Pseudogene YdfW in *Escherichia coli* decreases hydrogen production through nitrate respiration pathways

Marahaini Mokhtar [a], Mohd Zulkhairi Mohd Yusoff [a,b,],
Mohd Shukuri Mohamad Ali [c], Nurul Asyifah Mustapha [d],
Thomas K. Wood [e], Toshinari Maeda [d]*

[a] *Department of Bioprocess Technology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[b] *Laboratory of Biopolymer and Derivatives, Institute of Tropical Forestry and Forest Products (INTROP), Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[c] *Enzyme and Microbial Technology Research Center (EMTech), Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[d] *Department of Biological Functions Engineering, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan*
[e] *Department of Chemical Engineering, Pennsylvania State University, 161 Fenske Laboratory, University Park, PA 16802, USA*

ARTICLE INFO

ABSTRACT

Escherichia coli has approximately 4300 open reading frames and about 178 of them are annotated as pseudogenes. The existence of the pseudogenes in *E. coli* had raised the question of whether they contribute to any cell function. Recently, several pseudogenes have been studied; however, these genes are still considered rare and tend to be ignored as so-called junk. Our recent study has shown several pseudogenes in *E. coli* are essential in hydrogen metabolism. *ydfW* is one of the pseudogenes associated with hydrogen metabolism. Here, we found carbon dioxide and nitrate in the fermentation medium of mutant *ydfW* suggesting that this strain may induce nitrate pathways. The Fdh-N complex is comprised of FdnG, FdnH, and FdnI subunits which correspond to the nitrate-nitrite pathway. This finding is supported by transcription analysis that shows formate dehydrogenase (*fdnG*) and nitrate reductase (*narG*) are significantly induced (25- and 18-fold, respectively). FdnG and NarG play a significant role in the reduction of nitrate into nitrite; hence, their genes are induced by the build-up of nitrate. Based on *in silico* docking analysis, we predict that YdfW is covalently bound to the Fdh-N complex in the cytoplasm. Hence, deletion of *ydfW* is expected has influenced the nitrate-nitrite pathway. Subsequently, the situation will affect the hydrogen production in *E. coli*.

© 2019 Hydrogen Energy Publications LLC. Published by Elsevier Ltd. All rights reserved.

List of abbreviation

ERRAT	The computational analysis to evaluate the predicted 3D model
Fdh-H	Formate dehydrogenase - H
Fdh-N	Formate dehydrogenase-N
FdhO	Formate dehydrogenase-O
FdnG	Formate dehydrogenase
FHL	Formate hydrogenlyase
GRAVY	Grand average of hydropathy
GTP	Guanosine triphosphate
Hyd 1	Hydrogenase 1
Hyd 2	Hydrogenase 2
Hyd 3	Hydrogenase 3
Hyd 4	Hydrogenase 4
Km^R	Kanamycin resistance gene
NarG	Nitrate reductase
ORF	Open reading frame
PDB	Protein data bank
PIPER	Docking program is freely available for non-commercial applications
qRT-PCR	Quantitative real-time polymerase chain reaction
RAMPAGE	Ramachandran plot analysis

Introduction

Sustainable and environmentally-friendly use of energy will be essential in replacing fossil fuels. At present, almost all the energy we use in our daily life is generated from fossil fuel which is leading to climate change. According to Das and colleagues [1], this non-renewable source will be depleted and clean energy demands will increase six times in the next few years [2]. As a fuel with high energy content, hydrogen possesses about 122 kJ g^-1 and is classified as a renewable source. Hydrogen is one of the compelling future energy sources since it burns cleanly and it does not contribute to greenhouse emissions [3].

 Numerous efforts are being made to develop a hydrogen economy. Previously, hydrogen was extracted from natural gas and fossil fuels, neglecting the pollution effect [4]. Also, hydrogen can be produced through various methods including the electrolysis of water, steam reformation, and biological processes [5,6]. Among these methods, biological hydrogen production is promising because it can be utilized under ambient temperatures and pressures [6]. In addition, biological processes can utilize wastes like food waste and cellulosic biomass as a carbon source. Hence, the biological production of hydrogen is one of the best approaches to meet the increasing energy demands and also to replace fossil fuels [7]. Economically, fermentative hydrogen production is more convenient and efficient for industrial applications since it consumes less energy, utilizes organic waste, and requires a simple reactor [8]. However, the major barrier in this process is low hydrogen production and the technology used needs

further development prior to being commercialized [7]. In this study, apart from discussing hydrogen production, we elucidate the role of a targeted pseudogene that is present in a microorganism. Therefore, with the information obtained hydrogen production through fermentative method can be improved.

Escherichia coli has been used as a robust model strain for metabolic engineering and protein engineering to enhance bacterial hydrogen production [9]. The hydrogen is produced by *E. coli* through a mixed acid fermentation process [10]. However, even though the main genes related to hydrogen metabolism are well-characterized, there are still some under-studied proteins involved in hydrogen metabolism [11]. Mohd Yusoff et al. reported that the presence of uncharacterised genes and pseudogenes in *E. coli* influences the hydrogen yield [11,12].

A pseudogene is basically a fragment of a gene which is originally derived from a functional gene; however, their own function in the genome is not recognized [13,14]. The definition also has been supported by Goodhead and Darby [15], which explained that pseudogenes are a fragment of once-functional genes that have been silenced by one or more nonsense, frameshift, or missense mutations. Based on the EcoCyc database, mutant *ydfW* is truncated and small (150 bp). From the reported ORFs, 178 of them are annotated as pseudogenes, including 5 RNA pseudogenes and 116 are y-pseudogenes (www.ecogene.org). Recently, some of these pseudogenes genes have been studied; however, pseudogenes are still considered to be rare and tend to be ignored [12]. Beyond these, more small proteins in *E. coli* remain unidentified [16]. The existence of a large number of pseudogenes in *E. coli* raises the question of whether they actually contribute to any cell function [17].

Traditionally, prokaryotes are classified as having compact genomes; i.e., low amount of repetitive DNA. Today, this view has dramatically changed as many repetitive DNA segments have been revealed in bacterial cells [18]. These genes that were originally characterized as not functional seem to have evolutionary consequences due to the adaptation processes [19]. Pseudogenes are prevalent and can persist for a long time, representing a history of once-functional genetic characteristics [20]. Moreover, most researchers believe that proteins from pseudogenes are no longer necessary as the bacterium has adapted to a new and novel environment [20]. In another study conducted by Kuo and Ochman [20], they found that the retention time of pseudogenes in the bacterial genome is very short. The genome appears to be continually created and removed depending on the current surroundings. Later, by using a comparative approach among the same species, only a few of the pseudogenes persisted [14]. This indicates that in a bacterial cell, pseudogenes are produced in high numbers but will be deleted at a rapid rate.

Thus, although the role of pseudogenes still remains unclear, our previous study shows that a few pseudogenes (*ydfW*, *ypdJ*, *yicE*, and *yqiG*) in *E. coli* are related to hydrogen production [12]. In *E. coli*, theoretically from 1 mol of glucose, 2 mol of hydrogen are synthesized [21]. *E. coli* also synthesizes three types of formate dehydrogenases (FdnGHI, FdhF, and FdoGHI)

which oxidize formate into CO_2 [22]. These enzymes consume formate while passing the electrons either to the quinone pool or to the anaerobic respiratory substrates: nitrate, nitrite, dimethyl sulfoxide, trimethylamine N-oxide and fumarate [22]. Formate dehydrogenase-H (Fdh-H) is part of the FHL complex that is located at the cell inner surface of the cytoplasmic membrane for maximum synthesis in the presence of formate under anaerobic conditions [23]. The FHL system is responsible for converting formate to hydrogen and carbon dioxide. FhlA is suppressed by the presence of nitrate due to nitrate-dependent suppression of $fdhF$ gene expression. Thus, FDH-H is believed to play a role in fermentation and but is not involved in electron transfer to either of the respiratory nitrate reductases [22]. Additionally, formate dehydrogenase-N (Fdh-N) is encoded by $fdnGHI$ and is produced maximally under anaerobic conditions in the presence of nitrate [24]. It reduces nitrate to nitrite in the formate-nitrate respiratory chain along with Nar. $fdnG$ expression is activated by NarL and NarP [22]. Finally, formate dehydrogenase-O (Fdh-O) is encoded by $fdoGHI$, and Fdh-O is synthesized either oxygen or nitrate is available in the cell [25]. Fdh-O is coupled with NarZ which is almost similar to FdhN/NarG [26]. In this research, these genes were studied by using microarray and transcriptional analysis. Apart from that, there are four types of hydrogenases (Hyd) responsible for hydrogen production, the gene is encoding the corresponding large subunit: hyaB for Hydrogenase 1, hybC for Hydrogenase 2, hycE for Hydrogenase 3, and hyfG for Hydrogenase 4 [27]. The FDH-H identified as important in hyd 3 and hyd 4 in glucose degradation [28]. The absence of hyd 3 and FDH-H has significantly delayed the hydrogen production and inhibited the cell growth compared with wild [28].

Single-gene knockout mutants of *E. coli* K-12 known as the Keio collection are able to provide information especially to predict the function of unknown protein or pseudogene by deleting the target gene which might reveal its physiological function loss [29]. This method is regarded as a simple and efficient way for gene deletion and comes with extra advantages; easy to cure, low copy number and express λ Red recombinase. Besides, it also allows the researcher to delete target gene precisely and easily remove the antibiotic resistance marker. Integrated with a complete set of ORF clones of *E. coli* (ASKA), the predicted ORF is cloned into a multicopy vector. Thus, the function of the target protein can be studied by the oversupply of the target gene product [30]. Later, by using both of these complete set; Keio and ASKA, a complementation assay can be conducted in order to validate the result obtained [31]. Theoretically, a complementation test should be highly accurate and effective in most cases,

however, there are a few exception types of research that provide misleading answer [32]. However, no further investigation was done to explain the specific role of pseudogenes which in some cases are essential for hydrogen production in *E. coli* (e.g., $ydfW$). Therefore, in this study, the $ydfW$ pseudogene in *E. coli* BW25113 was selected as a model to determine its contribution to hydrogen metabolism.

Materials and methods

Bacteria strains, maintenance, and growth

The wild-type strain *E. coli* BW25113 was obtained from the Yale Coli Genetic Stock Center, and *E. coli* BW25113 Δ$ydfW$: Km^R was obtained from the Keio mutant collection [31]. The strains were streaked on Luria-Bertani (LB) plates and 50 μg/ml of kanamycin was added for the mutant.

Mutant verification

The mutant $ydfW$ was verified using polymerase chain reaction (PCR); primers are shown in Table 1. The PCR reaction was carried out using KOD plus reagent (TOYOBO CO., Ltd); the mixture consists of 2 μl 10 x Buffer for KOD plus, 2 μl dNTPs (2 mM), 1.6 μl $MgSO_4$ (25 mM), 0.3 ml each primer (20 mM), 1.0 ml template (0.2–4 ng/ml) and 0.5 ml KOD plus enzyme (1.0 U/ml). The total volume was 20 μl and the products were amplified for 30 cycles (denaturation at 95 °C for 1 min, annealing at 25 °C for 1 min and elongation at 68 °C for 5 min).

Fermentation and hydrogen assay

Cell suspensions (2 ml) from an overnight culture were inoculated into 18 ml of fresh medium which consists of 5.0 g tryptone, 7.0 g K_2HPO_4, 5.5 g KH_2PO_4, 0.5 g L-cysteine-$HCl.H_2O$, 1.0 g $(NH_4)_2SO_4$, 0.25 g $MgSO_4.7H_2O$, 0.021 g $CaCl_2.2H_2O$, 0.029 g $Co(NO_3)_2.6H_2O$, 0.039 g $Fe(NH_4)_2SO_4.6H_2O$, 2.0 mg nicotinic acid, 0.172 mg Na_2SeO_3, 0.02 mg $NiCl_2$ and 10 ml of trace element solution containing 0.5 g $MnCl_2.4H_2O$, 0.1 g H_3BO_3, 0.01 g $AlK(SO_4)_2.H_2O$, 0.001 g $CuCl_2.2H_2O$ and 0.5 g Na_2EDTA per liter [33]. The mixture was added into 60 ml serum vials with 100 mM of glucose or formate added in separate experiments. The final pH was adjusted using NaOH or HCl to pH 6.8 [33]. This step was done in an anaerobic chamber filled with nitrogen gas. The fermentation was conducted for 24 h, and 50 μl of headspace was analysed for using a gas chromatograph (GC) 6890-N (Agilent Technologies. Glastonbury, CT) equipped with an 80–100 mesh

Table 1 – Primers designed for strain verification.			
Name	Sequences	Description	Reference
$ydfW$-F	5' – CGCTGGAAGGCTATCGAC - 3'	Forward primer	This study
$ydfW$- R	5'- CACCGCCCAGACTGACTC -3'	Reverse primer	
K1	5'- CAGTCATAGCCGAATAGCCT — 3'	Confirmation presence of kanamycin gene	
K2	5'- CGGTGCCCTGAATGAACTGC — 3'		

Porapak Q column (Suppelco, Bellefonte, PA) and a thermal conductivity detector. The column temperature was maintained at 70 °C [34]. The calculation of the hydrogen produced is using equation (1) (Eq. (1)) [35]:

$$Hydrogen\ production\ (\mu mol) = peak\ area$$
$$\times\ standard\ hydrogen\ curve\ \times\ \frac{8500\mu L\ (headspace)}{50\mu L\ (injection)} \quad (1)$$

Analytical assays

For organic acid analysis, the fermentation broth was centrifuged at $18\,900 \times g$. The supernatants were filtered (0.45 μm) and analyzed by high-performance liquid chromatography (HPLC) LC-10AD (Shimadzu, Japan) with a Bio-Rad Aminex HPX-87H column with a reflective index detector at 80 °C. Sulphuric acid (0.005 M) was used as a mobile phase at a flow rate of 0.6 ml/min [36].

For nitrate and nitrite determinations, the analysis was done using the Nitrate/Nitrite Colorimetric Assay Kit (Cayman Chemical, Cat# 780001, USA). The strains were grown at 37 °C until reaching the log phase. Then, the samples were diluted and aliquoted into 96 wells plates. The concentration of nitrate and nitrite were read at 540 nm by using a 96-well ELISA plate reader (Labsystems, USA).

Microarray analysis

The anaerobic chamber was used in order to preserve the anaerobic condition during pellet isolation. Overnight cultures (2 mL) were inoculated into 18 ml of sparged, fresh complex glucose medium and crimp-sealed with butyl rubber stoppers. The cultures were incubated at 37 °C and 250 rpm. After 3 h incubation, about 10 mL of each sample was transferred to the RNA bench in chilled centrifuge tubes, and RNA was isolated from cell pellets following the Affymetrix DNA Microarray Protocol for the Genome 2.0 Array. Affymetrix *E. coli* Genome 2.0 Arrays (P/N: 900550, Lot: 4181865) were used, and the fragmented cDNA amount about was 1.65 μg.

Total RNA extraction and quantitative RT-PCR

The strains were grown aerobically until the exponential phase at 37 °C. One ml of fermentation broth was taken and mixed with an RNA*later* solution at a 1:1 ratio (Ambion, USA). To obtain the cell pellet, the mixture was centrifuged at $18\,900 \times g$ for 2 min. Then, the cell pellet was immersed in a dry ice-95% ethanol mixture for 10 s and stored in -80°C prior to use for RNA extraction. A mini bead-beater Wakenyaku Co. Ltd, Kyoto Japan, model 3011b was used to lyse the cells, and total RNA was isolated via an RNeasy kit (Qiagen, Inc., Valencia, CA). To conduct the transcriptional analysis, a StepOne Real-Time PCR system and Power SYBR green RNA-to CT 1- Step kit (Applied Biosystem, Foster City, CA) were used. In this study, *rrsG* (16S rRNA) was used as a housekeeping gene to normalize the values, and at least three technical replicate samples were analysed. Nine pairs of primers were designed to analyse the target genes; *pflB, hycA, fhlA, hycE, hybC, hyaB, fdhF, narG,* and *fdnG* (Table 2). RNA was extracted from the wild-type was used as the reference template. The concentration of the samples was measured using a Nanodrop spectrophotometer (Thermo Fisher Scientific, USA) and standardized at 100 ng/μl. All data obtained were analysed through relative quantification for qRT-PCR ($2^{-\Delta\Delta CT}$) [11,37].

Computational analysis

YdfW was analyzed using ProtParam and the predicted structure of YdfW was constructed by homology modelling using Phyre 2 [38], Raptor X [39], and Rosetta [40]. The quality of the models was evaluated using the RAMPAGE and PROCHECK [41]. The model was visualized using The PyMOLMolecular Graphics System, Version 1.8 Schrödinger, LLC(PYMOL).

Table 2 – List of primers used for gene expression analysis.			
Target gene	Primers	Description	Reference
rrsG	F- 5′-TATTG CACAATG GGCGCAAG-3′ R- 5′-ACTTAAC AAACC GCCTGCGT-3′	Housekeeping gene for expression data	[42]
hycA	F- 5′-AGCTGGCATCTCTGTTAAACG-3′ R- 5′-GTCATTTTCGGACACTCATCGAC-3	Repressor for FHL complex	[43]
fhlA	F-5′-GTGTATTGCAGGAACAGGAGTTTG-3′ R-5′-GAATACGTTCAGGCGGTAATAGAG-3′	Activator for FHL complex	[44]
HyaB	F- 5′- GGCGTTCGTTGAACGTATCT-3′ R- 5′- CAAGCTGATAGAAGTGCACCAG-3′	Large subunit of hydrogenase 1	[45]
hybC	F- 5′-AGCAGGTTTATAAGGTTGATACCG -3′ R- 5′-AGGTATTCATCGGAATGAGAAGTG-3′	Large subunit of hydrogenase 2	[45]
hycE	F-5′-ATCAG CTGACTG TCACCGTAAAG-3′ R-5′-GTAATCCAACACTTAGTGCCCTTC-3′	Large subunit of hydrogenase 3	[9]
fdhF	F- 5′-GGATT TCTACGGT GCGACTTAC-3′ R- 5′-GGTACT CGTCGGT GAGTTTGTC-3′	Formate dehydrogenase	[34]
narG	F- 5′- TAGGCACTAACCCGCTGACT -3 R- 5′- CCGTTTTCTGGTTGTTCTGC -3′	Respiratory nitrate reductase large subunit	[46]
fdnG	F- 5′-GCTGGCTGATCAGGTGAACT-3′ R- 5′-CCAGTCATAGCCCCAGTTGT-3′	Formate dehydrogenase N large subunit	[22]

Results and discussion

Hydrogen production from the wild-type and ydfW strains

The absence of *ydfW* caused the cell to cease hydrogen production in glucose [12]. The mechanism was not determined. To initiate our work, the *ydfW* mutation was confirmed via PCR. The expected product sizes were obtained for the mutant *ydfW* (1471 bp) and the wild-type (396 bp) (Fig. S1). For hydrogen production in *E. coli,* glucose is the preferred substrate where it is metabolized through a mixed acid fermentation under facultative anaerobic conditions [34]. During this process, some of the major intermediates such as succinate, lactate, acetate, and formate (Fig. 2) are produced. Furthermore, formate is converted into carbon dioxide and hydrogen through the formate hydrogenlyase (FHL) complex. Hence, in order to study the role of the *ydfW* pseudogene in hydrogen production, glucose and formate were supplied to the wild-type and mutant *ydfW* strains (Table 3). Hydrogen production was observed in the wild-type strain for both glucose and formate, 148 ± 8 μmol and 161 ± 5 μmol, respectively, compared the mutant *ydfW,* which produced no hydrogen (Fig. 1A). Hence, YdfW is necessary for the conversion of glucose to formate as well as for the conversion of formate to hydrogen. Also, the growth of the mutant *ydfW* was slower (about 0.3 h^{-1}) compared to the wild-type (0.7 h^{-1}) (Fig. 1B). Hence, both strains were able to ferment glucose and formate, but no hydrogen was produced from the mutant *ydfW*. In the fermentation process, the initial pH was set at 6.8. After 24 h, the final pH with glucose was reduced to pH 4.6 while with formate, only a small pH change was observed (pH 6.2).

Complementation analysis

A complementation analysis was performed using the wild-type and the *yfdW* mutant by utilizing an empty vector (pCA24N) and by producing YfdW (pCA24N-yfdW) in both strains. Plasmids were isolated using the Plasmid Mini Kit (Qiagen, Inc., Valencia, CA) from 5 ml of an overnight culture of each ASKA clone. About 1−2 μl (~100 ng) of isolated plasmid were resuspend in the prepared competent cells (host cells). Electroporated cells were immediately resuspended in 1 ml of LB broth and incubated at 37 °C in a dry bath prior to spreading on LB plates containing 30 μg/ml chloramphenicol (LB + Cm). A single colony from each overnight plate was purified by streaking onto a new LB + Cm plate before use. The Different concentrations of isopropyl β-ᴅ-1-thiogalactopyranoside (IPTG) were added (0, 0.1 and 1.0 mM) to induce the expression of the expressed gene

(Fig. 3). However, an unexpected result was that complementation did not occur. Similaly, no complementation was observed in an independent study of mutant *yieL* (Δ*yeiL*). Thus, the non-complementation may be due to polar effects [47]. On the other hand, several fusion proteins of IPTG-induced expression were failed because most of the amplified protein was insoluble. For these arguments, the mutant must be further analysed further and the actual knockout strain must be developed before any conclusion can be made. However, in this study, the development of a new mutant was not performed since the mutant and ASKA were obtained from the Keio collection and ASKA library. Furthermore, our focus to understand the necessity of the pseudogene ton hydrogen metabolism. However, determination of the ydfW sequence deeply important for our future investigation of the pseudogenes as reported earlier [48]. On the other hand, in genetic, false positive could be happened in complementation test by showing complementation between two mutation in same genes or in other case exhibiting failure complementation due to the two mutations are at different genes [49].

Organic acid production during hydrogen metabolism in E. coli

Fig. 4A shows the concentration of organic acids in the fermentation broth with glucose as the main substrate. In general, four types of organic acids were detected in the fermentation broth. This is one of the reasons that the final pH of the glucose medium fell, with acetic acid and lactic acid accumulating at the end of the fermentation. After lactic acid, acetic acid is the second most prevalent acid in the medium, and it strongly acidifies the final pH. Accumulation of the organic acids inhibits hydrogen production due to the change in internal pH [36]. Therefore, the pH is an indicator of fermentation performance.

Formate is the vital intermediate for hydrogen formation. No formic acid was detected for the wild-type while it was detected for the *ydfW* mutant during fermentation with glucose (Fig. 4A). It seems that in the wild-type strain, all the formate produced was converted into hydrogen but not for the mutant *ydfW*. For confirmation, glucose was replaced with formate (Fig. 4B). As a result, formic acid was observed in the fermentation broth of the *ydfW* mutant and, as expected, little formate was observed for the wild-type. In contrast, a small amount of acetic acid and lactic acid were detected in the broth for both strains. These data also explain why the final pH in formate medium was not significantly changed.

An additional experiment was carried out to investigate whether the formic acid accumulated at the beginning of the

Table 3 − Hydrogen production from wild-type and *ydfW* mutant using 100 mM glucose and formate, respectively.

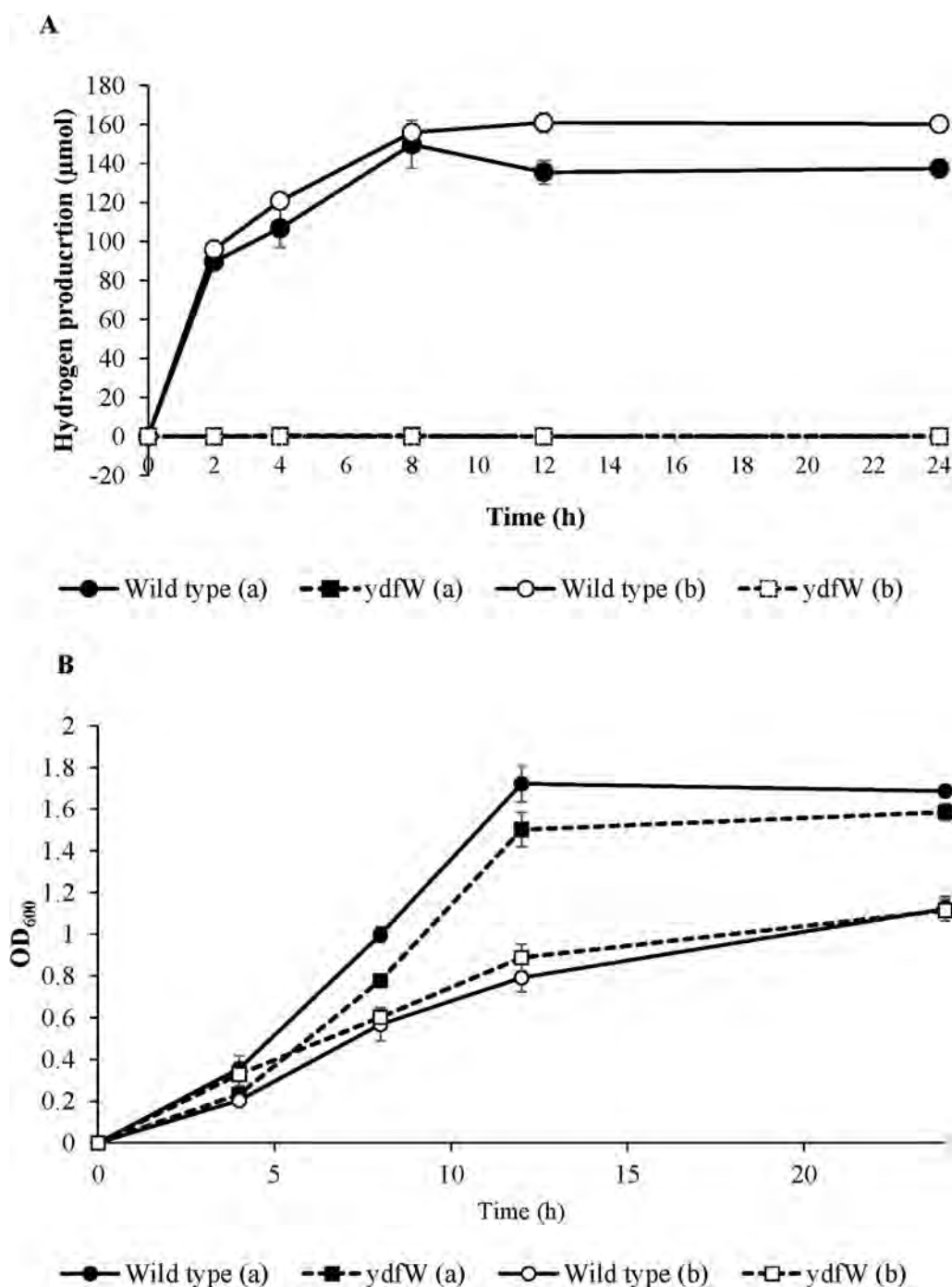| Strain	Glucose				Formate			
	Hydrogen production (μmol)	Specific growth rate (h^{-1})	Initial pH	Final pH	Hydrogen production (μmol)	Specific growth rate (h^{-1})	Initial pH	Final pH
Wild-type	148 ± 8	0.79 ± 0.03	6.8 ± 0.2	4.6 ± 0.1	161 ± 5	0.33 ± 0.03	6.8 ± 0.1	6.2 ± 0.1
Mutant *ydfW*	0.44 ± 0.03	0.74 ± 0.04	6.8 ± 0.1	4.5 ± 0.1	0.88 ± 0.09	0.27 ± 0.01	6.8 ± 0.1	6.5 ± 0.01

Fig. 1 — A) Hydrogen production in wild-type and mutant *ydfW* using glucose (a) and formate (b). B) Growth rate of the wild-type and the *ydfW* mutant using glucose (a) and formate (b).

fermentation or whether it was produced during the whole fermentation process for mutant *ydfW*. As the fermentation begins, with 100 mM of formate, a sample was taken every 6 h up 24 h. Fig. 4 shows the formic acid reduction about 95% in the wild-type while the mutant *ydfW* only 45% reduction. This confirms the mutant *ydfW* has a deficiency in consume formate for hydrogen production compared to the wild-type. The consumption of formate in the mutant *ydfW* was likely due to cell growth and maintenance. The concentration of formic acid should be maintained at 25 mM or 1.15 g/L in order

to obtain maximum hydrogen yield and once formate exceeds 50 mM or 2.30 g/L, the hydrogen cannot be produced [50]. Because the formate concentration for the mutant *ydfW* is greater than 50 mM (Fig. 5), most probably it halts hydrogen production.

Nitrate and nitrite accumulation in the fermentation broth

It was expected that nitrate and nitrite would accumulate in the medium based on the microarray results (below). To
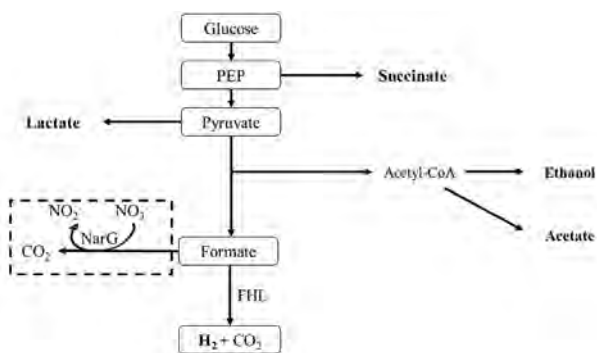
Fig. 2 — Schematic of fermentative hydrogen production and its by-products in *Escherichia coli* with glucose as the main substrate. The hydrogen is produced from formate by the formate hydrogen lyase (FHL) system. The dashed box is induced only in the presence of nitrate.

investigate this hypothesis, nitrate and nitrite concentrations in the *ydfW* culture were assayed (Fig. 6). In the wild-type strain, the concentration of nitrate was reduced significantly and after 8 h of incubation, no nitrate and nitrite was detected in the medium. These results are expected since formate dehydrogenase N (Fdh-N) is maximal under anaerobic conditions when nitrate is present [22]. In contrast, for *ydfW*, the concentration of nitrate was slowly depleted while at the same time the amount of nitrite increased.

In *E. coli*, the presence of nitrate will direct all three formate dehydrogenases towards formate-nitrite respiration [22,24]. Formate acts as an electron donor that is coupled to nitrate reduction which generates a proton gradient. Fdh-N, Fdh-O and Nar possess a similar modular architecture and are located in the cytoplasmic membrane. Moreover, Fdh-N and Nar form a redox loop. Fdh-N takes two protons from the cytoplasm and releases them to the periplasm via the menaquinol oxidation site in Nar. At the same time, two electrons
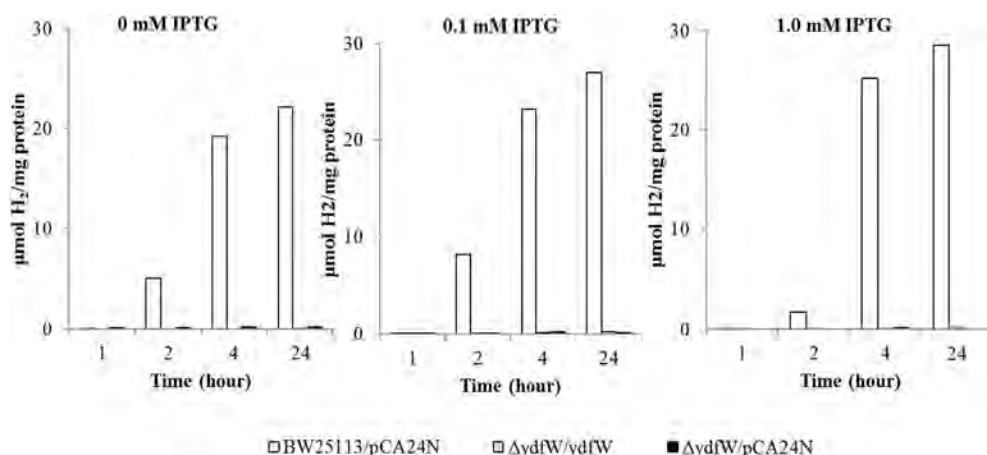

Fig. 3 — Biohydrogen productivity obtained during fermentation using wild-type/pCA24N, *ydfW*/YdfW and *ydfW*/pCA24N induced with different IPTG concentrations in 100 mM glucose complex medium.
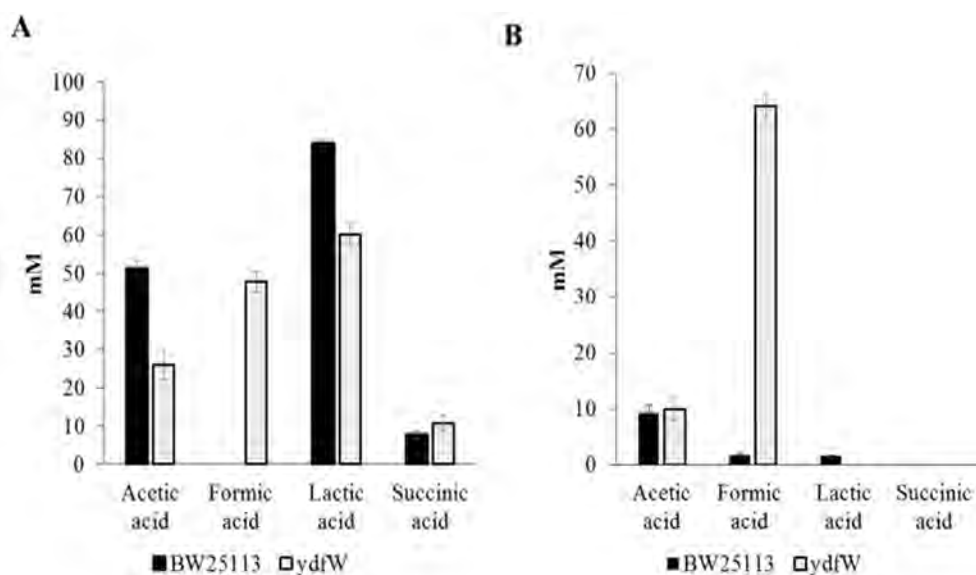

Fig. 4 — Organic acid composition detected using HPLC at the end of the fermentation (24 h). A) Glucose as the substrate, B) Formate as the substrate.
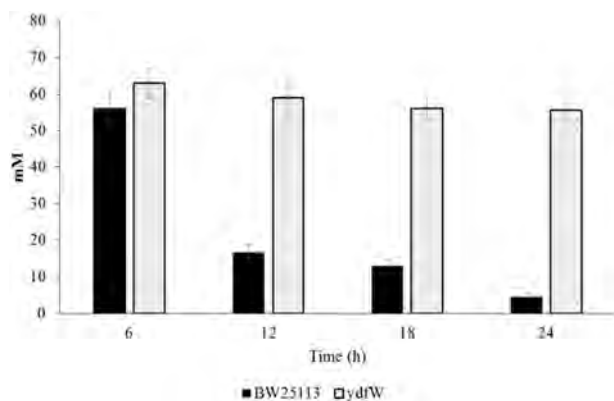
Fig. 5 – Formic acid consumption during fermentation using the wild-type and *ydfW* mutant.

from formate are transferred to the nitrate reduction site which generates $2H^+$ across the membrane [26]. This situation suggests that the accumulated nitrate in the *ydfW* medium is inducing *fdnG* expression and then consuming formate to produce carbon dioxide while oxidizing nitrate to nitrite. To investigate this further, carbon dioxide levels were measured. We found there was a 40% increase in carbon dioxide for the mutant *ydfW* (130 μmol) compared to the wild-type strain (80 μmol).

Microarray analysis

To investigate the impact of YdfW on metabolism, we conducted a microarray for the wild-type and mutant *ydfW*. Based on the microarray analysis, the highest induced gene was the nitrate-nitrite inducible gene *fdnG,* followed by *yeaR* and *fdnH,* respectively (Table 4). FdnG is one of the subunits of formate dehydrogenase N (Fdh-N) which is maximally produced during anaerobic respiration and in the presence of nitrate. In addition, *fdnH*, *fdnI* and *nirB* were also strongly induced and all of them are related to the nitrate respiratory pathway. The function of the product of the *yeaR* gene is unknown but the *yeaR* mutant is sensitive to hydrogen peroxide and cadmium

[51] and is related to the nitrate-regulated activator of *narL* [52]. In summary, all these differentially expressed genes are intimately related to nitrate metabolism. In contrast, *oxyS* was repressed 13-fold; *oxyS* is a small untranslated RNA, and it is responsible for detoxifying oxidative damage in *E. coli*. OxyS can also repress the translation of *fhlA* which subsequently halts formate metabolism and reduces hydrogen production. From this analysis, we hypothesized that the inactivation of YdfW increases nitrate production and it is one of the reasons hydrogen production is repressed.

Transcriptional analysis

To verify the microarray data, 10 of the genes were selected for qRT-PCR. *pflB* encodes pyruvate formate-lyase, a crucial enzyme for converting glucose to formate. Hence, the expression of *pflB* in *ydfW* gives an idea of whether PflB is affected. The FHL complex is comprised of hydrogenase 3 (encoded by *hycABCDEFGHI*) and formate dehydrogenase-H (encoded by *fdhF*) and is activated by FhlA while it is repressed by the *hycA* gene product [34,50]. In glucose metabolism, some of the evolved hydrogen is consumed by hydrogenase 1 (*hyaB*) and hydrogenase 2 (*hybC*) [34]. Also, NarG and FdnG involve in nitrate to nitrite reduction [53].

Fig. 7 shows the expression level of the targeted genes in the *ydfW* mutant compared to the wild-type strain. *pflB* gene was expressed similarly to the wild-type which means that the mutant *ydfW* was able to convert glucose to formate so the reduction in hydrogen production in the mutant *ydfW* is not due to a lack of formate production. Moreover, *fhlA* and *fdhF* were repressed 3.9 and 9.3-fold change in the mutant *ydfW*, respectively. Inactivation of *fhlA* leads to the depletion of hydrogen and disrupts transcription of *fdhF* [54,55]. In the mutant, *ydfW*, *fhlA* and *fdhF* were highly repressed which indicates the FHL complex cannot operate. The other hydrogen metabolism-related genes, *hycA, hycE, hybC,* and *hyaB*, were not significantly affected by the *ydfW* mutation. Three of these genes encode for hydrogenase large subunits *hyaB* for Hyd1, *hybC* for Hyd 2 and *hycE* for Hyd 3 while *hycA* encodes for repressor of hyc transcription, and its protein product will
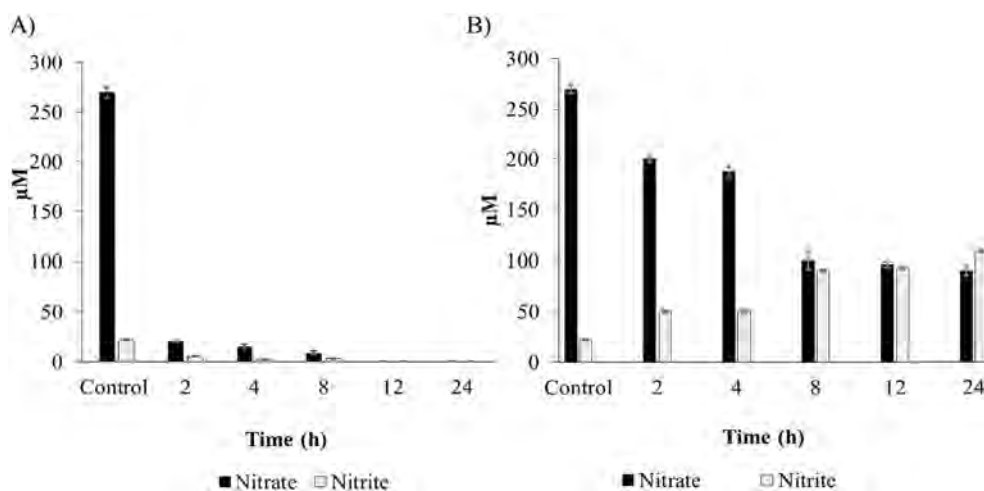

Fig. 6 – Accumulation of nitrate and nitrite in the culture medium during fermentation of glucose for A) wild-type B) and the *ydfW* mutant.

Table 4 – Fold change in promoter activity for the specific genes related to hydrogen metabolism for the *ydfW* mutant vs. the wild-type (microarray analysis).

Gene name	Fold change	Function
fdnG	34	α- subunit of FDH-N, catalytic subunit, tat-signal protein (fdnG operon)
yeaR	24	stress-induced protein, sensitive to hydrogen peroxide resistance (yeaR-yoaG operon)
fdnH	21	β- subunit of FDH-N, co-exported bound to Tat substrate FdnG (fdnG operon)
fdnI	15	γ- membrane subunit of FDH-N, (fdnG operon)
nirB	10	large subunit of Nitrite reductase [NAD(P)H] subunit
yghW	−26	predicted protein, unknown function
oxyS	−13	oxidative stress regulator
yjiH	−8	predicted inner membrane protein, function unknown
yjiG	−7	predicted inner membrane protein with three predicted transmembrane domains
trxC	−5	OxyR regulon, have a role as a redox switch

repress FHL complex [34]. Hyd 1 and Hyd 2 were responsible for hydrogen uptake activity.

Our hypothesis is that since the nitrate level was minimal in the wild-type strain, the FHL complex was able to function properly and hydrogen was produced efficiently. However, the nitrate level was high in the mutant y*dfW* (about 100 mM), which served to induce *fdnG* gene expression. Simultaneously, the *yfdW* mutation was affecting the *fdhF* gene which is crucial for the FHL complex. This condition jeopardizes hydrogen metabolism.

Computational analysis

YdfW has 60 amino acid residues with a molecular mass of 7.1 kDa, and contains Lys (18.3%), Asp (13.3%), and Pro (10.3%) residues. It is also an alkaline protein with an isoelectric point (pI) of 9.80. The low instability index value of 33.51 indicates the protein has a stable structure. The aliphatic index for YdfW is 55.17. This high aliphatic index indicates that the protein is stable for a broad range of temperatures [56]. In addition, this value indicates the relative volume occupied by the amino acids such as alanine, valine, isoleucine, and leucine, which have an aliphatic side chain in their structure. This data is regarded as a positive factor for the increase of
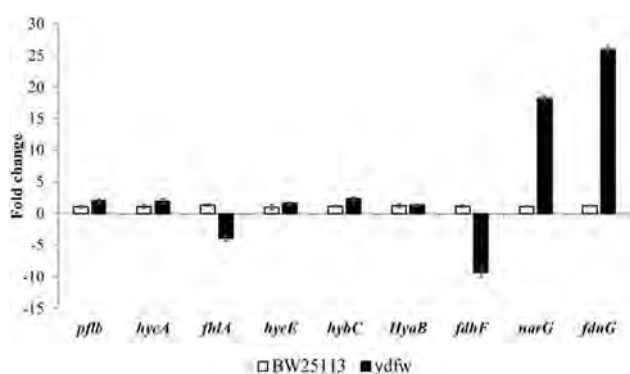

Fig. 7 – Fold change in promoter activity during transcriptional analysis obtained from targeted genes (*pflB, hycA, fhlA, hycE, hybC, hyaB, fdhF, narG, fdnG*) using quantitative real-time reverse transcription polymerase chain reaction.

thermal stability of globular proteins. The GRAVY value was found to be negative (−1.733) which indicates YfdW is a hydrophilic protein. The Protter web service was used to visualize the sequence, topology, and annotations of YdfW protein in *E. coli*. The protein is predicted to be an intracellular protein, in which the protein is located in the cytoplasm.

Table 5 shows the homology modelling and structural analysis of YdfW using three different tools (Phyre 2, Raptor X and SP2). The 3D structure was constructed according to homology modelling. The structure modelled by RaptorX was the best output with the highest score in the favoured region and absence residue in disallowed regions. The parental strain used for constructing the structure was Crystal Structure of viral RdRp in complex with GTP (Protein database ID: 4HDG). Then the structure was further analyzed by ERRAT to evaluate the most reliable predicted 3D model. The final score was 60%. Prior to being docked, the protein structure of Model 1 was further studied.

The docking of the receptor to ligand was conducted by using the ClusPro tool. The ligand will rotate 70,000 times over the receptor. All the possible docking structures were ranked according to the energy scores from PIPER; PIPER is a docking program for non-commercial applications. However, the scoring function is only to cluster it from the start of the rotation. For this reason, the best model should be selected based on prior experimental knowledge of the complex. In the ligand binding search approach, YfdW was assumed as the ligand while Fdh-N was chosen as the receptor. As shown in Fig. S1, YfdW was expected to bind to the lower part of Fdh-N.

Among all the docked structures, the best structure was selected by using the experimental data obtained from the fermentation process. This is because the PIPER scores estimate the true interaction energy, and the scores are for population clusters. Once the docking and clustering processes were done, the models and energy were analysed. The docked model of each protein was bound by the interaction of van der Walls forces and electrostatic binding. To evaluate and view the model, PyMol tool was used. By superimposing the Fdh-N complex with the docked model, it was predicted that YdfW is bound to the FdnH subunit in the Fdh-N complex. However, this data serves only as a prediction and extra work is needed to be done in order to elucidate the real function of YdfW.
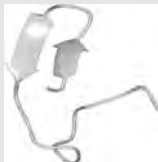
Table 5 – Homology modelling and structural analysis of YdfW using three different tools, Raptor X, Phyre 2, and SP2.

Model	Structure	Ramachandran analysis		
		Favoured	Allowed	Disallowed
Raptor X		90	0	10
Phyre 2		90	0	10
SP2		89.1	8.7	2.2

SDS-PAGE and ydfW sequence analysis

According to the microarray analysis, the *ydfW* mutant has induced *fdnG, fdnH* and *fdnI* (Table 4). These genes encode membrane-bound enzymes comprising which are synthesized when the bacterium grows anaerobically with nitrate as the exogenous electron acceptor. It has been reported in the absence of functioning FDH-N, intracellular formate accumulates and activates fFhlA in the FHL pathway [57]; hence, higher activity of FDH-N should complete with the FHL system and as a result, lower hydrogen production would be obtained. On the other hand, FDH-N is synthesized in the presence of nitrate under anaerobic conditions [58]. To confirm the expression of these genes, SDS-PAGE was carried out to observe the presence and absence of *ydfW*. The analysis was done between BW25113, BW25113 Δ*ydfW* and produced proteins FdnG, FdnH and FdnI in the wild-type (Fig. S2). However, none of the expressed proteins were seen in the mutant *ydfW*. Hence, reconsidered the actual sequence and protein size of YdfW in *E. coli*. *ydfW* in *Escherichia coli* K-12 was obtained from EcoGene (http://ecogene.org), and the gene sequence is for *E. coli* K-12 strain MG1655 [59]. The details about ydfW are as follow:

    Primary Gene Name: *intK′*
    EcoGene Accession Number: EG13835.
    K-12 Gene Accession Number: ECK1561.
    MG1655 Gene Identifier: **b1567**.
    Alternate Gene Symbols: *ydfW*.

    Description: Pseudogene, integrase fragment, Qin prophage.
    CDS position: 1645198–1645380.
    Amino acid size: 60aa.

Table S1 shows various reports for *ydfW* sequences and their protein sequence. As a summary, a previous 75 aa version of an intact IntK(YdfW) (Sequence 5) has been replaced with the current 60 aa intK′(ydfW') C-terminal pseudogene (Sequence 6). An intact version of 437 aa IntK is found in *E. coli* ATCC 8739 (EcolC_2069, NCBI: 169755010) (Sequence 7). The position of YdfW has been re-coordinated from genebank U00096.2 to correspond to the EcolC_2069 in ATCC 8739 (Sequence 1), thus the 75 aa (sequence 4) was replaced with the 60 aa (sequence 5) finally. The position of the YdFW sequence matches the end of the C-terminal EcolC_2069 (Sequence 7).

In addition, from a protein BLAST search, 60 aa of YdfW were predicted to function as either as a transposase, integrase fragment, or integrase core domain protein. YfdW matches at 100% with a 180 aa transposase of *E. coli* DEC9C and the rest of the matches are for integrase related proteins. Wang, Kim [60] have reported that Qin prophage is located at 163031–1650767 in the *E. coli* genome K-12 MG1655. Based on Motif protein analysis, protein function that relates to binding or regulation system was conducted. Both proteins of intact IntK (437 aa) and YdfW' (60 aa) were analyzed. Unfortunately, none of the predicted protein functions match for the binding role or any binding related family. Thus, probably the predicted sequence

of *ydfW* is not appropriate. Thus, no complementation was observed in our work with pCA24N-yfdW. As reported earlier by us, the *yqiG* mutant successfully complemented [48]. To conclude, *ydfW* is essential in hydrogen production. The *ydfW* mutation ceases hydrogen production in *E. coli*. However, the *ydfW* mutation could not be complemented. According to sequence alignments, it seems the annotation of the *ydfW* sequence and its amino acids are not accurate. The actual sequence and protein are deemed important to for further investigations.

Conclusions

Knockout of *ydfW* gene in *E. coli* ceases hydrogen production from glucose and formate. The accumulation of nitrate/nitrite in the medium of the *ydfW* mutant suggests the strain is primed for nitrate metabolism. The high amount of carbon dioxide in the headspace also had supported our hypothesis that the *ydfW* mutant produces CO_2 via the FDH-N complex. This finding is corroborated by the expression level of *fdnG, fdnH, fdnI, nirB* and *narG* in the *ydfW* mutant. The computational study via in silico docking suggests that YdfW and Fdh-H are bound by van der Walls interactions. Probably, YdfW is part of the nitrate pathway. However, the exact sequence of *ydfW* must be identified to retrieve and confirm its function in hydrogen production.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijhydene.2019.04.228.

REFERENCES

[1] Das D, Veziroglu T, Veziroä TN. Hydrogen production by biological processes : a survey of literature. Int J Hydrogen Energy 2001;26:13—28.

[2] Turner JA. Sustainable hydrogen production. Science 2004;305:972—4.

[3] Bockris JOM. The origin of ideas on a Hydrogen Economy and its solution to the decay of the environment. Int J Hydrogen Energy 2002;27:731—40.

[4] Kothari R, Buddhi D, Sawhney RL. Comparison of environmental and economic aspects of various hydrogen production methods. Renew Sustain Energy Rev 2008;12:553—63.

[5] Levin D. Biohydrogen production: prospects and limitations to practical application. Int J Hydrogen Energy 2004;29:173—85.

[6] Manish S, Banerjhee R. Comparison of biohydrogen production processes. Int J Hydrogen Energy 2008;33:279—86.

[7] Kapdan IK, Kargi F. Bio-hydrogen production from waste materials. Enzym Microb Technol 2006;38:569—82.

[8] Meher Kotay S, Das D. Biohydrogen as a renewable energy resource-Prospects and potentials. Int J Hydrogen Energy 2008;33:258—63.

[9] Maeda T, Sanchez-Torres V, Wood TK. Hydrogen production by recombinant *Escherichia coli* strains. Microb Biotechnol 2012;5:214—25.

[10] Fan Z, Yuan L, Chatterjee R. Increased hydrogen production by genetic engineering of *Escherichia coli*. PLoS One 2009;4.

[11] Mohd Yusoff MZ, Maeda T, Sanchez-Torres V, Ogawa HI, Shirai Y, Hassan MA, et al. Uncharacterized *Escherichia coli* proteins YdjA and YhjY are related to biohydrogen production. Int J Hydrogen Energy 2012;37:17778—87.

[12] Mohd Yusoff MZ, Hashiguchi Y, Maeda T, Wood TK. Four products from *Escherichia coli* pseudogenes increase hydrogen production. Biochem Biophys Res Commun 2013;439:576—9.

[13] Lawrence JG, Hendrix RW, Casjens S. Where are the pseudogenes in bacterial. Trends Microbiol 2001;9:535—40.

[14] Lerat E, Ochman H. Psi-Phi: exploring the outer limits of bacterial pseudogenes. Genome Res 2004;14:2273—8.

[15] Goodhead I, Darby AC. Taking the pseudo out of pseudogenes. Curr Opin Microbiol 2015;23:102—9.

[16] VanOrsdel CE, Kelly JP, Burke BN, Lein CD, Oufiero CE, Sanchez JF, et al. Identifying new small proteins in *Escherichia coli*. Proteomics 2018;18:e1700064.

[17] Wen YZ, Zheng LL, Qu LH, Ayala FJ, Lun ZR. Pseudogenes are not pseudo any more. RNA Biol 2012;9:27—32.

[18] Chan WL, Chang JG. Pseudogene-derived endogenous sirnas and their function. Methods Mol Biol 2014;1167:227—39.

[19] Gil R, Latorre A. Factors behind junk DNA in bacteria. Genes 2012;3:634—50.

[20] Kuo CH, Ochman H. The extinction dynamics of bacterial pseudogenes. PLoS Genet 2010;6.

[21] Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H. Enhanced hydrogen production from glucose using ldh- and frd-inactivated *Escherichia coli* strains. Appl Microbiol Biotechnol 2006;73:67—72.

[22] Wang H, Gunsalus RP. Coordinate regulation of the *Escherichia coli* formate dehydrogenase fdnGHI and fdhF genes in response to nitrate, nitrite, and formate: roles for NarL and NarP. J Bacteriol 2003;185:5076—85.

[23] Sawers RG. Formate and its role in hydrogen production in *Escherichia coli*. Biochem Soc Trans 2005;33:42—6.

[24] Darwin A, Tormay P, Page L, Griffiths L, Cole J. Identification of the formate dehydrogenases and genetic determinants of formate-dependent nitrite reduction by *Escherichia coli* K12. J Gen Microbiol 1993;139:1829—40.

[25] Abaibou H, Pommier J, Benoit S, Giordano G, Mandrand-Berthelot MA. Expression and characterization of the *Escherichia coli* fdo locus and a possible physiological role for aerobic formate dehydrogenase. J Bacteriol 1995;177:7141—9.

[26] Jormakka M, Tornroth S, Byrne B, Iwata S. Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. Science 2002;295:1863—8.

[27] Vardar-Schara G, Maeda T, Wood TK. Metabolically engineered bacteria for producing hydrogen via fermentation. Microb Biotechnol 2008;1:107—25.

[28] Trchounian K, Mirzoyan S, Poladyan A, Trchounian A. Hydrogen production by *Escherichia coli* growing in different nutrient media with glycerol: effects of formate, pH,

production kinetics and hydrogenases involved. Int J Hydrogen Energy 2017;42:24026—34.

[29] Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R. Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. Nature 2011;476:355—9.

[30] Ghatak S, King ZA, Sastry A, Palsson BO. The y-ome defines the 35% of Escherichia coli genes that lack experimental evidence of function. Nucleic Acids Res 2019;47:2446—54.

[31] Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2006;2. 2006 0008.

[32] Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, et al. Complete set of ORF clones of Escherichia coli ASKA library (A complete set of E. coli K-12 ORF archive): unique resources for biological research. DNA Res 2005;12:291—9.

[33] Rachman MA, Furutani Y, Nakashimada Y, Kakizono T, Nishio N. Enhanced hydrogen production in altered mixed acid fermentation of glucose by Enterobacter aerogenes. J Ferment Bioeng 1997;83:358—63.

[34] Maeda T, Sanchez-Torres V, Wood TK. Enhanced hydrogen production from glucose by metabolically engineered Escherichia coli. Appl Microbiol Biotechnol 2007;77:879—90.

[35] Maeda T, Sanchez-Torres V, Wood TK. Metabolic engineering to enhance bacterial hydrogen production. Microbial Biotechnology 2008;1:30—9.

[36] Mohd Yasin NH, Fukuzaki M, Maeda T, Miyazaki T, Hakiman Che Maail CM, Ariffin H, et al. Biohydrogen production from oil palm frond juice and sewage sludge by a metabolically engineered Escherichia coli strain. Int J Hydrogen Energy 2013;38:10277—83.

[37] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:45e.

[38] Kelley LA, Sternberg MJE. Protein structure prediction on the web: a case study using the phyre server. Nat Protoc 2009;4:363—73.

[39] Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, et al. Template-based protein structure modeling using the RaptorX web server. Nat Protoc 2012;7:1511—22.

[40] Raveh B, London N, Zimmerman L, Schueler-Furman O. Rosetta FlexPepDockab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. PLoS One 2011;6.

[41] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283—91.

[42] Lee J, Maeda T, Hong SH, Wood TK. Reconfiguring the quorum-sensing regulator SdiA of Escherichia coli to control biofilm formation via indole and N-acylhomoserine lactones. Appl Environ Microbiol 2009;75:1703—16.

[43] Sauter M, Bohm R, Bock A. Mutational analysis of the operon (hyc) determining hydrogenase 3 formation in Escherichia coli. Mol Microbiol 1992;6:1523—32.

[44] Schlensog V, Bock A. Identification and sequence analysis of the gene encoding the transcriptional activator of the formate hydrogenlyase system of Escherichia coli. Mol Microbiol 1990;4:1319—27.

[45] Forzi L, Sawers RG. Maturation of [NiFe]-hydrogenases in Escherichia coli. Biometals 2007;20:565—78.

[46] Wang H, Tseng CP, Gunsalus RP. The napF and narG nitrate reductase operons in Escherichia coli are differentially expressed in response to submicromolar concentrations of nitrate but not nitrite. J Bacteriol 1999;181:5303—8.

[47] Sharma O, Datsenko KA, Ess SC, Zhalnina MV, Wanner BL, Cramer WA. Genome-wide screens: novel mechanisms in colicin import and cytotoxicity. Mol Microbiol 2009;73:571—85.

[48] Zakaria MA, Mohd Yusoff MZ, Zakaria MR, Hassan MA, Wood TK, Maeda T. Pseudogene product YqiG is important for pflB expression and biohydrogen production in Escherichia coli BW25113. 3 Biotech 2018;8:435.

[49] Hawley RS, Gilliland WD. Sometimes the result is not the answer: the truths and the lies that come from using the complementation test. Genetics 2006;174:5—15.

[50] Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H. Enhanced hydrogen production from formic acid by formate hydrogen lyase-overexpressing Escherichia coli strains. Appl Environ Microbiol 2005;71:6762—8.

[51] Lin HY, Bledsoe PJ, Stewart V. Activation of yeaR-yoaG operon transcription by the nitrate-responsive regulator NarL is independent of oxygen-responsive regulator Fnr in Escherichia coli K-12. J Bacteriol 2007;189:7539—48.

[52] Constantinidou C, Hobman JL, Griffiths L, Patel MD, Penn CW, Cole JA, et al. A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as Escherichia coli K12 adapts from aerobic to anaerobic growth. J Biol Chem 2006;281:4802—15.

[53] Pinske C, Sawers RG. A-type carrier protein ErpA is essential for formation of an active formate-nitrate respiratory pathway in Escherichia coli K-12. J Bacteriol 2012;194:346—53.

[54] Mizuno O, Dinsdale R, Hawkes FR, Hawkes DL, Noike T. Enhancement of hydrogen production from glucose by nitrogen gas sparging. Bioresour Technol 2000;73:59—65.

[55] Sanchez-Torres V, Mohd Yusoff MZ, Nakano C, Maeda T, Ogawa HI, Wood TK. Influence of Escherichia coli hydrogenases on hydrogen fermentation from glycerol. Int J Hydrogen Energy 2013;38:3905—12.

[56] Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem 1980;88:1895—8.

[57] Maeda T, Sanchez-Torres V, Wood T. Enhanced hydrogen production from glucose by metabolically engineered Escherichia coli. Appl Microbiol Biotechnol 2007;77:879—90.

[58] Suppmann B, Sawers G. Isolation and characterization of hypophosphite–resistant mutants of Escherichia coli: identification of the FocA protein, encoded by the pfl operon, as a putative formate transporter. Mol Microbiol 1994;11:965—82.

[59] Zhou J, Rudd KE. EcoGene 3.0. Nucleic Acids Res 2013;41:D613—24.

[60] Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, et al. Cryptic prophages help bacteria cope with adverse environments. Nat Commun 2010;1:147.