcription of *pykA* gene under anaerobic condition could be improved up to 3-fold that of under aerobic condition, and knockout of *fnr* gene (encoding FNR regulator which represses genes involved in aerobic respiration and activates genes required for anaerobic respiration) decreased the *pykA* transcription by 63% (32, 33). Based on these transcriptional and our enzymatic analysis, we believe that *pykA* should not be considered as a dispensable pyruvate kinase gene, but plays an important role especially under anaerobic conditions.

In association with its activity of converting PEP into pyruvate, PykA was also reported to have a broad specificity on nucleotide diphosphates, which could serve as acceptors of energy-rich phosphate group from PEP, and could take place of nucleoside diphosphate kinase in supplying nucleoside triphosphates (26, 27). Determination of the specificity constant $k_{cat}/K_m$ allowed us to systematically understand the difference in catalytic efficiency of PykA and PykF on four NDPs. PykA is mostly and equally active on ADP and UDP, followed by GDP and CDP. Waygood et al (29) only reported the $K_m$ of PykA on four NDPs (ADP, 0.08 mM; GDP, 0.13 mM; UDP, 0.25 mM; CDP 0.39 mM). While the $K_m$ of PykA on ADP is consistent with our determination, the $K_m$ on the other three NDPs are quite different from our results. Interestingly, the PykF is not active on GDP and CDP in our test, whereas the PykF documented in an early literature showed a very high affinity to GDP ($K_m$ 0.05 mM) and very low affinity to CDP ($K_m$ 6.7 mM) (15).
K_m of PykF on ADP and UDP determined in our study is quite consistent with the literature though (15). As the catalytic efficiency of PykA on GDP and CDP is comparable with the catalytic efficiency of PykF on ADP and UDP (Table 6), it is therefore conceivable that the phosphate group originally used for generation of GDP and CDP by PykA might be redirected for generation of ADP and UDP by PykF, when pykA is deleted. However, we should also note that the in vitro conditions used for determining the kinetic parameters of PykA and PykF would not be the same as the actual conditions inside the cells - a comprehensive quantitative understanding why intracellular ATP concentration in strain EB216-ΔpykA increased by 150% need further investigation.

It is intriguing that our results (Fig. 1 and Table 4) showed that the deletion of pykA increased the conversion from pyruvate to n-butanol, but not from glucose to pyruvate. However, the n-butanol pathway from pyruvate to n-butanol does not contain reactions requiring ATP, thus the increased ATP level in pykA mutant cannot account for this phenomenon. We noticed the accumulation of pyruvate in strain EB216, suggesting the rate of converting glucose to pyruvate surpassed the rate of converting pyruvate into n-butanol. In another word, the glycolysis rate in strain EB216 is faster than the rate of the introduced n-butanol pathway. Notably, the pyruvate accumulation was slower in strain EB216-ΔpykA, and the peak value was lower, as compared to that of strain EB216. This suggests that with 30% decreased pyruvate kinase activity, the glycolysis rate in strain EB216-ΔpykA is reduced to a level that is more compatible with rate of the introduced n-butanol pathway, leading to a continuous carbon flux from glucose to n-butanol, resulting in an improved n-butanol production. Moreover, activation of the yhjX gene in strain EB216-ΔpykA
enabled reassimilation of pyruvate, which may explain the yield of n-butanol on glucose in strain EB216-ΔpykA also slightly improved.

In summary, the little pyruvate kinase activity of PykA under aerobic conditions let people think PykA is not important for glycolysis. Although previous in vivo studies have shown PykA has a broad specificity on nucleotide diphosphates, no literature has correlated this feature with the glycolysis ATP production. The retrieve of a PykA null mutant with stronger n-butanol producing ability encouraged us to reexamine the role of PykA in glycolysis. It turned out that PykA represents 30% of the pyruvate kinase activity under anaerobic conditions, and presumably the PykA is used for NTPs production in the cells. Knocking out PykA can lead to an increased ATP production while not obviously affecting glucose consumption, thus can be applied as a general strategy to improve production of pyruvate-derived metabolites.

MATERIALS AND METHODS

Bacterial strains, plasmids and primers

_E. coli_ BW25113 (16) was used for n-butanol producing strain construction as wild type. All primers were synthesized by Invitrogen (Beijing, China) followed by polyacrylamide gel electrophoresis purification. All strains and plasmids used in this study are listed in Table 1. Primers are listed in Table S2.

Construction of plasmids for gene integration

Plasmid pKmTsSacB (20) was used as a skeleton plasmid for constructing various recombinant plasmids. A constitutive tac promoter (miniPtac) (20, 34) which was introduced in the forward
primer was used for gene expression. Recombinant plasmid pKmTsSacB-yqhD::atoB was constructed for integration of the atoB gene into the genome of *E. coli* BW25113 at its yqhD site (ΔyqhD::atoB). Firstly, the atoB gene was amplified from the genomic DNA of BW25113 using primers yqhD::atoB-3, yqhD::atoB, and yqhD::atoB-4. Secondly, the yqhD upstream and downstream fragments were amplified from genomic DNA of BW25113 using primer pairs yqhD::atoB-1/yqhD::atoB-2 and yqhD::atoB-5/yqhD::atoB-6, respectively. Thirdly, the amplified atoB gene and the above two fragments were spliced through BamHI and Xbal sites to form the ‘yqhDup-atoB-yqhDdown’ cassette *in vitro*. Finally, this ‘yqhDup-atoB-yqhDdown’ cassette was inserted into the XhoI and PstI sites of plasmid pKmTsSacB to generate plasmid pKmTsSacB-yqhD::atoB. In a similar manner, plasmids pKmTsSacB-ldhA::hbd (ΔldhA::hbd, *hbd* from *Clostridium acetobutylicum*), pKmTsSacB-ackA-pta::crt (Δ(ackA-pta)::crt, *crt* from *C. acetobutylicum*), pKmTsSacB-adhE::ter (ΔadhE::ter, ter from *Treponema denticola*), pKmTsSacB-frdBC::adhE2 (ΔfrdBC::adhE2, *adhE2* from *C. acetobutylicum*), and pKmTsSacB-eutE::fdh (ΔeutE::fdh, *fdh* from *Candida boidinii*) were constructed using their respective primers listed in Table S2. PCR enzymes, restriction enzymes, and T4 DNA ligase were purchased from New England Biolabs (New England Biolabs, Beijing, China).

**Strain development**

Six genes involved in production of *n*-butanol were integrated into the BW25113 chromosome at their respective chosen sites, one after another. Briefly, the pKmTsSacB-based recombinant plasmid was transformed into *E. coli* BW25113 by a Gene pulser (Bio-Rad Laboratories, Inc., Richmond, CA) and plated on agar with kanamycin. Positive single-crossover integrants were separately streaked on agar plates supplemented with sucrose to screen for double-crossover
integration. The correct integrant verified by colony PCR and DNA sequencing was selected for integrating the next gene (20). Finally, a strain with six n-butanol pathway genes *atoB, hbd, crt, ter, adhE2, fdh* respectively integrated at the *yqhD, ldhA, ackA-pta, adhE, frdBC, eutE* sites, was obtained and designated as EB205. Subsequently, the *hyc* operon and *hyp* operon of strain EB205 were deleted following Red recombinase-based one-step inactivation of chromosomal genes technology (16), followed by integrating a modified *fdh* gene from *C. boidinii* which was codon-optimized for *E. coli* into the *fdhF* site under control of the native *fdhF* promoter, resulting in strain EB215. Furthermore, the NADH consumption gene *mdh* of strain EB215 was deleted to generate strain EB216. Red mediated recombination method (16) was used for further deletion of *pykA, pykF*, and *yhjX* genes. All strains developed in this study were listed in Table 1.

**Tn5 transposon based screening for new targets**

The genome-wide random mutagenesis of strain EB216 was carried out using the EZ-Tn5 <KAN-2> TNP Transposome Kit (Epicentre, Madison, WI), following the manufacturer’s instructions. From previous experiment (data not shown), it is noticed that high titer of n-butanol always associated with low extracellular concentration of pyruvate and high optical density. Therefore, mutants were screened by measuring the optical density (OD600) of the bacterial cells, the pyruvate concentration in the screening broth, and the ratio of pyruvate concentration over OD600 of each mutant. The mutant colonies were inoculated into deep 96-well plates (containing 1.25 mL fermentation medium per well) by a QPIX2 Robotic colony picker (Molecular Devices Co.), and incubated in a Shellab Bactron I-2 anaerobic chamber (Sheldon Manufacturing Inc.) filled with an anaerobic gas mixture (5% CO2, 5% H2 and 90% N2) for 3 days. For optical density measurements, 200 µL of culture was transferred to a standard 96-well
plate and the OD600 was determined. For pyruvate concentration measurements, 80 µL 25-fold diluted culture, 80 µL 0.0375% 2,4-dinitrophenylhydrazine (DNPH), 80 µL 2.81 mol/L NaOH were mixed and the absorbance at 520 nm was determined after 10-minute reaction (35). All measurements were performed by using a laboratory automation workstation Biomek 3000 (Beckman Coulter Inc.). A TECAN infinite M200 plate reader (Tecan Austria GmbH, Grödig, Austria) was used for absorbance detection. Mutants with low ratio of pyruvate concentration/OD600 were selected for the second round screening.

The second round screening was carried out using HPLC analysis of n-butanol after tube fermentation. The best n-butanol producing mutants were collected to extract genomic DNA using an E.Z.N.A Bacterial DNA Isolation Kit (Omega Biotek Inc.) for the NGS which was completed and analyzed by GENEWIZ, Inc. (GENEWIZ, Suzhou, China).

**Growth and maintenance conditions**

*E. coli* strains were grown aerobically at 30°C or 37°C in Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) supplemented, when necessary, with ampicillin (100 µg/mL), chloramphenicol (30 µg/mL), or kanamycin (50 µg/mL). Strains were maintained frozen in 15% glycerol at −80°C. Before culturing, fresh colonies were picked from LB plates and inoculated into LB medium at 37°C and 200 rpm for 12 h. Unless otherwise indicated, 5% of this culture was inoculated into the fresh medium. For batch culture, M9 mineral salt (17.1 g Na₂HPO₄·12H₂O, 3 g KH₂PO₄, 0.5 g NaCl, 1 g NH₄Cl, 2 mmol MgSO₄, and 0.1 mmol CaCl₂ per liter water) supplemented with 2 g/L yeast extract (M9Y) (36) and 20 g/L glucose was used. Tube fermentation was performed in 10 mL medium in a sealed 15 mL polypropylene conical
tube (BD Biosciences, San Jose, CA) in a 37°C incubator without shaking for 48 h or longer when necessary. Bioreactor fermentation was carried out in a 7.5 L BioFlo 110 bioreactor with 4 L medium (New Brunswick Scientific, Edison, NJ) for 72 h. Cells were grown at 37°C and 200 rpm without pH control.

Transcriptome sequencing

Cell cultures growing in fermentor for 6 h, 12 h, 24 h and 48 h were harvested by centrifugation and frozen at −80°C. RNA isolation, cDNA libraries construction and sequencing was performed by GENEWIZ, Inc. (GENEWIZ, Suzhou, China). Briefly, total RNAs were isolated using TRIzol (37, 38). For cDNA libraries construction, Illumina TruSeq RNA Sample Pre Kit (Illumina, San Diego, CA) was used. After purification, detection, and quantification, the libraries were sequenced by HiSeq 2000 system (Illumina, San Diego, CA).

Transcriptomic analysis

Transcriptomic analysis was also carried out by GENEWIZ, Inc. Raw data in FASTQ format was processed by Trimmomatic (v0.30) (39) to obtain the clean data which was then evaluated using FastQC (v0.10.1 at http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The average quality score exceeded 30 which indicated a good sequencing quality. Subsequently, clean data reads were mapped onto the E. coli BW25113 genome (GenBank accession number CP009273.1) using Bowtie2 (40). Quantification of gene expression and analysis of gene differential expression were performed using FPKM based rsem software (v1.2.4) (41) and edgeR (v3.4.2) (Bioconductor), respectively. Differentially expressed genes showed more than 2-
fold change and a smaller false discovery rate (FDR) (<=0.05) were defined as the significantly differentially expressed genes.

**Analytical procedure**

Cell density in tube or fermentor was determined by a UV-visible spectrophotometer (UV-2802PC; Unico, Shanghai, China) with the optical density at 600 nm (OD600). Sugar, organic acid, and alcohol concentrations of fermentation samples were analyzed by an Agilent 1260 HPLC system (Agilent Technologies, Santa Clara, CA). A Bio-Rad HPX-87H column (Bio-Rad Laboratories, Inc., Richmond, CA) with 5 mM H₂SO₄ as the mobile phase (10 μL injection, 0.5 mL/min, 15°C) was used for HPLC analysis.

Cells growing at the mid-exponential growth phase were harvested to prepare the crude extract by sonication (Hongdaxinchen biotechnology Co., Ltd., Beijing, China) for enzyme and intracellular metabolites assay. Specifically, pyruvate kinase assay was implemented based on the previous studies (15, 42, 43). The activities of PykF and PykA were detected with the addition of their respective activator FBP and AMP. 1 mM FBP could bring PykF to 80% of its maximal velocity without affecting the activity of PykA; while 1 mM AMP would bring PykA to its maximal velocity without affecting the activity of PykF (43). A classical lactate dehydrogenase coupled reaction was used to detect the decrease in absorbance of NADH at 340 nm using a microtiter plate reader. The standard assay mixture contains (in a final volume of 200 μL): 30 mmol/L HEPES, pH 7.5; 10 mmol/L MgCl₂; 10 mmol/L KCl; 2 mmol/L ADP; 1 mmol/L PEP; 0.15 mmol/L NADH; 20 units of bacterial lactate dehydrogenase and 1 mmol/L FBP or 1 mmol/L AMP for the activation of PykF or PykA. The reaction was started by the
addition of the crude extract and carried out at 25°C. A molar extinction coefficient for NADH of 6.22×10^3 L/mol·cm was used for converting changes in absorbance into moles of pyruvate formed per minute.

To further determine the kinetic parameters of the two pyruvate kinases, UDP, GDP, and CDP were used as substitutes for ADP. Pyruvate kinases expressed on pET30a vector in *E. coli* BL21(DE3) were purified using Immobilized Metal Affinity Chromatography (IMAC) (44). The intracellular concentration of pyruvate, ATP, and NADH were immediately determined after sampling using pyruvate assay kit (BioAssay Systems, Hayward, CA), ATP fluorometric assay kit (BioVision Incorporated, Milpitas, CA), and NADH fluorometric assay kit (BioVision Incorporated, Milpitas, CA) according to the manufacturer’s instructions, respectively.

Measurement was performed after sampling and disrupting the cells by sonication at 140 W for 1 min on ice (45, 46). The protein concentration in the cell free extracts was quantified by using a coomassie brilliant blue protein assay kit (Comin biotechnology Co., Ltd., Suzhou, China).

TECAN infinite M200 plate reader was used to detect the absorbance at different wavelengths.

All the mentioned reagents were purchased from Sigma (Sigma-Aldrich Co., Shanghai, China) or Yuanye (Yuanye biotek, Shanghai, China).

**Statistical analysis**

The statistical analysis of data and plots was performed using Student’s *t* test in SPSS software when necessary. *P* values of <0.05 were considered to indicate statistical significance.

**Accession number(s)**
High-throughput sequence data (RNA-seq) of EB216 and EB216-∆pykA at different fermentation time have been deposited into the NCBI GEO database under the accession number GSE96551.

**ACKNOWLEDGMENTS**

We are grateful to Huawei Zhu from our lab and Simi Jasmin Zhang (University of Waterloo, Waterloo) for critical reading of the manuscript. We thank Hongwei Yu and Huailin Zhou (Institute of Botany, CAS, Beijing) for their help with statistical analysis. We also thank Dr. Guoxia Liu from our lab for the kind help of enzyme kinetic data analysis.

This study was funded by the National Natural Science Foundation of China (grant number 31270107), the National High Technology Research and Development Program of China (grant number 2011AA02A208), and the CAS/SAFEA International Partnership Program for Creative Research Teams.
REFERENCES


FIGURE LEGENDS

**FIG 1** Fermentation profiles of strain EB216 (square) and EB216-ΔpykA (triangle).

Fermentation was performed in 7.5 L bioreactors with an initial working volume of 4 L. Data shown represent the average of three independent batch culture and the error bars represent standard deviation. (A) Cell growth; (B) Glucose; (C) *n*-Butanol; (D) Formate; (E) Lactate; (F) Ethanol; (G) Acetate; (H) Butyrate; (I) Pyruvate.

**FIG 2** The role of *yhjX* gene in fermentative production of *n*-butanol. Three repeats were performed and the error bars represent standard deviation. (A) The expression strength of *yhjX* gene in strain EB216 and EB216-ΔpykA. The expression strength is showed as the values of FPKM calculated by RSEM software tool (V 1.2.4). (B) The pyruvate and *n*-butanol production of *yhjX*-deleted strains derived from EB216 and EB216-ΔpykA. The strains were grown in static 15 mL sealed BD tubes for *n*-butanol production at 37°C for 48 h.
## TABLE 1 Bacterial strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strains or plasmids</th>
<th>Relevant characteristics</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW25113</td>
<td>lacI, rnb^T_14, ΔlacZ_{W116}, hsdR514, ΔaraBAD_{AE133}, ΔβhaBAD_{LD78}</td>
<td>(16)</td>
</tr>
<tr>
<td>BL21(DE3)</td>
<td>F- ompT hsdSB (rB^-, mB^-) galCmrrne13I (DE3) Lab storage</td>
<td></td>
</tr>
<tr>
<td>EB205</td>
<td>Δ(ackA-pta):cer, ΔadhE::cer, ΔfrdBC::adhE2, ΔeutE::fdh</td>
<td>This work</td>
</tr>
<tr>
<td>EB215</td>
<td>Derived from EB205, Δ(bye-hyp)::FRT, ΔfilhF::fdh</td>
<td>This work</td>
</tr>
<tr>
<td>EB216</td>
<td>Derived from EB215, Δmdh::FRT</td>
<td>This work</td>
</tr>
<tr>
<td>EB216-Tn5-599#</td>
<td>Derived from EB216, pykA gene was inserted by Tn5 transposon</td>
<td>This work</td>
</tr>
<tr>
<td>EB216-ΔpykA</td>
<td>Derived from EB216, ΔpykA</td>
<td>This work</td>
</tr>
<tr>
<td>EB216-ΔpykF</td>
<td>Derived from EB216, ΔpykF</td>
<td>This work</td>
</tr>
<tr>
<td>EB216-ΔyhjX</td>
<td>Derived from EB216, ΔyhjX</td>
<td>This work</td>
</tr>
<tr>
<td>EB216-ΔpykA&amp;pykF</td>
<td>Derived from EB216-ΔpykA, ΔpykF</td>
<td>This work</td>
</tr>
<tr>
<td>EB216-ΔpykA&amp;yhjX</td>
<td>Derived from EB216-ΔpykA, ΔyhjX</td>
<td>This work</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pKmTsSacB</td>
<td>Skeleton plasmid for cloning, constructed from pK18mobsacB (ATCC) with thermo-sensitive ori of pKD46, kan</td>
<td>(20)</td>
</tr>
<tr>
<td>pKmTsSacB-yqhD::atoB</td>
<td>pKmTsSacB-based vector, containing yqhDup-atoB-yqhDdown cassette, kan</td>
<td>This work</td>
</tr>
<tr>
<td>pKmTsSacB-ldhA::hbd</td>
<td>pKmTsSacB-based vector, containing ldhAup-hbd-</td>
<td>This work</td>
</tr>
</tbody>
</table>
ldhA

down cassette, kan

pKmTsSacB-(ackA-pta)::crt

pKmTsSacB-based vector, containing (ackA-pta)up-crt-
(ackA-pta)down cassette, kan

This work

pKmTsSacB-adhE::ter

pKmTsSacB-based vector, containing adhEup-ter-
adhEdown cassette, kan

This work

pKmTsSacB-frdBC::adhE2

pKmTsSacB-based vector, containing frdBCup-adhE2-
frdBCdown cassette, kan

This work

pKmTsSacB-eutE::fdh

pKmTsSacB-based vector, containing eutEup-fdh-
eutEdown cassette, kan

This work

pKmTsSacB-fdhF::fdh

pKmTsSacB-based vector, containing fdhFup-fdh-
fdhFdown cassette, kan

This work

pKD4

bla, FRT, kan

(16)

pKD46

bla, araC, gam-bet-exo

(16)

pCP20

bla, flp, cat

(16)

pET30a

kan

Lab storage

pET30a-pykA

pET30a-based vector, containing his-tagged PykA gene,
kan

This work

pET30a-pykF

pET30a-based vector, containing his-tagged PykF gene,
kan

This work

bla, ampicillin resistance gene; kan, kanamycin resistance gene; gam-bet-exo, red recombinase genes; flp, flipase

(Flp recombinase) gene; cat, chloramphenicol resistance gene; FRT, flipase recognition target.
TABLE 2 The n-butanol production of the basic strains

<table>
<thead>
<tr>
<th>Strains</th>
<th>Glucose consumed (g/L)</th>
<th>n-Butanol (g/L)</th>
<th>Yield (g/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB205</td>
<td>4.83±0.22</td>
<td>0.47±0.03</td>
<td>9.7%</td>
</tr>
<tr>
<td>EB215</td>
<td>13.45±0.04</td>
<td>0.87±0.03</td>
<td>13.2%</td>
</tr>
<tr>
<td>EB216</td>
<td>11.12±0.18</td>
<td>1.60±0.12</td>
<td>18.0%</td>
</tr>
</tbody>
</table>

The strains were grown in static 15 mL sealed BD tubes for n-butanol production at 37°C for 72 h. Three repeats were performed and the errors represent standard deviation.
The effect of pykA inactivation on utilizing glucose for n-butanol production

<table>
<thead>
<tr>
<th>Strains</th>
<th>OD600</th>
<th>Glucose consumed (g/L)</th>
<th>Products (g/L)</th>
<th>Yield (g/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB216</td>
<td>1.44±0.02</td>
<td>8.09±0.40</td>
<td>1.28±0.04</td>
<td>1.38±0.16</td>
</tr>
<tr>
<td>EB216-Tn5-599#</td>
<td>1.83±0.04</td>
<td>10.88±0.21</td>
<td>0.59±0.07</td>
<td>2.52±0.18</td>
</tr>
<tr>
<td>EB216-ΔpykA</td>
<td>1.74±0.02</td>
<td>10.49±0.15</td>
<td>0.43±0.09</td>
<td>2.37±0.10</td>
</tr>
</tbody>
</table>

The strains were grown in static 15 mL sealed BD tubes for n-butanol production at 37°C for 48 h. EB216 is the control strain. The pykA gene was inactivated by insertion of Tn5 transposon in strain EB216-Tn5-599#, or deletion in strain EB216-ΔpykA. Three repeats were performed and the errors represent standard deviation.
TABLE 4 Comparison of end products titers and glucose distribution between strain EB216 and EB216-ΔpykA

<table>
<thead>
<tr>
<th>Strains</th>
<th>Biomass</th>
<th>Pyruvate</th>
<th>Lactate</th>
<th>Acetate</th>
<th>Ethanol</th>
<th>Butyrate</th>
<th>n-Butanol</th>
<th>Gluc. cons.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB216</td>
<td>mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.42</td>
<td>15.45</td>
<td>8.64</td>
<td>1.38</td>
<td>18.53</td>
<td>2.06</td>
<td>22.84</td>
<td>46.39</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>6%</td>
<td>16%</td>
<td>9%</td>
<td>1%</td>
<td>20%</td>
<td>4%</td>
<td>49%</td>
</tr>
<tr>
<td>EB216-ΔpykA</td>
<td>mM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.83</td>
<td>4.43</td>
<td>9.49</td>
<td>2.01</td>
<td>27.46</td>
<td>4.48</td>
<td>55.54</td>
<td>78.50</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>5%</td>
<td>3%</td>
<td>6%</td>
<td>1%</td>
<td>17%</td>
<td>5%</td>
<td>68%</td>
</tr>
</tbody>
</table>

The calculation is based on the data collected at the final point (72 h) in Fig. 1.

*a* The value indicates the glucose distribution for specific end product, which is the value of the theoretical amount of glucose consumed for one end product divided by the practical amount of total consumed glucose.

*b* The biomass was calculated from OD600 using an equation that biomass (g/L) = OD600×0.25. The molecular formula for *E. coli* biomass is CH₁.₇7O₀.₄9N₀.₂₄ (47).

*c* The ‘gluc. cons.’ indicates total glucose consumed in 72 hours.

*d* The value indicates the carbon balance, which is calculated as the theoretically used glucose for all the detected end products divided by the practically used glucose. The reason that the total value exceeds 100% is due to the presence of yeast extract.

Light grey shading indicates decreased glucose distribution. Dark grey shading indicates increased glucose distribution.
<table>
<thead>
<tr>
<th>Strains</th>
<th>Anaerobic culture</th>
<th>Aerobic culture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PykF (U/mg)</td>
<td>PykA (U/mg)</td>
</tr>
<tr>
<td>EB216</td>
<td>0.72±0.09</td>
<td>0.31±0.03</td>
</tr>
<tr>
<td>EB216-ΔpykF</td>
<td>n.d.</td>
<td>0.35±0.05</td>
</tr>
<tr>
<td>EB216-ΔpykA</td>
<td>0.76±0.09</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

PykF: 1 unit of pyruvate kinase will generate 1 μmole pyruvate and 1 μmole ATP from PEP and ADP per minute under test conditions (25°C, pH 7.5). The mg refers to the mg crude protein extract. n.d. indicates ‘not detectable’. For anaerobic culture, the strains were grown in static 15 mL sealed BD tubes at 37°C, while for aerobic culture, the strains were grown in 50 mL half-sealed BD tubes in a 220 rpm shaker at 37°C. The working volume were both 10 mL. Three repeats were performed and the errors represent standard deviation.
# TABLE 6 The kinetic parameters of the purified his-tagged PykA and PykF

<table>
<thead>
<tr>
<th>Substrates</th>
<th>PykA</th>
<th>PykF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (mM)</td>
<td>$k_{cat}$ (s$^{-1}$)</td>
</tr>
<tr>
<td>ADP</td>
<td>0.060±0.009</td>
<td>129±14</td>
</tr>
<tr>
<td>UDP</td>
<td>0.089±0.019</td>
<td>164±20</td>
</tr>
<tr>
<td>GDP</td>
<td>0.345±0.036</td>
<td>209±17</td>
</tr>
<tr>
<td>CDP</td>
<td>0.560±0.028</td>
<td>43±2</td>
</tr>
</tbody>
</table>

n.d. indicates “not detectable”. *E. coli* BL21(DE3) and pET30a vector were used to express the his-tagged pyruvate kinases. *pykA* gene was amplified from genomic DNA of BW25113 with primers pykA-NdeI-1 and pykAChis-XhoI-2 (Table S2) and cloned into pET30a vector through NdeI and XhoI sites while *pykF* gene was amplified from genomic DNA of BW25113 with primers pykF-NdeI-1 and pykFChis-HindIII-2 (Table S2) and cloned into pET30a vector through NdeI and HindIII sites. Michaelis-Menten equation was used to calculate the kinetic parameters. Three repeats were performed and the errors represent standard deviation.
<table>
<thead>
<tr>
<th>Strains</th>
<th>Intracellular pyruvate level (μmol/mg protein)</th>
<th>Intracellular ATP level (nmol/mg protein)</th>
<th>Intracellular NADH level (pmol/mg protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB216</td>
<td>0.33±0.04</td>
<td>3.37±0.48</td>
<td>3690±355</td>
</tr>
<tr>
<td>EB216-ΔpykF</td>
<td>0.31±0.02</td>
<td>2.33±0.22</td>
<td>3203±277</td>
</tr>
<tr>
<td>EB216-ΔpykA</td>
<td>0.39±0.01</td>
<td>8.43±0.47</td>
<td>2903±193</td>
</tr>
</tbody>
</table>

The mg protein refers to the mg crude protein extract. The strains were grown in static 15 mL sealed BD tubes at 37°C. Three repeats were performed and the errors represent standard deviation.