PROTEIN ENGINEERING OF FORMATE DEHYDROGENASE H (Fdh-H) TO ENHANCE HYDROGEN PRODUCTION VIA ERROR-PRONE PCR

JESSICA PAOLA RAMIREZ ANGULO

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICOQUIMICAS
ESCUELA DE INGENIERIA QUIMICA
BUCARAMANGA
2009

PROTEIN ENGINEERING OF FORMATE DEHYDROGENASE H (Fdh-H) TO ENHANCE HYDROGEN PRODUCTION VIA ERROR-PRONE PCR

JESSICA PAOLA RAMIREZ ANGULO

Trabajo de grado presentado como requisito para optar el titulo de Ingeniera Química

Director
ALVARO RAMIREZ GARCIA
Ingeniero Quimico Ph.D
Universidad Industrial de Santander

Co-Director
THOMAS K. WOOD
Ingeniero Quimico Ph.D
Texas A&M University

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICOQUIMICAS
ESCUELA DE INGENIERIA QUIMICA
BUCARAMANGA
2009

ACKNOWLEDGEMENTS

A DIOS.

A MI MADRE, quien siempre me ha apoyado en mis decisiones, me ha guiado y quien nunca ha dejado de creer en mi, para ti con todo el amor del mundo.

A MIS ABUELOS, quienes han permanecido fieles a mi lado y a quienes amo.

A MI FAMILIA, que me ha tendido su mano amiga y nunca me ha dejado sola.

AL DR THOMAS K. WOOD, por la oportunidad de trabajar en su laboratorio, por creer en mi y por asesorarme durante el tiempo que duro mi investigación.

AL DR ALVARO RAMIREZ GARCIA, quien me apoyo en la finalización de este proyecto de grado y por sus palabras de vida.

A MIS COMPAÑEROS DE GRUPO, Viviana, Andrea, Brian y al Dr Michael, que estuvieron ahi para darme sugerencias y por ser unos amigos tan incondicionales.

AL DEPARTAMENTO DE INGENIERIA QUIMICA DE LA UNIVERSIDAD DE TEXAS A&M y de la UNIVERSIDAD INDUSTRIAL DE SANTANDER, que me brindaron la oportunidad de realizar esta pasantía que amplio mis conocimientos.

A TODOS MIS AMIGOS, que con su apoyo y amistad incondicional hicieron que mi paso por la Universidad fuera una época memorable e involvidable, los quiero.

Jessica Paola Ramirez Angulo

TABLE OF CONTENTS

LIST OF ATTACHMENTS

1. LIST OF FIGURES

2. LIST OF TABLES

ABSTRACT

TITLE: PROTEIN ENGINEERING OF FORMATE DEHYDROGENASE H (Fdh-H) TO ENHANCE HYDROGEN PRODUCTION VIA ERROR-PRONE PCR*

AUTHOR: Jessica Paola Ramirez Angulo **

KEYWORDS: Protein engineering. Formate hydrogen lyase (FHL) complex. P1 transduction. Error-prone PCR. Saturation mutagenesis

The formate hydrogen lyase (FHL) complex of *Escherichia coli* is responsible for the conversion of formate to carbon dioxide (CO_2) and hydrogen (H_2) under anaerobic conditions and in the absence of electron acceptors. The active complex consists of seven proteins, six from the hydrogenase 3 encodes by the *hyc* operon (HycBCDEFG) and formate dehydrogenase H (Fdh-H). Fdh-H, encoded by the *fdhF* gene, is a 79 kDa cytoplasmic protein and contains selenocysteine, molybdenum and a [4Fe-4S] cluster at its active site. In this study, we engineered *fdhF* (715 amino acids) for enhanced hydrogen production by an error-prone polymerase chain reaction (epPCR) using a host that lacked hydrogenase 1 (*hyaB* encodes the large subunit) and hydrogenase 2 (*hybC* encodes the large subunit) activity and the formate dehydrogenase H via the *hyaB hybC fdhF* mutations. To generate random mutations we use the plasmid *pCA24N-fdhF* as the template to perform the epPCR.

One enhanced *fdhF* variant was obtained with a novel chemochromic membrane screen that directly detected hydrogen from individual colonies. The epPCR variant contained one mutation (Y513H) and had 12-fold higher hydrogen-producing activity than wildtype *fdhF*.

In addition, we perform saturation mutagenesis in this variant to obtain one mutant (H513Q) that had 6-fold higher hydrogen-producing activity than mutant (Y513H).

---

* Proyecto de Grado desarrollado en Texas A&M University, College Station, Texas, USA.

** Facultad de Ingenierías Físico-Químicas, Escuela de Ingeniería Química, Director: Dr. Alvaro Ramirez Garcia, Co-Director: Thomas Wood

RESUMEN

TITULO: INGENIERIA DE LAS PROTEINAS APLICADA A LA DEHIDROGENASA DE FORMATO H (Fdh-H) PARA INCREMENTAR LA PRODUCCION DE HIDROGENO POR MEDIO DEL USO DE ERROR-PRONE PCR*

AUTOR: Jessica Paola Ramirez Angulo **

PALABRAS CLAVES: Ingeniería de las Proteínas. Complejo Liasa para producir Hidrógeno del Formato (FHL). Transducción P1. Error-prone PCR. Mutagénesis Saturada.

El complejo Liasa para producir Hidrógeno del Formato (FHL) de la *Escherichia coli* es el responsable de la conversión del formato en dióxido de carbono (CO_2) e hidrógeno (H_2) bajo condiciones anaeróbicas y en la ausencia de receptores de electrones. El principio activo del complejo consta de siete proteínas, seis provenientes de la hidrogenada 3 codificada por el operador hyc (HycBCDEFG) y uno de la deshidrogenasa del formato H (Fdh-H). Fdh-H, codificada por el gen fdhF, es una proteína citoplasmática de 79 kDa y contiene selenocisteina, molibdeno y un grupo de [4Fe-4S] en su sitio activo. En este estudio, aplicamos ingeniería a la fdhF (715 amino ácidos) para aumentar la producción de hidrógeno por un proceso de error-prone y reacción en cadena de la polimerasa (epPCR) usando un huésped que carecía de actividad de la hidrogenada 1 (*hyaB* codifica la subunidad mas grande) y la hidrogenada 2 (*hybC* codifica la subunidad mas grande) y de la deshidrogenasa del formato H a través de las mutaciones hyaB hybC fdhF. Para generar las mutaciones aleatorias mediante la implementación de la técnica de epPCR, se utilizo como plantilla el plasmido *pCA24N-fdhF.*

Una variante mejorada de la fdhF se obtuvo con una membrana chemochromica que detecta directamente la producción de hidrógeno a partir de colonias individuales. La variante obtenida a travéz de epPCR tiene una mutación (Y513H) y produce 12 veces más hidrógeno que la cepa sin mutaciones de la fdhF.

Además, realizamos mutagénesis saturada en esta variante para obtener una mutante (H513Q) que produce 6 veces más hidrógeno que la mutante (Y513H).

---

* Proyecto de Grado desarrollado en Texas A&M University, College Station, Texas, USA.

** Facultad de Ingenierías Físico-Químicas, Escuela de Ingeniería Química, Director: Dr. Alvaro Ramirez Garcia, Co-Director: Thomas Wood

INTRODUCTION

Hydrogen is the most abundant element in the universe, is renewable, efficient, and clean, and is utilized for fuel cells in portable electronics, power plants, and the internal combustion engine. Also, Hydrogen is a promising fuel as it has a higher energy content than oil (142 MJ kg^{-1} for H_2 versus 42 MJ kg^{-1} for oil). Most of the hydrogen now produced globally is by the process of steam reforming and the water–gas shift reaction, or as a by-product of petroleum refining and chemicals production. Use of biological methods of hydrogen production should significantly reduce energy costs, as these processes do not require extensive heating (or extensive electricity as in electrolysis plants). Biological methods depend on hydrogenases that catalyse the reaction $2H^+ + 2e^- \leftrightarrow H_2$ (g). Hydrogen gas may be produced through either photosynthetic or fermentative processes, but fermentative hydrogen production is more efficient than photosynthetic ones.

We chose to metabolically engineer *Escherichia coli* for hydrogen production as this is the best-characterized bacterium (i.e. has well-established metabolic pathways) and it is one of the easiest strains to manipulate genetically via P1 phage transduction, this technique allows us to easily introduce mutations into E. coli cells.

Previously, we used the isogenic E. coli K-12 KEIO collection of the Genome Analysis Project in Japan, which contains all non-lethal deletion mutations (3985 genes), to introduce as many as four mutations in a single E. coli strain for directing cell metabolism from formate to hydrogen without diminishing cell growth.

E. coli produces hydrogen from formate by the formate hydrogen lyase system (FHL) that consists of hydrogenase 3 (encoded by *hycABCDEFGHI*) and formate

dehydrogenase-H (encoded by *fdhF*); these enzymes catalyze the reaction $HCOO^-$ + $H_2O \leftrightarrow H_2 + HCO_3^-$ (Fig. 1) and are probably used to help regulate internal pH and to regulate external pH by removing toxic formate.

The evolved hydrogen from the FHL is consumed by E. coli hydrogenase 1 (*hyaB* encodes the large subunit) and hydrogenase 2 (*hybC* encodes the large subunit). Hence, by deleting *hyaB* and *hybC*, hydrogen production should be enhanced, and we have found that a four mutant (BW25113 *hyaB hybC fdhF*/ pCA24N-*FdhF*) increases hydrogen production from formate by over eighteen (18) orders of magnitude and increases hydrogen production from acetate by over twenty four (24) orders of magnitude. The simple technique consisted of removing the *hyaB*, *hybC* and *fdhF* genes after each round of P1 transduction. Along with the ease of this process (each round of mutagenesis takes 2 days), the resulting deletion mutations are more stable for eliminating target genes compared with point mutations or frameshift mutations (reversion is far more difficult).

A novel chemochromic membrane to easily detect hydrogen produced by single colonies by colorimetric response by binding of hydrogen to a thin-film WO3 sensor is now available and should speed research in this area. Additionally, a facile method to measure hydrogenase activity based on formate consumption has been reported recently. Herein, we report on evolving *fdhF* for hydrogen production using epPCR, with this technique we found one mutant with one amino acid change that increases hydrogen production from acetate by over twelve (12) orders of magnitude; and using saturation mutagenesis, we found one mutant that increases hydrogen production by over six (6) orders of magnitude.

1. THEORY

1.1 FERMENTATION VERSUS PHOTOSYNTHESIS FOR HYDROGEN

Hydrogen, the smallest biological substrate, has great potential as an alternative to limited fossil fuel resources. In addition to its higher energy content than fossil fuels, it is renewable if it is derived from renewable feedstocks, and the product of hydrogen oxidation is water; hence, the impact of hydrogen on the environment is not relevant.

Production of hydrogen by microorganisms is at ambient temperature and pressure; hence, it requires less energy compared with conventional thermal systems (steam methane reforming, 850°C, 25 bar) and compared with electrolytic processes. Microorganisms produce hydrogen via two main pathways: photosynthesis and fermentation. Photosynthesis is a light-dependent process, including direct biophotolysis, indirect biophotolysis and photo-fermentation, whereas, anaerobic fermentation, also known as dark fermentation, is a light-independent process. Photosynthetic hydrogen production is performed by photosynthetic microorganisms, such as algae, photosynthetic bacteria and cyanobacteria. Fermentative hydrogen production is conducted by fermentative microorganisms, such as strict anaerobes [*Clostridium* strains, thermophiles, rumen bacteria and methanogens], facultative anaerobes [*Enterobacter* strains, *Escherichia coli* and *Citrobacter* species], or mixed cultures.

Compared with photosynthetic processes, fermentative hydrogen production generally yields two orders of magnitude higher rates, does not rely on the availability of light, utilizes a variety of carbon sources such as organic compounds, low-cost wastes, or insoluble cellulosic and cellobiose substrates, requires less energy, and is technically much simpler and more stable. Although hydrogen

production yields are usually higher with photosynthetic processes, oxygen is evolved during photosynthesis which inhibits the hydrogenase enzyme which is responsible for H_2 production. In addition, fermentative microorganisms have rapid growth and are not affected by oxygen as much as the main process is anaerobic (any residual oxygen is rapidly consumed at the onset). Therefore, fermentative hydrogen production is more advantageous than the photosynthetic hydrogen production and appears to have more potential for practical applications.

1.2 FORMATE HYDROGEN LYASE SYSTEM

The formate hydrogen lyase (FHL) system is a multienzyme complex responsible for molecular hydrogen production from formate. Formate hydrogen lyase activity has been found in various bacteria such as *Salmonella typhimurium*, *Klebsiella pneumoniae*, *Rhodospirillum rubrum*, *Methanobacterium formicicum*, *E. coli*, and many other coli-aerogenes bacteria; however, the FHL complex of *E. coli* is the most studied and was discovered in 1931. The *E. coli* genome encodes four nickel-iron hydrogenases: hydrogenase-1 (Hyd-1), hydrogenase-2 (Hyd-2), hydrogenase-3 (Hyd-3) and hydrogenase-4 (Hyd-4). From these four hydrogenases, Hyd-3 is a part of the active anaerobic FHL complex and is encoded by the *hyc* operon (Fig. 2, Table 1).

Hyd-1 and Hyd-2 are known as uptake hydrogenases which catalyse hydrogen oxidation and are encoded by the *hya* and *hyb* operons. Hyd-3 was first thought to only have hydrogen-producing activity; however, recent studies show that Hyd-3 has hydrogen uptake activity as well. Hyd-4, encoded by the *hyf* operon, has high homology with the *hyc* operon and was first proposed to possess a second FHL complex. However, some studies show that the Hyd-4 did not replace the Hyd-3 in hydrogen production and the *hyf* operon is not expressed in *E. coli*, but can be activated in the presence of effector-independent FhlA (transcriptional activator of the FHL complex) mutant proteins or HyfR, which is an FhlA homologue for Hyd-4.

The FHL complex of *E. coli* is responsible for the conversion of formate to CO_2 and H_2 under anaerobic conditions and in the absence of electron acceptors such as oxygen and nitrate. The active complex consists of seven proteins, six from the *hyc* operon (HycBCDEFG) and formate dehydrogenase H (Fdh-H encoded by *fdhF*). In addition, the FHL system is controlled by a transcriptional activator FhlA, which is required for the transcription of *fdhF* and the *hyc* operon, and a negative transcriptional regulator, HycA. Transcription of the FHL complex is dependent on the presence of formate, acidic pH and the σ^{54} factor. Fdh-H, encoded by the *fdhF* gene, is a 79 kDa cytoplasmic protein and contains selenocysteine, molybdenum and a [4Fe-4S] cluster at its active site.

1.3 THEORETICAL YIELD OF HYDROGEN FROM GLUCOSE

The hydrogen yield is defined as the moles of H_2 produced per mole of substrate. Carbohydrates, mainly glucose, are the preferred substrates for fermentative hydrogen production. Starch, cellulose, as well as organic wastes can also be used as substrates. Many microorganisms such as *Enterobacter* sp., *Clostridium* sp. and *E. coli* are capable of producing hydrogen.

The fermentative route of hydrogen production starts with the conversion of glucose to pyruvate and NADH through glycolysis in both strict and facultative anaerobic bacteria. In facultative anaerobes, such as *E. coli* (Fig. 3), pyruvate is then converted to acetyl-CoA and formate, which is catalysed by pyruvate formate lyase (PFL). Hydrogen is produced from formate by the FHL complex. Because a maximum of two molecules of formate are produced from two pyruvate molecules, facultative anaerobic bacteria have a theoretical maximum yield of 2 mol of H_2 per mole of glucose. There are several factors that influence the yield, such as whether some of the pyruvate is converted to lactate; these competing paths for obtaining reducing power from pyruvate lower the yield.

1.4 GROWTH CELLS CONDITIONS

All living creatures need a suitable medium for growth: also, the bacteria need some nutrients to grow. First, we will present the difference between the macronutrients and the micronutrients, and then, we will present the medium using during the development of the investigation.

1.4.1 MACRONUTRIENTS

The macronutrients are defined as the chemical elements that humans, plants or bacteria consume in the largest quantities such as carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur.

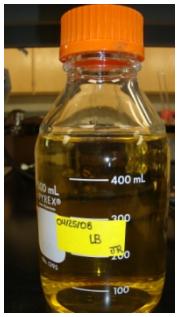
1.4.2 MICRONUTRIENTS

The micronutrients are nutrients needed for life in small quantities such as iron, cobalt, chromium, copper, iodine, manganese, selenium, zinc and molybdenum. They are dietary minerals needed by the human body, plant or bacteria in very small quantities (generally less than 100 mg/day).

1.4.3 LURIA BERTANI (LB)

Luria-Bertani medium is a liquid media that is used in the growth of bacteria. This medium is made using the following reagents per liter of dionized water: 10 g Bactotryptone, 5 g Bacto Yeast Extract and 10 g NaCl. (Fig. 4)

Fig. 4. Luria-Bertani (LB) Medium



1.4.4 COMPLEX MEDIA

This medium is rich in a variety of nutrients that gives to the cells certain characteristics in the fermentation process, or serves as a cofactor of several enzymes for its development; however, this kind of medium limits the growth of the cells. This was the medium that we used for the measurements of the H_2 production. This medium is made using the following reagents per liter of dionized water: 5 g Bacto Yeast Extract, 5 g Bacto Tryptone, 7 g K_2HPO_4, 5.5 g KH_2PO_4, 0.5 g L-Cysteine, 1 g $(NH_4)_2SO_4$, 0.25 g $MgSO_4 \cdot 7H_2O$, 0.021 g $CaCl_2 \cdot 2H_2O$, 0.029 g $Co(NO_3)_2 \cdot 6H_2O$, 0.039 g $Fe(NH_4)_2SO_4 \cdot 6H_2O$, 2 mg (1 mL 1000x) Nicotonic Acid, 0.172 mg (0.1 mL 10000x) Na_2SeO_3, 0.02 mg (0.1 mL 10000 x) $NiCl_2$, 10 mL Trace Solution and 0.1 mL 10000x $(NH_4)_6Mo_7O_{24}$. Also this medium contains an acid, in our case, we use formic acid, acetic acid and lactic acid for the assays, in each case we add: 6.8 g Sodium formate, 8.2 g Sodium acetate or 18.7 g Sodium lactate. (Fig. 5)

Fig. 5. Complex Media

Trace Solution contains: 0.5 g $MnCl_2 \cdot 4H_2O$, 0.1 g H_3BO_3, 0.01 g $AlK(SO_4)_2 \cdot H_2O$, 0.001 g $CuCl_2 \cdot 2H_2O$ and 0.5 g Na_2EDTA per liter.

1.5 P1-TRANSDUCTION

The P1-Transduction is a quick method for transferring a knock-out from one strain to another when there is an antibiotic resistance gene at the knock out site of

the original strain (so you can easily select for the new mutation in the new strain) and when the two strains are related enough to share homology of DNA around the knocked-out gene. We use the following plates to perform the experiment.

R-plate (bottom plate): 10 g Bacto Tryptone, 1 g Bacto Yeast Extract, 8 g Sodium Chloride and 10 g Agar, after autoclave, must add Calcium Chloride (final concentration: 2 mM) and glucose (final concentration: 0.1%) per liter.

R-top medium: 10 g Bacto Tryptone, 1 g Bacto Yeast Extract, 8 g Sodium Chloride and 8 g Agar, after autoclave, must add Calcium Chloride (final concentration: 2 mM) and glucose (final concentration: 0.1%) per liter, and should keep this at 50^{o}C).

1.6 POLYMERASE CHAIN REACTION (PCR)

PCR is used to amplify specific regions of a DNA strand (the DNA target). A basic PCR set up requires several components and reagents. These components include:

- DNA template that contains the DNA region (target) to be amplified.
- Two primers, which are complementary to the DNA regions at the 5' (five prime) or 3' (three prime) ends of the DNA region.
- Taq polymerase or another DNA polymerase with a temperature optimum at around 70°C.
- Deoxynucleoside triphosphates (dNTPs), the building blocks from which the DNA polymerases synthesizes a new DNA strand.

The PCR usually consists of a series of 20 to 40 repeated temperature changes called cycles; each cycle typically consists of 2-3 discrete temperature steps. Most

commonly PCR is carried out with cycles that have three temperature steps (Fig. 6). We will present and explain the steps of the polymerase chain.

1.6.1 INITIALIZATION STEP

This step consists of heating the reaction to a temperature of 94-96°C, which is held for 1-9 minutes. It is only required for DNA polymerases that require heat activation.

1.6.2 DENATURATION STEP

This step is the first regular cycling event and consists of heating the reaction to 94-98°C for 20-30 seconds. It causes melting of DNA template and primers by disrupting the hydrogen bonds between complementary bases of the DNA strands, yielding single strands of DNA.

1.6.3 ANNEALING STEP

The reaction temperature is lowered to 50-65°C for 20-40 seconds allowing annealing of the primers to the single-stranded DNA template. Typically the annealing temperature is about 3-5 degrees Celsius below the Tm of the primers used. Stable DNA-DNA hydrogen bonds are only formed when the primer sequence very closely matches the template sequence. The polymerase binds to the primer-template hybrid and begins DNA synthesis.

1.6.4 EXTENSION/ELONGATION STEP

The temperature at this step depends on the DNA polymerase used; Taq polymerase has its optimum activity temperature at 75-80°C, and commonly a temperature of 72°C is used with this enzyme. At this step the DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand by adding dNTPs that are complementary to the template in 5' to 3' direction, condensing the 5'-phosphate group of the dNTPs with the 3'-hydroxyl group at the

end of the nascent (extending) DNA strand. The extension time depends both on the DNA polymerase used and on the length of the DNA fragment to be amplified. As a rule-of-thumb, at its optimum temperature, the DNA polymerase will polymerize a thousand bases per minute.

1.6.5 FINAL ELONGATION

This single step is occasionally performed at a temperature of 70-74°C for 5-15 minutes after the last PCR cycle to ensure that any remaining single-stranded DNA is fully extended.

1.6.6 FINAL HOLD

This step at 4-15°C for an indefinite time may be employed for short-term storage of the reaction.

1.7 ERROR-PRONE PCR

The Error-Prone PCR is a technique used to generate randomized genomic libraries. The Error-Prone PCR allow us to initiate DNA amplification starting with tiny amounts of parent molecule to produce considerable amounts of the mutated gene. This technique is based on the principle that Taq polymerase is capable of annealing incompatible base-pairs to each other during amplification under imperfect PCR conditions. (Fig. 7)

After the implementation of Error-Prone PCR the mutant gene is in abundant concentration and thus has an extremely high likelihood of ligating into a viable plasmid. These random matching “errors” in the global transcription factors will presumably lead to desirable mutations ultimately causing increase in current output. Following the ligation step, the random variants will be screened and selected using novel chemochromic membrane.

1.8 SATURATION MUTAGENESIS

Saturation Mutagenesis is an in vitro mutagenesis strategy wherein one tries to generate all (or most) possible mutations within a narrow region of a gene. The two most popular ways to make the mutations are cassette mutagenesis, or oligo-directed mutagenesis. Both depend on synthesizing oligonucleotides covering the regions to be mutagenized such that the oligos contain a random assortment of incorrect bases. A third way is to conduct PCR under conditions that favor a high degree of misincorporation.

1.9 SDS-PAGE (SODIUM DODECYL SULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS)

The SDS-PAGE is a technique used to separate proteins according to their electrophoretic mobility (a function of length of polypeptide chain or molecular weight as well as higher order protein folding, posttranslational modifications and other factors). We use the following conditions for making the gels.

Table 2. Reagents used to prepare the SDS-PAGE Gels.

	Mini-Protein II (quantities for 2 gels)	
Reagent	12 % Resolving	5% Stacking
ddH_2O	3.3 mL	2.7 mL
Tris buffer	2.5 mL (1.5 M Tris-HCl pH 8.8)	0.5 mL (1.0 M Tris-HCl pH 6.8)
10% SDS	0.1 mL	40 μL
30% Acrylamid/Bis	4.0 mL	0.67 mL
10% APS (ammonium persulfate)	0.1 mL	40 μL
TEMED	4 μL	4 μL

2. EXPERIMENTAL SECTION

2.1 BACTERIAL STRAINS

Parent strain *E. coli* BW25113 *hyaB hybC* was obtained from previous studies in Wood's lab. *E. coli* BW25113 *hyaB hybC fdhF::kan* was constructed using P1-Transduction and used as the host for screening the *fdhF* variants since it lacks hydrogen production and hydrogen consumption.

The colonies we select for PCR analysis were analyze using forward primer mdtPforward 5'-CAACGGCACGCGTCTGCAAA-3' (targeting the *mdtP* gene which is upstream of *fdhF*) and reverse primer K1 3'-TCCGATAAGCCGATA CTGAC-5' (targeting the Kanamycin resistance gene).

2.2 GROWTH

E. coli strains were initially streaked from −80°C glycerol stocks on Luria–Bertani (LB) agar plates containing 30 μg/mL of CM (chloramphenicol) and incubated at 37°C.

2.3 ERROR-PRONE PCR

Plasmid pCA24N-*fdhF* was obtained from previous studies in Wood's lab and has *fdhF* under control of a constitutive lac promoter. To introduce random mutations into the whole *fdhF* locus, epPCR was performed using *pCA24N-fdhF* as the template with two primers (fhlA front 5′-CACCGATCGCCCTTCCCAACAGTTGC-3′ and fhlA rev 5′-CATCCGCTTACAGACAAGCTGTGACC-3′). The epPCR products were cloned in pCA24N, and the ligation mixture was electroporated into *E. coli* BW25113 *hyaB hybC fdhF*.

2.4 SATURATION MUTAGENESIS

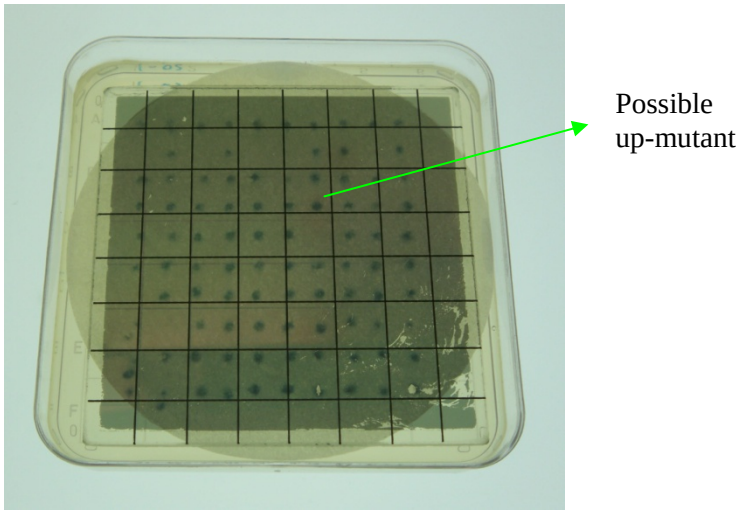
Saturation mutagenesis was performed at codon Y513 of *fdhF* using a QuikChange® XL Site-directed Mutagenesis Kit and DNA primers were designed to vary the codon to allow for substitution of all 3 amino acids. The 5′-ACT codon for *fdhF* Y513 was varied using primers with the variable NNS codon where N is A, G, C, or T and S is G or C: BfdhF1 5′-CGTCATCTGTGCCCGGATTTC**NNS**GGTGCGACTTACGAG-3′ and BfdhF2 5′-CTCGTAAGTCGCACC**NNS**GAAATCCGGGCACAGATGACG-3′. The constructed plasmids were electroporated into BW25113 *hyaB hybC fdhF::kan*, and 385 of the generated colonies were screened with the chemochromic membranes to find variants showing high hydrogen-producing activity.

2.5 SCREENING

Chemochromic membranes were used to identify fdhF variants with beneficial mutations; that is, colonies were chosen on the basis of enhanced hydrogen production. The colonies from the epPCR and saturation mutagenesis libraries were transferred to square agar plates (100×100×15 mm) containing modified complex-formate medium (100 mM formate), complex-acetate medium (100 mM acetate) and complex-lactate medium (100 mM lactate).

The plates were incubated anaerobically at 37°C for 14 h using a Gas-Pak anaerobic system. In the presence of oxygen, Whatman filter paper was placed firmly on top of the colonies on each plate, and the glass plates coated with the chemochromic membrane were placed on top of the Whatman paper. Colonies showing deep blue were chosen as candidates. The negative control, *BW25113 hyaB hybC fdhF* did not produce hydrogen and remained colorless on the membrane, we use *BW25113 hyaB hybC fdhF*/pCA24N-*fdhF* as positive control. Five replicates were checked for each candidate before proceeding to a gas chromatography (GC)-based hydrogen assay. (Fig. 8)

Fig. 8. Screening of *E. coli* variants using chemochromic membranes.



2.6 CLOSED HYDROGEN ASSAY

Overnight aerobic modified complex-formate, modified complex-acetate or modified complex-lactate medium with CM (chloramphenicol) (30 µg/mL) cultures (25 mL) and fresh modified complex-formate, modified complex-acetate or modified complex-lactate medium were sparged for 5 min with nitrogen to remove oxygen. Sealed crimp-top vials (27 mL) were sparged for 5 min with nitrogen, and 0.5 mL of the cell suspension and 9.5 mL of fresh medium were added to the bottles inside the nitrogen hood (Fig. 9), which were incubated at 37°C with shaking for 0.5 to 5 h (Fig. 10). The amount of hydrogen generated in the head space of the recombinant system was measured using a 50-µL aliquot by GC using a 6890N gas chromatograph (Fig. 11 and Fig. 12). Add the end of the experiment we measure OD_{600} (Fig. 13). Total proteins are calculated from 0.22 mg/OD/mL.

Total hydrogen in head sample

$$H_2\ prod. = peak\ area \times \frac{0.224 * 10^{-3}\ \mu mol}{peak\ area} \times \frac{17000\ \mu L\ (head\ space)}{50\ \mu L\ (injection)} \times \frac{1}{10\ mL\ (culture) \times OD} \times \frac{1}{0.22\ mg\ protein / OD / mL}$$

Fig. 9. Nitrogen Hood, which is anaerobic.

Fig. 10. Sealed Teflon Caps Vials.

Fig. 11. Gas Chromatograph (GC).

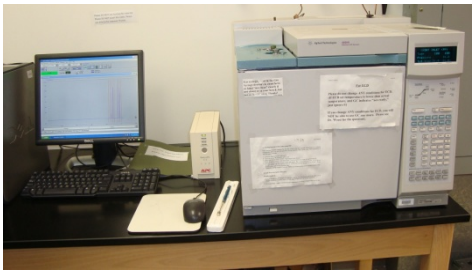


Fig. 12. Inyection of the sample into the Gas Chromatograph (GC).

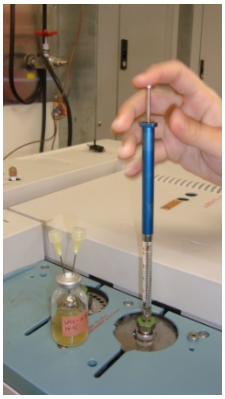


Fig. 13. Spectrophotometer to measure the $O.D._{600}$.

2.7 SDS-PAGE

Expression of recombinant proteins was analyzed with standard Laemmli discontinuous sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis technique. The volume loaded (V_L) of the sample, is calculating using the next equation, where V_c is the volume of the culture that was centrifuge and V_s is the final volume of the sample.

$$\frac{0.25\ AU}{well} \cdot \frac{1\ mL\ cells}{OD_{600}} \cdot \frac{well\ V_s\ mL}{V_T\ mL} = V_C\ \mu L$$

2.8 DNA SEQUENCING

A dideoxy chain termination technique with the ABI™ Prism BigDye Terminator Cycle Sequencing Ready Kit was used to determine the *fdhF* nucleotide sequences.

2.9 MODELING

Fdh-H was modeled using the Swiss Model Software and was visualized using the Pymol Software, based on the formate dehydrogenase H Fdh-H derived from Desulfovibrio gigas.

3. RESULTS

In this study, BW25113 *hyaB hybC fdhF* was used as the host for cloning since it produces minimal hydrogen due to inactivating hydrogen consumption by hydrogenases 1 and 2, and the formate hydrogen lyase (FHL) encoded by the *fdhF* gene.

Our goal was to engineer for hydrogen synthesis using epPCR and to identify beneficial mutations as there have been no structure/function studies for this enzyme beyond studies concerning the selenocysteine (SeCys), molybdenum, two molybdopterin guanine dinucleotide (MGD) cofactors, and an Fe4S4 cluster at the active site. Plasmid pCA24N-*fdhF*, which can complement the chromosomal *fdhF* mutation, was used for protein engineering of *fdhF*.

3.1 P1-TRANSDUCTION

We picked 8 colonies to streak onto LB containing 50 μg/mL of Kan (kanamycin) plates for further PCR analysis. Only three of them grew well on LB/Kan50 agar and broth. The three colonies were chosen for PCR analysis, after this procedure we find that only one of the three (colony No. 3) has the correct recombination (Fig. 14).

3.2 ERROR-PRONE PCR

To introduce random mutations into *fdhF*, epPCR was performed. From sequencing ten colonies, the maximum error rate was 0.53%. Ten thousand three hundred and seventy five (10375) colonies with ep- *fdhF* alleles were screened using the chemochromic sensor method, which resulted in the identification of one variant with beneficial mutations (Table 3) that showed high hydrogen-producing activity compared to BW25113 *hyaB hybC fdhF* / pCA24N-*fdhF* when modified complex acetate 100 mM is using as the carbon source. To confirm that the

phenotype was due to plasmid-based *fdhF*, the plasmids were isolated and re-electroporated into the original BW25113 *hyaB hybC fdhF* host.

Table 3. Comparison of hydrogen production rates measured in this study. The area shaded blue represents relative rate.

Strain	Production rate average (µmol/mg-protein/h) and relative					
	0.5 h	1 h	2 h	3 h	4 h	5 h
ep-fdhF 7- 8	0.21	0.07	0.83	2.11	1.3	1.29
	0.77	0.47	6.94	12.65	4.22	0.64
BW25113 *hyaB hybC fdhF*/pCA24N-fdhF	0.27	0.16	0.12	0.17	0.31	2
	1	1	1	1	1	1

3.3 CLOSED HYDROGEN ASSAY

Hydrogen production with the one plasmid was assayed using a closed hydrogen assay. The *fdhF* variant ep-fdhF 7-8 had twelvefold higher hydrogen production than BW25113 hyaB hybC *fdhF* / pCA24N- *fdhF* (Table 3). This increase in hydrogen production was due to one amino acid change (Fig. 15).

3.4 SATURATION MUTAGENESIS

The *fdhF* variant ep-fdhF 7-8 had one amino acid change at position Y513H (tyrosine is replaced for histidine), we investigated the importance of this position of *fdhF* for enhanced hydrogen production by substituting all possible amino acids at this position via saturation mutagenesis. After screening 385 colonies to ensure with a probability of 99.999% that all possible codons were utilized, two variants (SMA 1-24, this strain had sixfold higher hydrogen production; and SMA 3-58, this strain had fivefold higher hydrogen production) were identified that had elevated hydrogen producing activity compared to the ep-fdhF 7-8 variant using the chemochromic membranes and the hydrogen production was assayed using the closed hydrogen assay (Table 4).

Table 4. Comparison of hydrogen production rates measured in this study. The area shaded blue represents relative rate.

Strain	Production rate average (µmol/mg-protein/h) and relative					
	0.5 h	1 h	2 h	3 h	4 h	5 h
SMA 1-24	0.33	0.21	2.66	4.43	4.16	3.76
	0.9	0.82	6.03	2.36	1.83	1.7
SMA 3-58	0.28	0.17	1.61	9.81	10.81	8.79
	0.75	0.66	3.65	5.22	4.76	3.96
7 – 8	0.37	0.26	0.44	1.88	2.27	2.22
	1	1	1	1	1	1

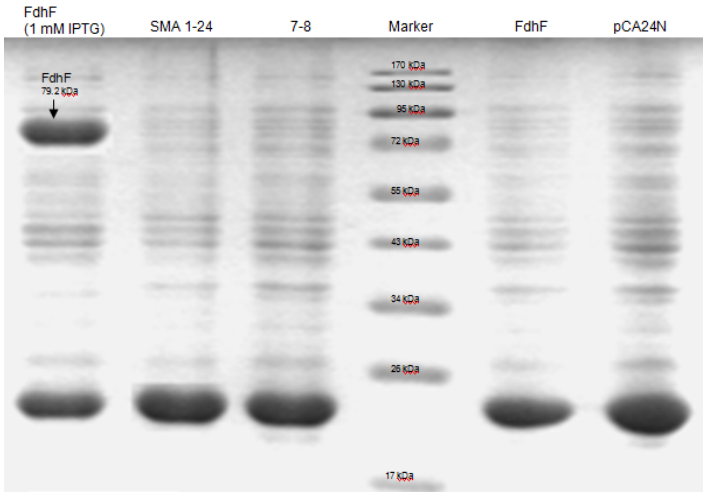
3.5 MODELING

DNA sequencing revealed that the two mutants had the same amino acid change H513Q (histidine is replaced for glutamine) (Fig. 16 and Fig 17). After re-electroporating the plasmid into the original host BW25113 *hyaB hybC fdhF* to confirm the phenotype was due to plasmid based fdhF.

3.6 SDS-PAGE

The fdhF gene of the saturation mutagenesis variant was verified using SDS-PAGE (Fig. 18).

Fig. 18. SDS-PAGE gel of *fdhF* variants.



CONCLUSIONS AND RECOMMENDATIONS

In this study, we constructed two *fdhF* variants that can produce up to 12-fold higher hydrogen than BW25113 *hyaB hybC fdhF* / pCA24N-*fdhF* through ep-PCR and saturation mutagenesis. Notably, the hydrogen yield was increased by an order of magnitude to become nearly equal to the theoretical maximum of 1 mol H_2/mol formate but using acetate as the carbon source.

These results may be improved by eliminating formate dehydrogenase-N and formate dehydrogenase-O.

From the two *fdhF* variants (one epPCR variant and one saturation mutagenesis variant), there are no mutations from amino acid positions 1 to 512 and 514 to 715; hence, these regions appear important for hydrogenase activity. Furthermore, ep-fdhF 7-8 and SMA 1-24 have amino acid change at the same position (513) indicating that the amino acid change Y513H and H513Q, respectively, may be important for improved activity. Therefore, this position is significant for enhanced hydrogen production and was a good target for saturation mutagenesis.

Since metabolic engineering (rational pathway engineering approach) has been used to enhance hydrogen production with E. coli, the *fdhF* variants described here using a random approach may be combined with these systems to increase hydrogen production further. In addition, since all of the metabolic pathways in E. coli are not fully elucidated, other random technologies might be performed including classical chemical mutagenesis followed by genome breeding and DNA microarray analysis to discern where the random chromosomal mutations lie. Hence, using E. coli as a reference system for producing hydrogen has many advantages.

The hydrogen production in E. coli changes with time and differs with nutrients employees during the assays, we did not find any good mutant when modified complex formate and modified complex lactate was using as carbon source, also, as a sub-product of the fermentation process we get small or large quantities of organic acids, such as maleic acid, lactate, butyrate, propionate and acetate, which acidify the medium and inhibit the growth rate and the hydrogen production, therefore, if we can control the pH of the medium, we should increase the yield.

REFERENCES

Axley MJ, Grahame DA, Stadtman TC (1990). Escherichia coli formate-hydrogen lyase. Purification and properties of the selenium-dependent formate dehydrogenase component. J Biol Chem 265:18213–18218

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008

Bagramyan K, Trchounian A (2003). Structural and functional features of formate hydrogen lyase, an enzyme of mixed-acid fermentation from Escherichia coli. Biochemistry (Mosc) 68: 1159–1170

Bisaillon A, Turcot J, Hallenbeck PC (2006). The effect of nutrient limitations on hydrogen production by batch cultures of Escherichia coli. Int J Hydrogen Energy 31:1504–1508

Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997). The complete genome sequence of Escherichia coli K-12. Science 277:1453–1474

Boyington Jeffrey, Gladyshev Vadim, Khangulov Sergei, Stadtman Thressa, Sun Peter. Crystal Structure of Formate Dehydrogenase H: Catalysis Involving

Mo, Molybdopterin, Selenocysteine, and an Fe_4S_4 Cluster. *Science* **275**, 1305 (1997); DOI: 10.1126/science.275.5304.1305

Bujard Hermann, Gentz Reiner, Lanzer Michael, Stueber Dietrich, Mueller Michael, Ibrahimi Ibrahim, Haeuptle Marie Therese, Dobberstein Bernhard. A T5 promoter -based transcription-translation system for the analysis of proteins in vitro and in vivo. Methods in Enzymology (1987), 155 (Recomb. DNA, Pt. F), 416-33.

Gokarn R. R., Eiteman M. A., Altman E. Metabolic Analysis of *Escherichia coli* in the Presence and Absence of the Carboxylating Enzymes Phosphoenolpyruvate Carboxylase and Pyruvate Carboxylase. Applied and Environmental Microbiology, May 2000, p. 1844–1850, Vol. 66, No. 5

Maeda Toshinari, Sanchez-Torres Viviana, Wood Thomas. Biotechnologically relevant enzymes and proteins. *Escherichia Coli* hydrogenase 3 is a reversible enzyme possessing hydrogen uptake and sintesis activities.

Maeda Toshinari, Sanchez-Torres Viviana, Wood Thomas. Protein engineering of hydrogenase 3 to enhance hydrogen production. Appl Microbiol Biotechnol (2008) 79:77–86. DOI 10.1007/s00253-008-1416-3

Mark Redwood, Iryna Mikheenko, Frank Sargent, Lynne Macaskie. Dissecting the roles of Escherichia coli hydrogenases in biohydrogen production. Federation Of European Microbiological Sciences. DOI:10.1111/j.1574-6968.2007.00966.x.

Mahyudin A. Rachman, Yoshinori Furutani, Yutaka Nakashimada, Toshihide Kakizono and Naomichi Nishio. Journal of Fermentation and Bioengineering. Vol. **83,** No. 4, 358-363. 1991

Neish, A.C. (1952). Analytical Methods for Bacterial Fermentations. Report 46-8-3, 2nd revision. Nat. Research Council of Can., Saskatoon

P.C. Hallenbeck. Fundamentals of the fermentative production of hydrogen. Water Science & Technology Vol 52 No 1-2 pp 21–29

QuikChange® XL Site-Directed Mutagenesis Kit. Instruction manual.

Sambrook J., Fritsch E. F., Maniatis T. (1989). Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sang Jun Lee, Dong-Yup Lee, Tae Yong Kim, Byung Hun Kim, Jinwon Lee, and Sang Yup Lee. Metabolic Engineering of *Escherichia coli* for Enhanced Production of Succinic Acid, Based on Genome Comparison and In Silico Gene Knockout Simulation. Applied and Environmental Microbiology, Dec. 2005, p. 7880–7887, Vol. 71 No. 12. DOI:10.1128/AEM.71.12.7880–7887.2005

Sanger F., Nicklen S., Coulson A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sambrook Joseph, Russel David. (2001). Molecular Cloning: A Laboratory Manual (3rd ed.). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. ISBN 0-87969-576-5.

Self, W. T., A. Hasona, and K. T. Shanmugam. (2001). N-terminal truncations in the FhlA protein result in formate- and MoeA-independent expression of the *hyc* (formate hydrogenlyase) operon of *Escherichia coli*. Microbiol **147.**

Sokatch, J. R. Bacterial Physiology and Metabolism. London: Academic Press Inc., 1969.

Yoshida A., Nishiamura T., Kawaguechi H., Inui M., Yukawa H. (2005). Enhance Hydrogen Production from Formic Acid by Formate Hydrogen Lyase – Overexpressing *Escherichia Coli* strains. Appl Microbiol Biotechnol **71**: 6762-6768.

ATTACHMENTS

Fig. 1. Schematic of fermentative hydrogen production in *Escherichia coli*. Hydrogen is produced from formate by the formate hydrogen lyase (FHL) system [hydrogenase 3 and formate dehydrogenase-H (FDHH)], which is activated by FhlA (that is regulated by Fnr) and repressed by HycA. Evolved hydrogen is consumed through the hydrogen uptake activity of hydrogenase 1 and hydrogenase 2. Formate is exported by FocA and/or FocB and is metabolized by formate dehydrogenase-N (FDHN) which is linked with nitrate reductase A and formate dehydrogenase-O (FDHO). Cyanobacterial hydrogenases (HoxEFUYH) derived from *Synechocystis* sp. PCC 6803 inhibit the activity of *E. coli* hydrogenase 1 and hydrogenase 2 resulting in enhanced hydrogen yield.

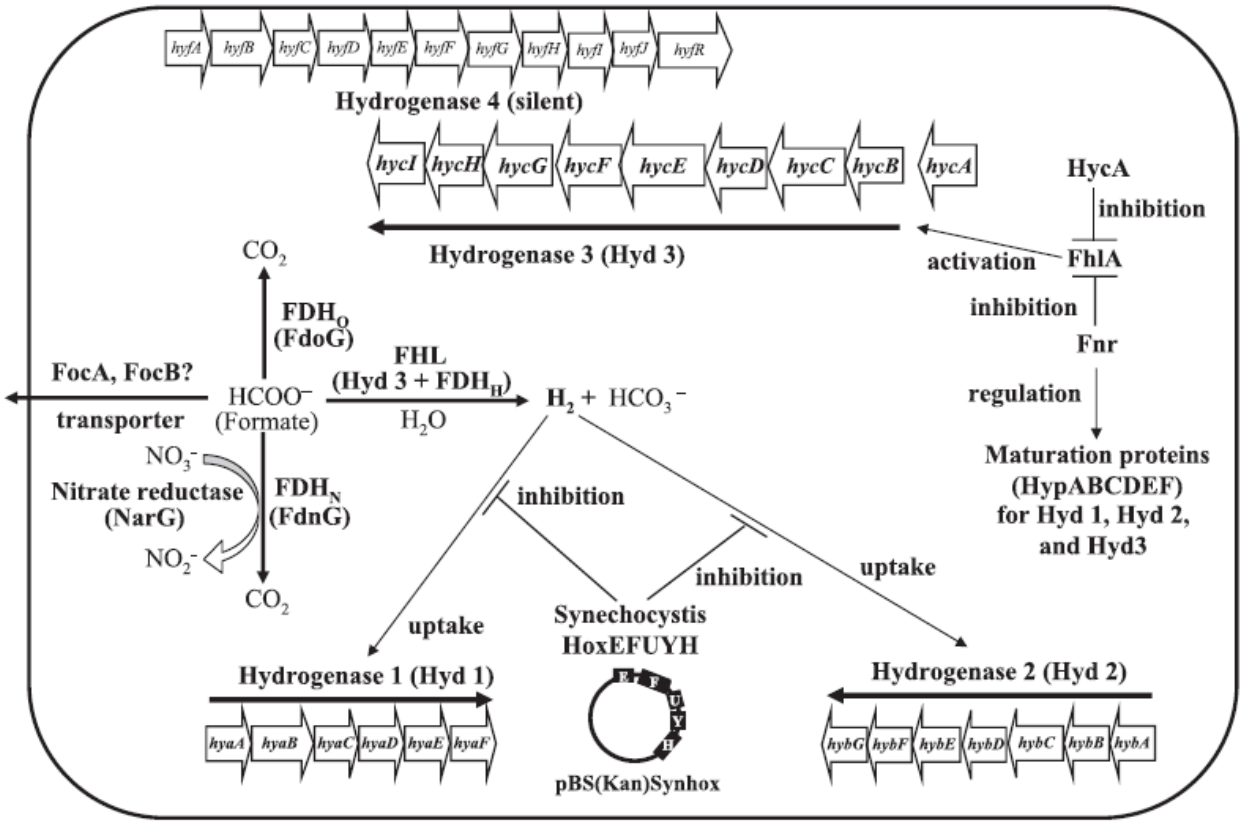


Fig. 2. A. Location of the four structural hydrogenase operons on the *E. coli* K12 chromosome (AC000091). The values in brackets signify the locations of the respective genes on the genome map.

B. Organization of the genes of each hydrogenase operon in *E. coli.* Arrows indicate the direction of transcription.

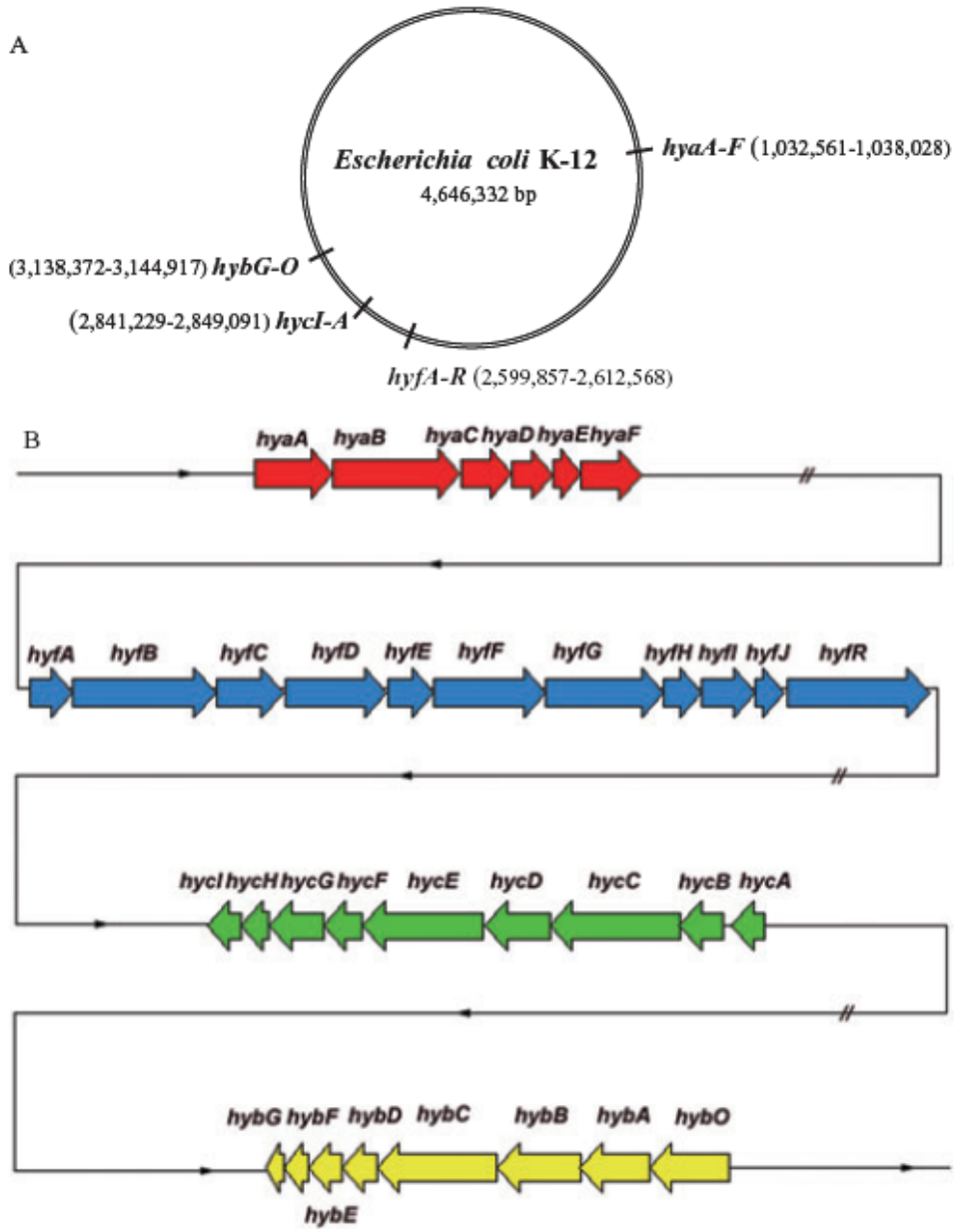


Table 1. Genes of four hydrogenase operons (*hya*, *hyb*, *hyc*, *hyf*) in *E. coli*.

Gene	Size, bp	Description
hyaA	372	Hydrogenase 1, small subunit
hyaB	597	Hydrogenase 1, large subunit
hyaC	235	Hydrogenase 1, *b*-type cytochrome subunit
hyaD	195	Protein involved in processing of HyaA and HyaB proteins
hyaE	132	Protein involved in processing of HyaA and HyaB proteins
hyaF	285	Protein involved in nickel incorporation into hydrogenase-1 proteins
hyfA	205	Hydrogenase 4, 4Fe-4S subunit
hyfB	672	Hydrogenase 4, membrane subunit
hyfC	315	Hydrogenase 4, membrane subunit
hyfD	479	Hydrogenase 4, membrane subunit
hyfE	216	Hydrogenase 4, membrane subunit
hyfF	526	Hydrogenase 4, membrane subunit
hyfG	555	Hydrogenase 4, subunit
hyfH	181	Hydrogenase 4, Fe-S subunit
hyfI	252	Hydrogenase 4, Fe-S subunit
hyfJ	137	Predicted processing element hydrogenase
hyfR	670	DNA-binding transcriptional activator, formate sensing
hycI	156	Protease involved in processing C-terminal end of HycE
hycH	136	Protein required for maturation of hydrogenase 3
hycG	255	Hydrogenase 3, small subunit
hycF	180	Formate hydrogenlyase complex Fe-S protein
hycE	569	Hydrogenase 3, large subunit
hycD	307	Hydrogenase 3, membrane subunit
hycC	608	Hydrogenase 3, membrane subunit
hycB	203	Hydrogenase 3, Fe-S subunit
hycA	153	Regulator of the transcriptional regulator FhlA
hybG	82	Hydrogenase 2 accessory protein
hybF	113	Protein involved with the maturation of hydrogenases 1 and 2
hybE	162	Hydrogenase 2-specific chaperone
hybD	164	Predicted maturation element for hydrogenase 2
hybC	567	Hydrogenase 2, large subunit
hybB	392	Predicted hydrogenase 2 cytochrome *b*-type component
hybA	328	Hydrogenase 2 4Fe-4S ferredoxin-type component
hybO	372	Hydrogenase 2, small subunit

Fig. 3. Scheme of hydrogen production in E. Coli via fermentation. Cells metabolize glucose into Phosphoenolpyruvate, Pyruvate, and Formate. Phosphoenolpyruvate is converted to Succinate by Fumarate Reductase (FrdC), and Pyruvate is converted to either Lactate by Lactate Dehydrogenase (LdhA), to Carbon Dioxide (CO_2) and Acetate by Pyruvate Oxidase (PoxB), to Carbon Dioxide by Pyruvate Dehydrogenase (AceE), or to Formate by Pyruvate Formate Lyase (PFL). Hydrogen is produced from Formate by the Formate Hydrogen Lyase (FHL) system consisting of Hydrogenase 3 (Hyd 3) and Formate Dehydrogenase-H (FDHH). Evolved hydrogen is consumed through the hydrogen uptake activity of Hydrogenase 1 (Hyd 1) and Hydrogenase 2 (Hyd 2). Formate is exported by FocA and/or FocB and is metabolized by Formate Dehydrogenase-N (FDHN; FdnG), which is linked with nitrate reductase A (NarG) and Formate Dehydrogenase-O (FDHO; FdoG).

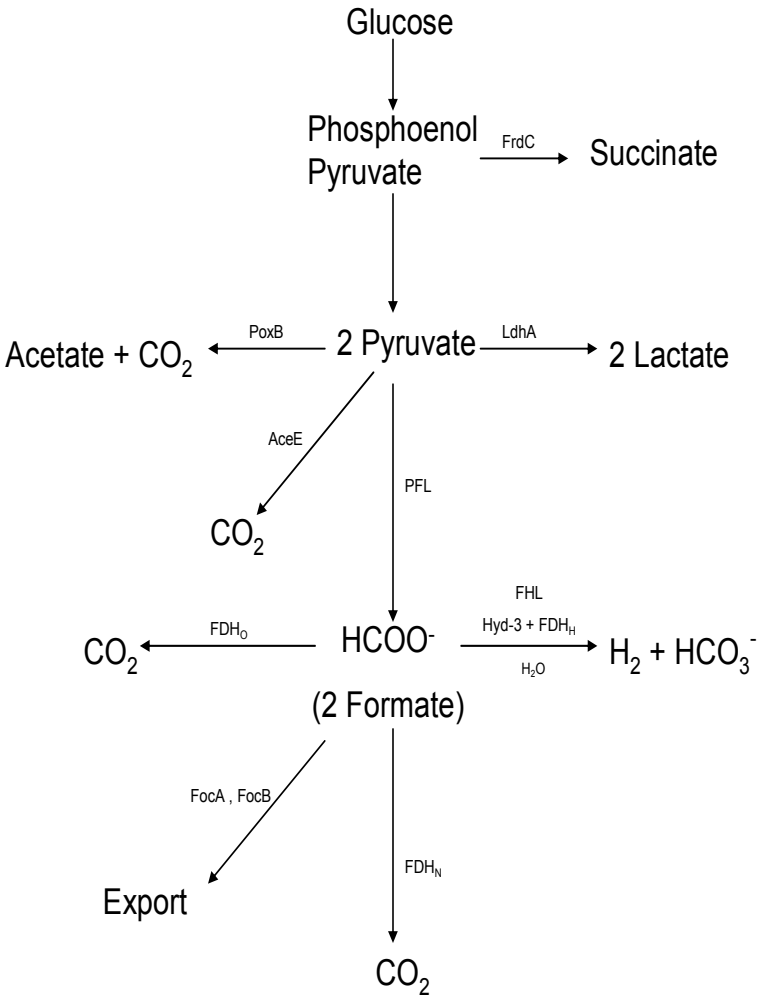


Fig. 6. Schematic drawing of the PCR cycle. (1) Denaturing at 94-96°C. (2) Annealing at ~65°C (3) Elongation at 72°C. Four cycles are shown here. The blue lines represent the DNA template to which primers (red arrows) anneal that are extended by the DNA polymerase (light green circles), to give shorter DNA products (green lines), which themselves are used as templates as PCR progresses.

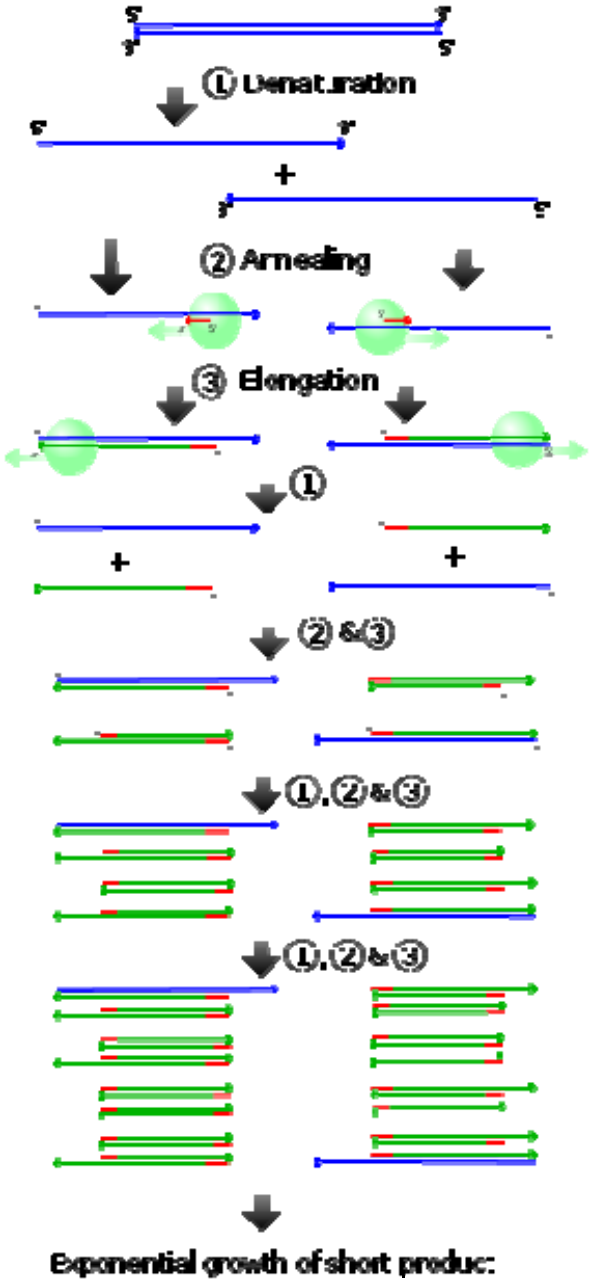


Fig. 7. Contrasts of the external conditions between PCR and Error-Prone PCR.

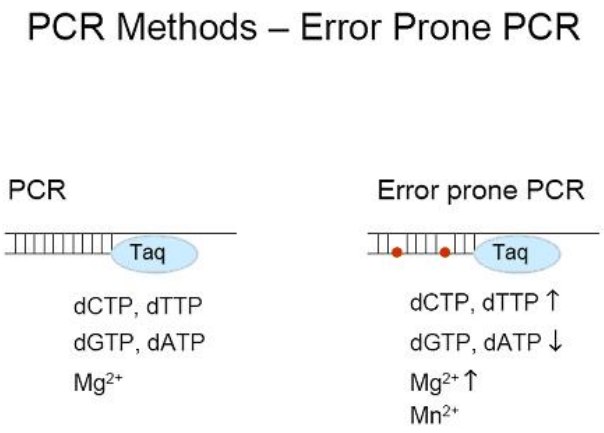


Fig. 14. Image of 0.8% agarose gel electrophoresis M: 1 Kb DNA ladder; 1 : colony 1; 2 : colony 2; 3 : colony 3; P : positive control (BW25113 ΔfdhF); N : negative control (BW25113 hyaB hybC Δkan)

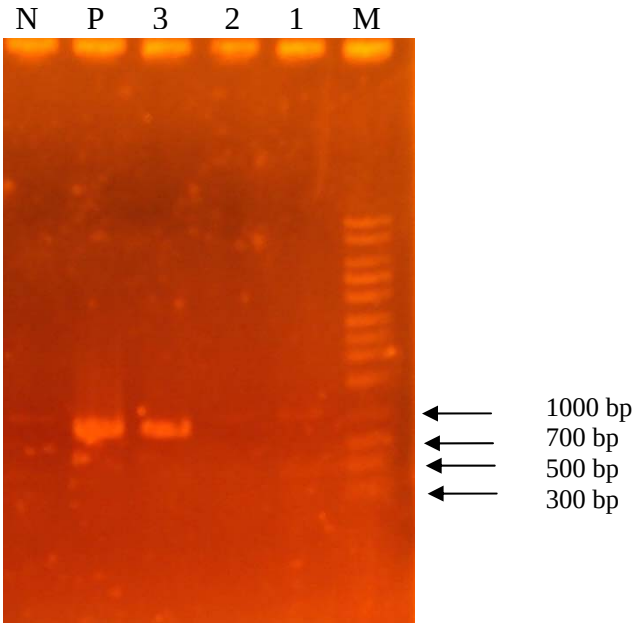


Fig. 15. Representation of Tyrosine (blue color) and Histidine (in purple) at position 513 of BW25113 *hyaB hybC fdhF* / pCA24N – *fdhF.* The sphere represents the molybdenum at the active site.

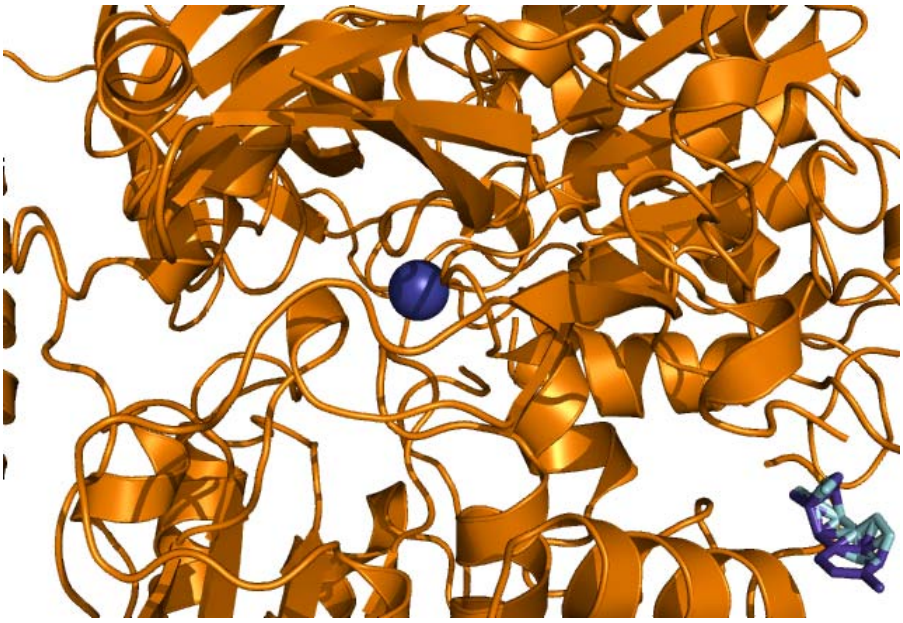


Fig. 16. Representation of Histidine (in purple) and Glutamine (red color) at position 513 of BW25113 *hyaB hybC fdhF* / pCA24N – *fdhF.* The sphere represents the molybdenum at the active site.

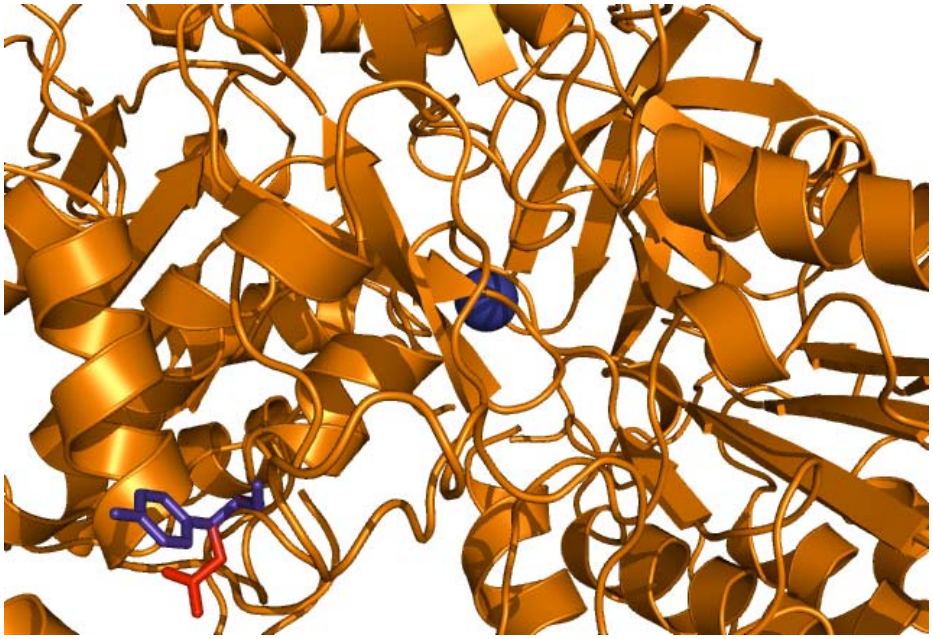


Fig. 17. Representation of Tyrosine (blue color), Histidine (in purple) and Glutamine (red color) at position 513 of BW25113 *hyaB hybC fdhF* / pCA24N – *fdhF.* The sphere represents the molybdenum at the active site.