Biochemical and Biophysical Research Communications 439 (2013) 576–579

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Four products from *Escherichia coli* pseudogenes increase hydrogen production ☆

Mohd Zulkhairi Mohd Yusoff [a,b,c,d,1], Yuya Hashiguchi [a,1], Toshinari Maeda [a,*], Thomas K. Wood [b,c]

[a] Department of Biological Functions and Engineering, Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0196, Japan
[b] Department of Chemical Engineering, Pennsylvania State University, 161 Fenske Laboratory, University Park, PA 16802, USA
[c] Department of Biochemistry and Molecular Biology, Pennsylvania State University, 161 Fenske Laboratory, University Park, PA 16802, USA
[d] Department of Bioprocess Technology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 26 August 2013
Available online 8 September 2013

Keywords:
Biohydrogen
Escherichia coli pseudogene
ydfW
ypdJ
yqiG
ylcE

ABSTRACT

Pseudogenes are considered to be nonfunctional genes that lack a physiological role. By screening 3985 *Escherichia coli* mutants using chemochromic membranes, we found four pseudogenes involved in hydrogen metabolism. Knockouts of pseudogenes *ydfW* and *ypdJ* had a defective hydrogen phenotype on glucose and formate, respectively. Also, the knockout of pseudogene *yqiG* formed hydrogen from formate but not from glucose. For the *yqiG* mutant, 100% hydrogen recovery was obtained by the complementation of YqiG via a plasmid. The knockout of pseudogene *ylcE* showed hydrogen deficiency in minimal media which suggested that the role of YlcE is associated with cell growth. Hence, the products of these four pseudogenes play an important physiological role in hydrogen production in *E. coli*.

1. Introduction

Hydrogen is a clean energy source since it does not produce any harmful compounds when combusted with O_2, and production of hydrogen as a renewable fuel is important as a means to address the problems associated with fossil fuels. Hydrogen is attractive since it has 3-fold higher energy content compared to oil [1]. Another important feature of hydrogen is that it can be easily produced from biological carbon sources through dark fermentation [2,3]. *Escherichia coli* is the most extensively utilized strain for hydrogen production and metabolic engineering studies [4,5]. There are four hydrogenases in the *E. coli* genome, hydrogenase 1, hydrogenase 2, hydrogenase 3, and hydrogenase 4. These hydrogenases possess different functions as comprehensively described beforehand [4,6]. In addition, there are some of uncharacterized genes in *E. coli*, which are responsible for hydrogen production [7]. Thus, the mechanism of hydrogen metabolism is still under investigation.

In *E. coli*, 4296 genes were annotated as open reading frames (ORF) [8,9]. From the reported ORFs, 74 genes were annotated as pseudogenes. However, in the current database in EcoGene, 182 ORFs are listed as pseudogenes and 116 of them are y-genes (uncharacterized genes). Homma et al. found 95 pseudogenes in *E. coli* K-12 MG1655 [10]. Hence, the number of pseudogenes in *E. coli* has not been completely determined.

These pseudogenes are annotated as hypothetical genes with uncertain functions [11]. Wen has suggested pseudogenes were originally functional genes. However, they lost their function due to some genetic reorganization such as deletion/insertion or mutation that occurred at a start or stop codon and might have arisen due to genetic evolution [12–14]. Pseudogenes are also observed in eukaryotes [10].

The presence of pseudogenes raises the question of whether they contribute to any cell function since they consist of 5–20% of the genome. To address whether pseudogenes play an important physiological role, we examined the role of pseudogenes using hydrogen production as the assay since hydrogen production is closely related to cell metabolism and involves a complex pathway [4,5].

2. Materials and methods

2.1. Bacteria strains, maintenance, and growth

Table 1 describes the strains used in this study [9,15]. These mutants were routinely streaked on lysogenic broth (LB) plates

* Corresponding author. Fax: +81 93 695 6005.
E-mail address: toshi.maeda@life.kyutech.ac.jp (T. Maeda).

[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.09.016

containing 100 μg/ml kanamycin or 30 μg/ml chloramphenicol where appropriate and grown at 37 °C for 12–15 h. For the overnight inoculum, single colonies were grown aerobically in 50 ml LB medium in a shaking incubator at 120 rpm for 12–15 h at 37 °C. Cell growth was measured based on cell turbidity at OD_{600} by UV/VIS spectrophotometric (JASCO V-530) while total cell amounts were determined based on cell turbidity at OD_{600} using the relationship of 0.22 mg ml^{-1} OD_{600}^{-1} [16]. All experiments were conducted using at least three independent biological replicates.

2.2. Biohydrogen membrane screening and biohydrogen assay

Biohydrogen production screening from the Keio mutant library was carried out using chemochromic membranes (GVD Corp., Cambridge, MA). This membrane is used to detect hydrogen gas produced from the colonies on LB plate medium via a colorimetric response [17].

For the biohydrogen production assay, 1.0 ml of overnight culture was inoculated into 9 ml of fresh complex medium containing glucose or formate in 34 ml glass vials inside an anaerobic chamber. In order to provide anaerobic conditions throughout the process, the empty vials and overnight culture were sparged with nitrogen for 2 min and 5 min prior to use, respectively [7]. The biohydrogen assay was conducted for 24 h. The amount of biohydrogen generated in the headspace was measured using gas chromatography (GC) with a thermal conductivity detector as described [18]. All the media preparation and analytical analysis such as complex and minimal medium, sample preparation and organic acids analysis were as described [7,18–20].

2.3. Polymerase chain reaction verification

The four gene knockouts *ydfW*, *ylcE*, *ypdJ* and *yqiG* were verified using the polymerase chain reaction (PCR). A pair of specific primers for each mutant was designed to confirm the correct mutant allele along with k1 and k2 primers for verifying the presence of the kanamycin marker (Table 2). PCR amplification and verification was done using KOD plus reagent (TOYOBO CO., Ltd) as described [7].

2.4. Complementation analysis test

Plasmids were isolated using the Plasmid Mini Kit (Qiagen, Inc., Valencia, CA) from 5 ml of overnight culture of each ASKA clone. Competent cells were prepared according to our lab protocol modified from Datsenko and Wanner [21]. The 30 μl of each competent cell preparation was gently resuspended with 1–2 μl (~100 ng) of plasmid. Transformation was carried out at 12.5 kV/cm using the Bio-Rad Gene Pulser II. Electroporated cells were immediately resuspended in 1 ml of LB broth and incubated at 37 °C in a dry bath incubator for about 1 h prior to spreading on LB plates containing 30 μg/ml chloramphenicol (LB + Cm) and incubated overnight. A single colony from each overnight plate was purified by streaking onto a new LB + Cm plate before use.

Table 1
E. coli strains and plasmids used in this study.

[a]Strains and plasmids	Genotype/relevant characteristics	References
BW25113	F^-Δ(araD-araB)567ΔlacZ4787 (::rrnB-3)λ^- rph-1 Δ(rhaD-rhaB) 568 hsdR514; parental strain for the Keio collection.	Yale Coli Genetic Stock Center
BW25113 *ydfW*	BW25113 *ΔydfW::kan* Km^R	[9]
BW25113 *ylcE*	BW25113 *ΔylcE::kan* Km^R	[9]
BW25113 *ypdJ*	BW25113 *ΔypdJ::kan* Km^R	[9]
BW25113 *yqiG*	BW25113 *ΔyqiG::kan* Km^R	[9]
ASKA JW1559	Harboring YdfW Cm^R	[15]
ASKA JW5386	Harboring YpdJ Cm^R	[15]
ASKA JW5507	Harboring YqiG Cm^R	[15]

Km^R is kanamycin resistance, Cm^R is chloramphenicol resistance.
[a] All Keio mutant and ASKA clones provided by National of Institute of Genetics (Japan).

3. Results and discussion

In this study, the role of *E. coli* pseudogenes in biohydrogen production was determined. By screening 3985 Keio mutants with hydrogen membranes, four pseudogenes were identified, *ydfW*, *ylcE*, *ypdJ*, and *yqiG*, that upon deletion, reduced hydrogen production. The knockouts for these four strains were verified through PCR with three pairs of primers.

3.1. Biohydrogen production

The four pseudogene knockout strains from hydrogen membrane screening were tested further by measuring biohydrogen production using glucose and formate fermentation. Theoretically, a mole of glucose may produce 2 mol of formate which can be converted to 2 mol of hydrogen through pyruvate as an intermediate [22]. Thus, glucose and formate are important for bacterial hydrogen production with *E. coli*.

Fig. 1A shows biohydrogen production from 100 mM complex glucose using mutants of *ydfW*, *ylcE*, *ypdJ* and *yqiG* compared to BW25113 (wild-type). After 24 h, the wild type strain had a biohydrogen productivity of 45 μmol/mg protein. Similar productivity was obtained with the *ylcE* mutant at 43 μmol/mg protein, respectively. In contrast, the *ydfW*, *ypdJ* and *yqiG* mutants had no appreciable biohydrogen production. Hence, the products of these three pseudogenes are required for biohydrogen production in *E. coli*.

Biohydrogen production was also assayed using 100 mM formate as the initial substrate. The same trend in productivity was obtained from all strains except the *yqiG* mutant formed more biohydrogen with formate (Fig. 1B). Hence, the product of *yqiG* probably plays a role during glucose conversion to formate. The consistent significant deficiencies in biohydrogen production due to deletions of *ydfW* and *ypdJ* confirm the products of these genes are important for hydrogen production.

Unlike in complex formate and complex glucose medium, the *ylcE* mutant in minimal glucose medium had significantly lower biohydrogen production than the wild-type strain (Fig. 1C). The lower productivity probably due to inadequate of growth supplement in the minimal medium. Based on this result, YlcE probably plays a role in cell growth during glycolysis since lower biohydrogen only observed in minimal medium.

3.2. Organic acids analysis from the mutants using minimal medium

pH is one of the vital parameters that influences biological hydrogen production [20,23]. The pH value generally decreases due to the production of volatile fatty acids during fermentation [23,24]. By altering initial pH from 6 to 7, these acids influence biohydrogen production [3]. Thus, we determined the organic acid content during biohydrogen production from the mutant strains that lack the pseudogenes in comparison with wild-type. Fig. 2 shows the organic acid concentrations analyzed after fermentation in minimal glucose medium. In general, four main organic acids were detected, succinic acid, lactic acid, formic acid and acetic acid. All acids were detected in all strains at different amounts compared to the wild-type strain. Notably, only the wild-strain lacked formic acid. This phenomena indicates the wild-type strain consumed the formic acid for biohydrogen production. Hence, deletion

Table 2
Primers designed according to the target genes allele for strains verification.

Name/targeted gene/function	Sequence	Relevant characteristics
ydfW-confirm	f-5′-CTGAAGCATTCTGGGCTTTG-3′ r-5′-GATCAATTATGAATTGCAAC-3′	Confirmation absence of *ydfW* gene (this study)
ylcE-confirm	f-5′-GGTGTGTAAATTCTAGCCCC-3′ r-5′-GGGGCGCAAAAGCATAACATC-3′	Confirmation absence of *ylcE* gene (this study)
ypdJ-confirm	f-5′-GACTCTCGGTTGGATCATCTG-3′ r-5′-GTGTTGCATGAGTGCTATCTCC-3′	Confirmation absence of *ypdJ* gene (this study)
yqiG-confirm	f-5′-GAACCTTGCAGAGGCGTTCG-3′ r-5′-TCATTATGATACGGGTACGATCC-3′	Confirmation absence of *yqiG* gene (this study)
K1	5′-CAGTCATAGCCGAATAGCCT-3′	Confirmation presence of kanamycin [21]
K2	5′-CGGTGCCCTGAATGAACTGC-3′	Confirmation presence of kanamycin [21]

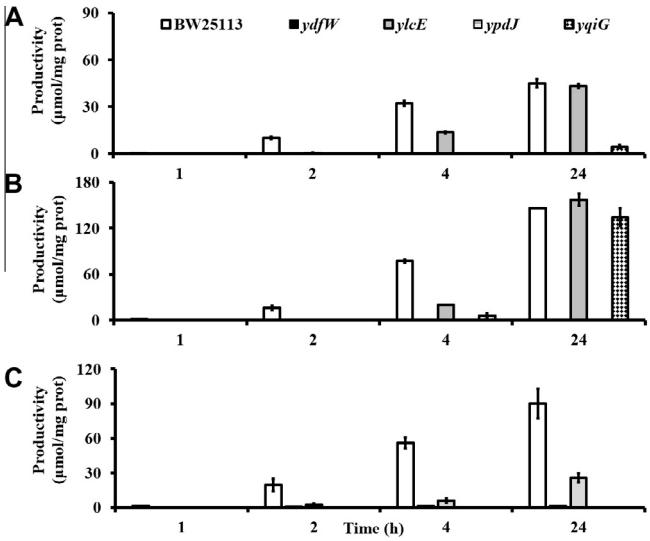

Fig. 1. Biohydrogen productivity obtained during fermentation using BW25113 (wild type, open column), *ydfW* (black column), *ylcE* (grey column), *ypdJ* (strike column) and *yqiG* (diamond column). (A) 100 mM glucose complex medium, (B) 100 mM formate complex medium, and (C) 100 mM minimal glucose medium.

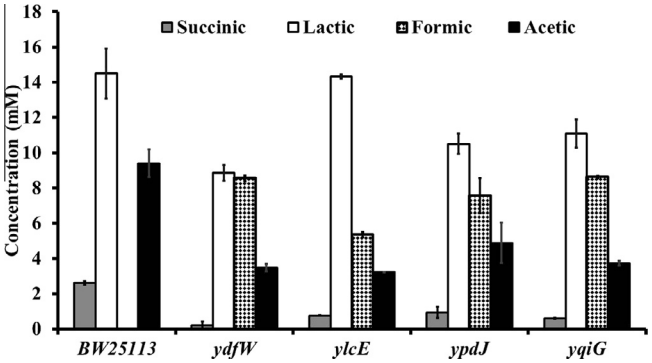

Fig. 2. Organic acid composition detected using HPLC from BW25113 (wild type), *ydfW*, *ylcE*, *ypdJ* and *yqiG* at the end of fermentation in minimal glucose medium.

of these pseudogenes affected the production of organic acids that influence hydrogen metabolism.

3.3. Growth and glucose consumption analysis

Based on the decrease in hydrogen produced by the *ylcE* mutant in minimal glucose medium, we assayed glucose consumption by the strains with deletions in the four pseudogenes with complex glucose. For the wild-type strain, there was 57% utilization of glucose after 24 h fermentation period. Similar glucose consumption was obtained with the *ylcE* pseudogene mutant with 58% glucose utilized. This result confirms that the *ylcE* mutation only has an effect in minimal medium. In contrast, the non-hydrogen producing strains, *ydfW*, *ypdJ* and *yqiG*, had less glucose consumed with 31%, 30% and 32%, respectively. The reduced glucose consumption phenomena observed with all these strains was believed due to the acidic conditions since at the end of the fermentation similar pH values were determined at pH 4.4–4.6. This low pH reduces hydrogen production in the broth through organic acids secretion [24]. Moreover, the slightly higher glucose consumption observed in the *ylcE* mutant is consistent with the higher cells density obtained at the end of fermentation after 24 h with 6.5 ± 0.4 mg protein, compared to the wild-type, 6.0 ± 0.2 mg protein, respectively. The *ydfW*, *ypdJ* and *yqiG* mutants had lower cell densities of 5.0 ± 0.7, 2.98 ± 0.01 and 3.3 ± 0.3 mg protein, respectively. Less glucose consumed demonstrates that the glucose was used to maintain cells but that it was not used for other by-products especially hydrogen. These data for glucose consumption corroborate the hydrogen generation data for the *ydfW*, *ypdJ* and *yqiG* mutants. Lower glucose utilization led to less hydrogen production.

3.4. Complementation analysis test

For the three strains with dramatically reduced biohydrogen production, *ydfW*, *ypdJ*, and *yqiG*, the corresponding protein was produced from the ASKA plasmid and biohydrogen production was assayed to see if the phenotype could be complemented. The empty pCA24N plasmid was used as a control for each strain and all the strains with the plasmid were induced by 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at the initial stage of fermentation. We found that biohydrogen production could be 100% complemented for the *yqiG* mutant but not for the *ydfW* and *ypdJ* mutants.

yqiG encodes for a 821 amino acid (aa) protein [25], and *yqiG* is classified as an interrupted gene due to insertion sequence (IS) insC9 and insD9 at the initial sequence [11,26]. The IS insertion probably interrupts the promoter of the coding sequence, and there is no function listed for the encoded protein. However, from our analysis, the *yqiG* mutation abolishes biohydrogen production so this protein is necessary for biohydrogen production.

ydfW and *ypdJ* encode for 75 and 46 aa polypeptides, respectively. However, different sizes of YpdJ are reported in genome databases with a size of 17 aa also listed. Thus, the actual sequence and protein encoded by pseudogene *ypdJ* is still unresolved. As discussed by Hemm and fellow researchers, proteins between 16 and 50 aa are referred to as small proteins [27]. From their study, small proteins are difficult to isolate and identify. However, they indicate these small proteins are important as leader peptides, ribosomal

proteins and toxic proteins [27]. Our inability to complement the hydrogen-deficient phenotype for the *ydfW* and *ypdJ* mutants shows that probably the proteins encoded on the plasmids are not the correct size.

In summary, we identified four proteins from pseudogenes, YdfW, YlcE, YpdJ, and YqiG, that increase biohydrogen production in *E. coli*. Although further work needs to be conducted to further characterize these proteins, we have demonstrated that pseudogenes play an important role in cell physiology.

Acknowledgments

Special acknowledgement goes to the NBRP-*E. coli* at the National of Institute of Genetics (Japan) for providing Keio mutants and ASKA clones. In addition, this research was performed as an international collaboration under the Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation (R2217) supported by the Japan Society for the Promotion of Science. In addition, we acknowledge the Japan Science Society through Sasakawa Scientific Research Grant (24-521) and JGC-S scholarship foundation for supporting this research.

References

[1] T. Maeda, G. Vardar, W. Self, T.K. Wood, Inhibition of hydrogen uptake in *Escherichia coli* by expressing the hydrogenase from the Cyanobacterium *Synechocystis* sp. PCC 6803, BMC Biotechnol. 7 (2007) 25.

[2] M.Z.M. Yusoff, M.A. Hassan, S. Abd-Aziz, N.A.A. Rahman, Start-up of biohydrogen production from palm oil mill effluent under non-sterile condition in 50 L continuous stirred tank reactor, Int. J. Agric. Res. 4 (2009) 163–168.

[3] N.H. Mohd Yasin, N.A.A. Rahman, H.C. Man, M.Z.M. Yusoff, M.A. Hassan, Microbial characterization of hydrogen-producing bacteria in fermented food waste at different pH values, Int. J. Hydrogen Energy 36 (2011) 9571–9580.

[4] G. Vardar-Schara, T. Maeda, T.K. Wood, Metabolically engineered bacteria for producing hydrogen via fermentation, Microb. Biotechnol. 1 (2008) 107–125.

[5] T. Maeda, V. Sanchez-Torres, T.K. Wood, Metabolic engineering to enhance bacterial hydrogen production, Microb. Biotechnol. 1 (2008) 30–39.

[6] K. Trchounian, A. Trchounian, Hydrogenase 2 is most and hydrogenase 1 is less responsible for H2 production by *Escherichia coli* under glycerol fermentation at neutral and slightly alkaline pH, Int. J. Hydrogen Energy 34 (2009) 8839–8845.

[7] M.Z. Mohd Yusoff, T. Maeda, V. Sanchez-Torres, H.I. Ogawa, Y. Shirai, M.A. Hassan, T.K. Wood, Uncharacterized *Escherichia coli* proteins YdjA and YhjY are related to biohydrogen production, Int. J. Hydrogen Energy 37 (2012) 17778–17787.

[8] N. Yamamoto, K. Nakahigashi, T. Nakamichi, M. Yoshino, Y. Takai, Y. Touda, A. Furubayashi, S. Kinjyo, H. Dose, M. Hasegawa, K.A. Datsenko, T. Nakayashiki, M. Tomita, B.L. Wanner, H. Mori, Update on the Keio collection of *Escherichia coli* single-gene deletion mutants, Mol. Syst. Biol. 5 (2009) 335.

[9] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner, H. Mori, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection, Mol. Syst. Biol. 2 (2006).

[10] K. Homma, S. Fukuchi, T. Kawabata, M. Ota, K. Nishikawa, A systematic investigation identifies a significant number of probable pseudogenes in the *Escherichia coli* genome, Gene 294 (2002) 25–33.

[11] J. Zhou, K.E. Rudd, EcoGene 3.0, Nucleic Acids Res. 41 (2013) D613–D624.

[12] W. Li, W. Yang, X.-J. Wang, Pseudogenes: pseudo or real functional elements?, J Genet. Genomics 40 (2013) 171–177.

[13] G. Drouin, Processed pseudogenes are more abundant in human and mouse X Chromosomes than in autosomes, Mol. Biol. Evol. 23 (2006) 1652–1655.

[14] Y.Z. Wen, L.L. Zheng, L.H. Qu, F.J. Ayala, Z.R. Lun, Pseudogenes are not pseudo any more, RNA Biol. 9 (2012) 27–32.

[15] M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, 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. 12 (2006) 291–299.

[16] T. Maeda, V. Sanchez-Torres, T.K. Wood, *Escherichia coli* hydrogenase 3 is a reversible enzyme possessing hydrogen uptake and synthesis activities, Appl. Microbiol. Biotechnol. 76 (2007) 1035–1042.

[17] T. Maeda, V. Sanchez-Torres, T.K. Wood, Protein engineering of hydrogenase 3 to enhance hydrogen production, Appl. Microbiol. Biotechnol. 79 (2008) 77–86.

[18] N.H. Mohd Yasin, M. Fukuzaki, T. Maeda, T. Miyazaki, C.M. Hakiman Che Maail, H. Ariffin, T.K. Wood, Biohydrogen production from oil palm frond juice and sewage sludge by a metabolically engineered *Escherichia coli* strain, Int. J. Hydrogen Energy 38 (2013) 10277–10283.

[19] H. Shinagawa, R. Nishiyama, S. Miyao, M. Kozaki, Organic acid composition and quality of Japanese shibazuke pickles, Food Sci. Technol. Int. (Tokyo) 3 (1997) 170–172.

[20] V. Sanchez-Torres, M.Z. Mohd Yusoff, C. Nakano, T. Maeda, H.I. Ogawa, T.K. Wood, Influence of *Escherichia coli* hydrogenases on hydrogen fermentation from glycerol, Int. J. Hydrogen Energy 38 (2013) 3905–3912.

[21] K.A. Datsenko, B.L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, Proc. Natl. Acad. Sci. 97 (2000) 6640–6645.

[22] E. Seol, Y. Jang, S. Kim, Y.-K. Oh, S. Park, Engineering of formate-hydrogen lyase gene cluster for improved hydrogen production in *Escherichia coli*, Int. J. Hydrogen Energy 37 (2012) 15045–15051.

[23] Z. Rasdi, T. Mumtaz, N.A. Abdul Rahman, M.A. Hassan, Kinetic analysis of biohydrogen production from anaerobically treated POME in bioreactor under optimized condition, Int. J. Hydrogen Energy 37 (2012) 17724–17730.

[24] H.M.Y. Nazlina, A.R. Nor Aini, F. Ismail, M.Z.M. Yusof, M.A. Hassan, Effect of different temperature, initial pH and substrate composition on biohydrogen production from food waste in batch fermentation, Asian J. Biotechnol. 1 (2009) 42–50.

[25] I.M. Keseler, J. Collado-Vides, A. Santos-Zavaleta, M. Peralta-Gil, S. Gama-Castro, L. Muniz-Rascado, C. Bonavides-Martinez, S. Paley, M. Krummenacker, T. Altman, P. Kaipa, A. Spaulding, J. Pacheco, M. Latendresse, C. Fulcher, M. Sarker, A.G. Shearer, A. Mackie, I. Paulsen, R.P. Gunsalus, P.D. Karp, EcoCyc: a comprehensive database of *Escherichia coli* biology, Nucleic Acids Res. 39 (2011) D583–D590.

[26] K. Hayashi, N. Morooka, Y. Yamamoto, K. Fujita, K. Isono, S. Choi, E. Ohtsubo, T. Baba, B.L. Wanner, H. Mori, T. Horiuchi, Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110, Mol. Syst. Biol. 2 (2006). 2006 0007.

[27] M.R. Hemm, B.J. Paul, T.D. Schneider, G. Storz, K.E. Rudd, Small membrane proteins found by comparative genomics and ribosome binding site models, Mol. Microbiol. 70 (2008) 1487–1501.