Appl Microbiol Biotechnol (2008) 79:77–86
DOI 10.1007/s00253-008-1416-3

BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Protein engineering of hydrogenase 3 to enhance hydrogen production

Toshinari Maeda · Viviana Sanchez-Torres · Thomas K. Wood

Received: 19 January 2008 / Revised: 7 February 2008 / Accepted: 12 February 2008 / Published online: 12 March 2008

Abstract The large subunit (HycE, 569 amino acids) of *Escherichia coli* hydrogenase 3 produces hydrogen from formate via its Ni–Fe-binding site. In this paper, we engineered HycE for enhanced hydrogen production by an error-prone polymerase chain reaction (epPCR) using a host that lacked hydrogenase activity via the *hyaB hybC hycE* mutations. Seven enhanced HycE variants were obtained with a novel chemochromic membrane screen that directly detected hydrogen from individual colonies. The best epPCR variant contained eight mutations (S2T, Y50F, I171T, A291V, T366S, V433L, M444I, and L523Q) and had 17-fold higher hydrogen-producing activity than wild-type HycE. In addition, this variant had eightfold higher hydrogen yield from formate compared to wild-type HycE. Deoxyribonucleic acid shuffling using the three most-active HycE variants created a variant that has 23-fold higher hydrogen production and ninefold higher yield on formate due to a 74-amino acid carboxy-terminal truncation. Saturation mutagenesis at T366 of HycE also led to increased hydrogen production via a truncation at this position; hence, 204 amino acids at the carboxy terminus may be deleted to increase hydrogen production by 30-fold. This is the first random protein engineering of a hydrogenase.

Keywords Protein engineering · Hydrogenase 3 · Error-prone PCR · DNA shuffling · Saturation mutagenesis

T. Maeda · V. Sanchez-Torres · T. K. Wood (✉)
Artie McFerrin Department of Chemical Engineering,
Texas A & M University,
200 Jack E. Brown Building,
College Station, TX 77843-3122, USA
e-mail: thomas.wood@chemail.tamu.edu

T. K. Wood
Department of Biology, Texas A & M University,
College Station, TX 77843-3258, USA

T. K. Wood
Zachry Department of Civil and Environmental Engineering,
Texas A & M University,
College Station, TX 77843-3136, USA

Introduction

Fermentative hydrogen production has much potential as a renewable energy source (Das and Veziroglu 2001), and *Escherichia coli* is amenable to genetic manipulation (Blattner et al. 1997). *E. coli* produces hydrogen from formate by the formate hydrogen lyase (FHL) complex that consists of formate dehydrogenase-H (encoded by *fdhF* [Axley et al. 1990]) for forming $2H^+$, $2e^-$, and CO_2 from formate and hydrogenase 3 (encoded by *hycABCDEFGHI* [Bagramyan and Trchounian 2003; Sauter et al. 1992]) for synthesizing hydrogen from $2H^+$ and $2e^-$ (Sawers et al. 1985).

Hydrogenase 3 is the large subunit (537 amino acids after a 32-amino acid truncation) of a [Ni–Fe]-type hydrogenase (Drapal and Böck 1998) whose Ni-Fe active site contains one CO and two CN ligands (Blokesch et al. 2002). The CN and CO ligands may be synthesized from carbamoylphosphate via the postulated reaction: $2L_nFe + 2H_2NCOOPO_3^- \rightarrow 2L_nFeCONH_2 \rightarrow L_nFeCO + L_nFeCN)$; Paschos et al. 2001); however, CO synthesis is nebulous (Forzi et al. 2007). For [NiFe]-hydrogenase from *Ralstonia eutropha*, carbamoylphosphate serves as the source of CN^- but not of the intrinsic CO (Lenz et al. 2007).

Seven accessory proteins are required for maturing HycE (Drapal and Böck 1998) including the HycI protease that catalyzes a 32-amino acid C-terminal proteolytic cleavage of the HycE apoenzyme (Rossmann et al. 1994) and HypA,

HypB, HypC, HypD, HypE, and HypF for assembling the Ni–Fe metallocenter (Drapal and Böck 1998). HypA and HypB play a role in the nickel insertion followed by proteolytic removal of a C-terminal extension (Blokesch et al. 2002). HypB has guanosine triphosphate (GTP) activity (Maier et al. 1993); GTP hydrolysis is required for Ni insertion (Maier et al. 1995). HypC is a chaperone (Drapal and Böck 1998), HypD is a Fe–S protein (Blokesch et al. 2004a), and seven conserved cysteine residues in HypD are essential for hydrogenase maturation (Blokesch and Böck 2006); the HypC–HypD complex is involved in Fe insertion (Blokesch et al. 2002). In addition, HypC interacts with the HycE apoenzyme to avoid misfolding or to maintain a suitable conformation for metal incorporation (Drapal and Böck 1998). The cysteine residue in the motif MC(L/I/V)(G/A)(L/I/V)P at the amino terminus in HypC is necessary for the interaction with the HycE apoenzyme residues Cys241, Cys244, Cys531, and Cys534 for Fe binding, Ni binding, and cysteine bridging (Magalon and Böck 2000). HypE has ATPase activity (Blokesch et al. 2002), and HypF has carbamoyl transferase activity (Blokesch et al. 2004b), and these proteins contribute to the synthesis of the CN ligands (Blokesch et al. 2004b). Thus, maturation for hydrogenase 3 is complex (Blokesch et al. 2002).

In formate dehydrogenase-H, the cofactors are selenocysteine, molybdenum, two molybdopterin guanine dinucleotides, and a Fe_4S_4 cluster at the active site (Boyington et al. 1997). The *fdhF* messenger ribonucleic acid (RNA) hairpin structure promotes selenocysteine incorporation through binding to SelB, which is a selenocysteinyl-transfer RNA-specific translation factor (Hüttenhofer et al. 1996). There is little known about how hydrogenase 3 interacts with formate dehydrogenase-H, but it is postulated that formate dehydrogenase-H binds to HycB, a [4Fe-4S] ferredoxin-type peptide (Bagramyan and Trchounian 2003). Hence, random mutagenesis should help determine how HycE functions.

Metabolic engineering has been used successfully to enhance hydrogen production from *E. coli*. For example, we have created a quintuple mutant by inactivating hydrogenase 1 via *hyaB* (to prevent hydrogen consumption), inactivating hydrogenase 2 via *hybC* (to prevent hydrogen consumption), inactivating the FHL repressor via *hycA*, inactivating formate dehydrogenase-O via *fdoG* (to prevent formate consumption), and overexpression of the FHL activator via $fhlA^+$; these five mutations enhanced hydrogen production from formate 141-fold and achieved the theoretical hydrogen yield (1 mol H_2/mol formate; Maeda et al. 2008). In addition, we created a septuple mutant (*hyaB hybC hycA fdoG frdC ldhA aceE*) that increased hydrogen production fivefold from glucose and improved the hydrogen yield twofold from 0.65 to 1.3 mol H_2/mol glucose (Maeda et al. 2007b). Previously, a threefold increase in hydrogen production from formate was also obtained by inactivating the FHL repressor (HycA) and by overexpressing the FHL activator (FhlA) (Yoshida et al. 2005). Inactivation of the twin-arginine translocation system for transporting proteins into the periplasm led to a twofold increase of hydrogen production from glucose, although this mutation led to a significant decrease in cell viability (Penfold et al. 2006). Furthermore, deleting lactate dehydrogenase (*ldhA*) for converting pyruvate to lactate and fumarate reductase (*frdBC*) for converting phosphopyruvate to succinate resulted in a 1.4-fold increase in hydrogen production (Yoshida et al. 2006).

In contrast, protein-engineering studies for hydrogenases to enhance bacterial hydrogen production have not been developed extensively as there have been no high-throughput methods to readily measure hydrogenase activity (either directly or indirectly), whereas many beneficial proteins such as epoxide hydrolase (van Loo et al. 2004), amine oxidase (Carr et al. 2003), alkane hydroxylase (Glieder et al. 2002), toluene *para*-monooxygenase (Fishman et al. 2005), toluene-*o*-xylene monooxygenase (Vardar and Wood 2005), 2,4-dinitrotoluene dioxygenase (Leungsakul et al. 2005), and acetyltransferase (Castle et al. 2004) have been engineered for enhanced catalytic function through deoxyribonucleic acid (DNA) shuffling, error-prone polymerase chain reaction (epPCR), and saturation mutagenesis followed by novel high-throughput screening methods that enable the evaluation of many samples (100 to 1,000 colonies per screen). There is only one report concerning the evolution of a hydrogenase, one derived from *Clostridia* sp. via DNA shuffling (Nagy et al. 2007); however, no screening method was used, and little improvement occurred.

A novel chemochromic membrane to easily detect hydrogen produced by single colonies by colorimetric response by binding of hydrogen to a thin-film WO_3 sensor (Seibert et al. 1998) 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 (Maeda and Wood 2008). Herein, we report on evolving HycE for hydrogen production using epPCR, DNA shuffling, and saturation mutagenesis.

Materials and methods

Bacterial strains, growth, and total protein Parent strain *E. coli* K-12 BW25113 was obtained from the Yale University Coli Genetic Stock Center, and its isogenic deletion *hycE* was obtained from the Genome Analysis Project in Japan (Keio collection; Baba et al. 2006); *E. coli* BW25113 *hyaB hybC hycE Δkan* was constructed as described previously (Maeda et al. 2007c) and used as the host for screening the *hycE* variants since it lacks hydrogen production, hydrogen consumption, and kanamycin (Kan) resistance. All *E. coli*

strains were initially streaked from −80°C glycerol stocks on Luria–Bertani (LB) agar plates (Sambrook et al. 1989) containing 100 μg/mL Kan (for maintaining pBS(Kan)-based plasmids) and incubated at 37°C. Overnight cultures were made from fresh, single colonies (37°C, 250 rpm) in LB medium–Kan. The total protein concentrations were 0.22 mg OD^{-1} mL^{-1} (Fishman et al. 2005).

epPCR and DNA shuffling Plasmid pBS(Kan)HycE was constructed as described previously (Maeda et al. 2007a) and has *hycE* under control of a constitutive *lac* promoter. To introduce random mutations into the whole *hycE* locus, epPCR was performed using pBS(Kan)HycE as the template with two primers (epHycE-forward: 5′-ACAGCTATGAC CATGATTACGCC-3′ and epHycE-reverse: 5′-AAGGCG ATTAAGTTGGGTAA CGC-3′) as described previously (Cadwell and Joyce 1992). The epPCR products were cloned into the multiple cloning site (*Kpn*I and *Eco*RI) in pBS(Kan) after double digestion with *Kpn*I and *Eco*RI, and the ligation mixture was electroporated into BW25113 *hyaB hybC hycE*.

DNA shuffling was conducted as described previously (Canada et al. 2002). To isolate template DNA for DNA shuffling, the ep-HycE alleles harboring increased HycE activity were PCR-amplified using *Taq* polymerase with two primers (shufHycE-forward: 5′-TGCAGCTGGC ACGACAGGTTTCC-3′ and shufHycE-reverse: 5′-CAGG CTGCGCAACTGTTGGGAAGG-3′). Fragments (20 to 100 bp) for DNA shuffling were created by digesting the cleaned PCR product with DNase I and purified by using a Centri-Sep spin column (Princeton Separations, Adelphia, NJ, USA). The fragments were reassembled by PCR without primers. The 2.0-kb HycE allele was recovered by PCR with a nested front primer (epHycE-forward: 5′-ACAGCTATGACCATGATTACGCC-3′) and a nested rear primer (epHycE-reverse: 5′-AAGGCGATTAAGTTGGG TAACGC-3′). The shuffled PCR products were then cloned into pBS(Kan), and these plasmids were electroporated into BW25113 *hyaB hybC hycE*.

Saturation mutagenesis Saturation mutagenesis was performed at codon T366 of HycE using a QuikChange® XL Site-directed Mutagenesis Kit (Stratagene; La Jolla, CA, USA), and DNA primers were designed to vary 32 codons to allow for substitution of all 20 amino acids as described previously (Leungsakul et al. 2006). The 5′-ACT codon for HycE T366 was varied using primers with the variable NNS codon where N is A, G, C, or T and S is G or C: HycE366front 5′-GGTGGATGTGCTGCTGAGCNNSCC GAACATGGAACAGC-3′ and HycE366Rear 5′-CGACAG TGCGCTGTTCCATGTTCGGSNNGCTCAGCAGC-3′. The constructed plasmids were electroporated into BW25113 *hyaB hybC hycE*, and 360 of the generated colonies were screened with the chemochromic membranes to find variants showing high hydrogen-producing activity.

Screening Chemochromic membranes (GVD, Cambridge, MA, USA; Seibert et al. 1998) were used to identify HycE variants with beneficial mutations; that is, colonies were chosen on the basis of enhanced hydrogen production. The colonies from the epPCR, DNA shuffling, and saturation mutagenesis libraries were transferred to square agar plates (100×100×15 mm) containing modified complex-formate medium (100 mM formate; Maeda et al. 2008), and 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 hycE*, did not produce hydrogen (Maeda et al. 2007a) and remained colorless on the membrane. At least three replicates were checked for each candidate before proceeding to a gas chromatography (GC)-based hydrogen assay.

Closed hydrogen assay Overnight aerobic LB–Kan cultures (25 mL) and fresh modified complex-formate 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 (formate concentration 100 mM) were added to the bottles, which were incubated at 37°C with shaking for 2 to 20 h. 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 as described previously (Maeda et al. 2007c).

Low-partial-pressure hydrogen assay Overnight, aerobic LB cultures (25 mL, turbidity at 600 nm of 3.5) were used to inoculate 75 mL of the modified complex-formate medium in 250-mL shake flasks, and the cultures were sparged for 5 min with nitrogen, sealed, and incubated anaerobically at 37°C for 6 h. After 6 h, the cultures were poured anaerobically into 250-mL centrifuge tubes in an anaerobic glove box and were centrifuged (7,350×*g*) for 5 min at 4°C. The supernatant was decanted in the glove box, 30 mL of fresh modified complex medium without formate was added anaerobically, and then the cells were resuspended. Sealed crimp-top vials (60 mL) were sparged for 5 min with nitrogen, and 18 mL of the cell suspension and 2 mL of 1 M formate (final concentration 100 mM) were added to the vials, which were then incubated at 37°C anaerobically with shaking as described previously (Maeda et al. 2008). The hydrogen gas generated passed through a needle in the septum via tubing that directed the gas

through 1 M NaOH (to remove carbon dioxide [Klibanov et al. 1982]) and was collected in an inverted graduated cylinder, which was used to measure the volume of the gas (Maeda et al. 2008). Hence, low partial pressure was maintained in the head space of the vials. The vials were incubated at 37°C with stirring for 2.5 h, and hydrogen was assayed with the GC as described above. For yield calculations, the vials were incubated for 16 h.

SDS-PAGE, DNA sequencing, and modeling Expression of recombinant proteins was analyzed with standard Laemmli discontinuous sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE; 12%; Sambrook et al. 1989). A dideoxy chain termination technique (Sanger et al. 1977) with the ABI™ Prism BigDye Terminator Cycle Sequencing Ready Kit (PerkinElmer, Wellesley, MA, USA) was used to determine the *hycE* nucleotide sequences. HycE was modeled using the MOE software (Chemical Computing group, Montreal, Canada) and was based on the large subunit of hydrogenase derived from *Desulfovibrio gigas* (Volbeda et al. 1995).

Results

Hydrogenase 3 expression host E. coli cells possess four hydrogenases (hydrogenase 1 encoded by *hyaABCDEF* [Richard et al. 1999], hydrogenase 2 encoded by *hybOABCDEFG* [Richard et al. 1999], hydrogenase 3 encoded by *hycABCDEFGHI* [Bagramyan and Trchounian 2003; Sauter et al. 1992], and hydrogenase 4 encoded by *hyfABCDEFGHIR* [Andrews et al. 1997]). Hydrogenases 1 and 2 have hydrogen uptake activity only (Ballantine and Boxer 1986; King and Przybyla 1999), hydrogenase 4 appears to be inactive and remains cryptic (Self et al. 2004), and hydrogenase 3 has not only hydrogen synthesis activity but also hydrogen uptake activity (Maeda et al. 2007a). In this study, BW25113 *hyaB hybC hycE* was used as the host for cloning since it produces minimal hydrogen due to inactivating hydrogen consumption by hydrogenases 1, 2, and 3 and synthesis by hydrogenase 3 (Maeda et al. 2007a). Our goal was to engineer HycE for hydrogen synthesis using epPCR and DNA shuffling and to identify beneficial mutations as there have been no structure/function studies for this enzyme beyond studies concerning the four cysteines in its active site for binding nickel and iron (Magalon and Böck 2000). Plasmid pBS(Kan)HycE, which can complement the chromosomal *hycE* mutation (Maeda et al. 2007a), was used for protein engineering of HycE.

Error-prone polymerase chain reaction To introduce random mutations into *hycE*, epPCR was performed. From sequencing ten colonies, the maximum error rate was 0.53%. Four thousand five hundred forty colonies with epHycE alleles were screened using the chemochromic sensor method, which resulted in the identification of seven variants with beneficial mutations (Table 1) that showed high hydrogen-producing activity compared to BW25113 *hyaB hybC hycE* with wild-type HycE. To confirm that the phenotype was due to plasmid-based HycE, the plasmids were isolated and re-electroporated into the original BW25113 *hyaB hybC hycE* host. It was also confirmed via gel electrophoresis that each plasmid from the seven HycE variants was correct by digesting with *Kpn*I and *Eco*RI, and SDS-PAGE showed no change in expression of mutated HycE among the variants (data not shown), indicating that the difference of the hydrogen-producing activity in HycE variants is due to the HycE mutations.

Closed-vial hydrogen assay for the epHycE variants Hydrogen production with the seven plasmids was assayed using a closed hydrogen assay. The HycE variants epHycE17, epHycE23-2, and epHycE39 had seven- to eightfold higher hydrogen production than BW25113 *hyaB hybC hycE* expressing wild-type HycE from pBS(Kan) HycE in the closed hydrogen assay (Table 1). Furthermore, epHycE67 and epHycE70 produced over 11- to 13-fold more hydrogen, and epHycE21 and epHycE95 produced 16- to 17-fold more hydrogen. These increases in hydrogen production were due to one to eight amino acid changes (Table 1 and Fig. 1). epHycE39 had a truncation of 16 amino acids, and epHycE70 also had a truncation of 78 amino acids as well as a mutation in the ribosome-binding site.

Low-partial-pressure hydrogen assay in epHycE variants To confirm the results from the closed-vial assay, a low-partial-pressure assay was performed on the seven HycE variants. By maintaining low partial pressures of hydrogen, feedback inhibition is avoided (Maeda et al. 2008). The results agreed well with hydrogen production values slightly higher in the low-partial-pressure assay (the relative rates demonstrated similar trends in both assay systems; Table 1). During long incubations (16 h), these epHycE variants also had significantly improved hydrogen yields from formate with three- to eightfold enhancement relative to BW25113 *hyaB hybC hycE* expressing wild-type HycE. This indicates that the beneficial mutations make the strains more capable of producing hydrogen for long periods.

DNA shuffling To identify additional HycE variants that produce more hydrogen than the best epPCR variant epHycE95, DNA shuffling was conducted with two pools consisting of either three best variants (epHycE21, epHycE67, and epHycE95) or four best variants

Table 1 Hydrogen production and yield in *E. coli* BW25113 *hyaB hybC hycE* expressing wild-type HycE or re-electroporated HycE variants via pBS(Kan)HycE using the closed and low-partial-pressure assays (n=2)

HycE variants	HycE amino acid changes	H_2 production rate[a]		H_2 production rate[b]		H_2 yield[c]	
		μmol/mg protein/h	Relative	μmol/mg protein/h	Relative	mol H_2/mol formate	Relative
Wild-type	Wild-type HycE	0.3±0.4	1	0.6±0.5	1	0.09±0.03	1
epHycE17	F297L, L327Q, E382K, L415M, A504T, D542N	2.1±0.5	7	2.4±0.3	4	0.40±0.04	4
epHycE21	Q32R, V112L, G245C, F409L	4.6±0.2	15	4.68±0.06	8	0.558±0.004	6
epHycE23-2	D210N, I271F, K545R	2.3±0.6	8	2.3±0.6	4	0.33±0.04	4
epHycE39	I333F, K554[d]	2±1	7	1.43±0.03	2.4	0.23±0.01	3
epHycE67	S2P, E4G, M314V, T366S, V394D, S397C	4±2	13	4±1	7	0.44±0.03	5
epHycE70	D202V, K492[d]	3.2±0.1	11	5±2	8	0.60±0.08	7
epHycE95	S2T, Y50F, I171T, A291V, T366S, V433L, M444I, L523Q	5.0±0.6	17	10.3±0.5	17	0.7±0.1	8
shufHycE1-9	Y464[d]	6.8±0.5	23	13.6±0.5	23	0.84±0.02	9
satHycE12T366	T366[d]	9±4	30	–	–	–	–
satHycE19T366	T366[d]	8±1	27	–	–	–	–

[a] Hydrogen production rate by the closed assay system; the rate was calculated from hydrogen production after 5 h in complex-formate medium
[b] Hydrogen production rate by the low-partial-pressure assay system; the rate was calculated from hydrogen production after 2.5 h in complex-formate medium
[c] Hydrogen yield was calculated as mol H_2/mol formate from hydrogen production after 16 h in complex-formate medium by the low-partial-pressure assay system
[d] Nonsense mutations that result in truncation

(epHycE21, epHycE67, epHycE70, and epHycE95). Eight thousand one hundred sixty colonies with shuffled HycE alleles were screened using the chemochromic sensor-screening system, and out of nine possible variants studied via the closed-system assay (with the GC-based hydrogen assay), one variant with significantly greater hydrogen production was identified, shufHycE1-9. The shufHycE1-9 plasmid was digested by *Kpn*I and *Eco*RI to verify correct construction, and no change in expression of the shufHycE1-9 HycE variant was observed with SDS-PAGE (Fig. 2). After re-electroporating the plasmid, shuf HycE1–9 was found to have 23-fold higher hydrogen-synthesis compared to wild-type HycE in both the closed and low-partial-pressure hydrogen assays (Table 1). In addition, the yield was increased 9.3-fold. These increases in hydrogen production were the result of truncating 74 amino acids from the C terminus of the mature protein; since 32 amino acids are removed from the C terminus upon HycE maturation, this mutation results in a deletion of 106 amino acids compared to the apoenzyme (Fig. 1). It is surprising to note that the variant shuf HycE1-9 did not have any mutations from epHycE21, epHycE67, and epHycE95; however, the results of DNA sequencing of two random colonies from the shuffling library showed that DNA shuffling worked well since one random clone had a silent mutation derived from epHycE21 and the V394D mutation derived from epHycE67, and the other random clone had a silent mutation from epHycE95, mutation F409L from epHycE21, and one new mutation (data not shown). This indicates that in the 8,000 variant protein space sampled, the 74-amino acid truncation was superior to a recombination of the beneficial mutations from epHycE21, epHycE67, and epHycE95. The benefit of the carboxy-terminal truncation of the shuffling mutant was corroborated by the epPCR variant epHycE70 that had a similar mutation (Fig. 1). SDS-PAGE also showed that both the shufHycE1-9 (Fig. 2) and epHycE70 variants had truncated HycE alleles.

Saturation mutagenesis Since two of the epPCR HycE variants (epHycE67 and epHycE95) had the same amino acid change, T366S (Table 1), we investigated the importance of this position of HycE for enhanced hydrogen production by substituting all possible amino acids at this position via saturation mutagenesis. After screening 360 colonies to ensure with a probability of 99.999% that all possible codons were utilized (Rui et al. 2004), three variants (satHycE12T366, satHycE18T366, and satHycE19T366) were identified that had elevated hydrogen-producing activity compared to the epHycE95 epPCR variant using the chemochromic membranes. DNA sequencing revealed that all three of these mutants had the TAG stop codon at T366 (note that with NNS mutagenesis,

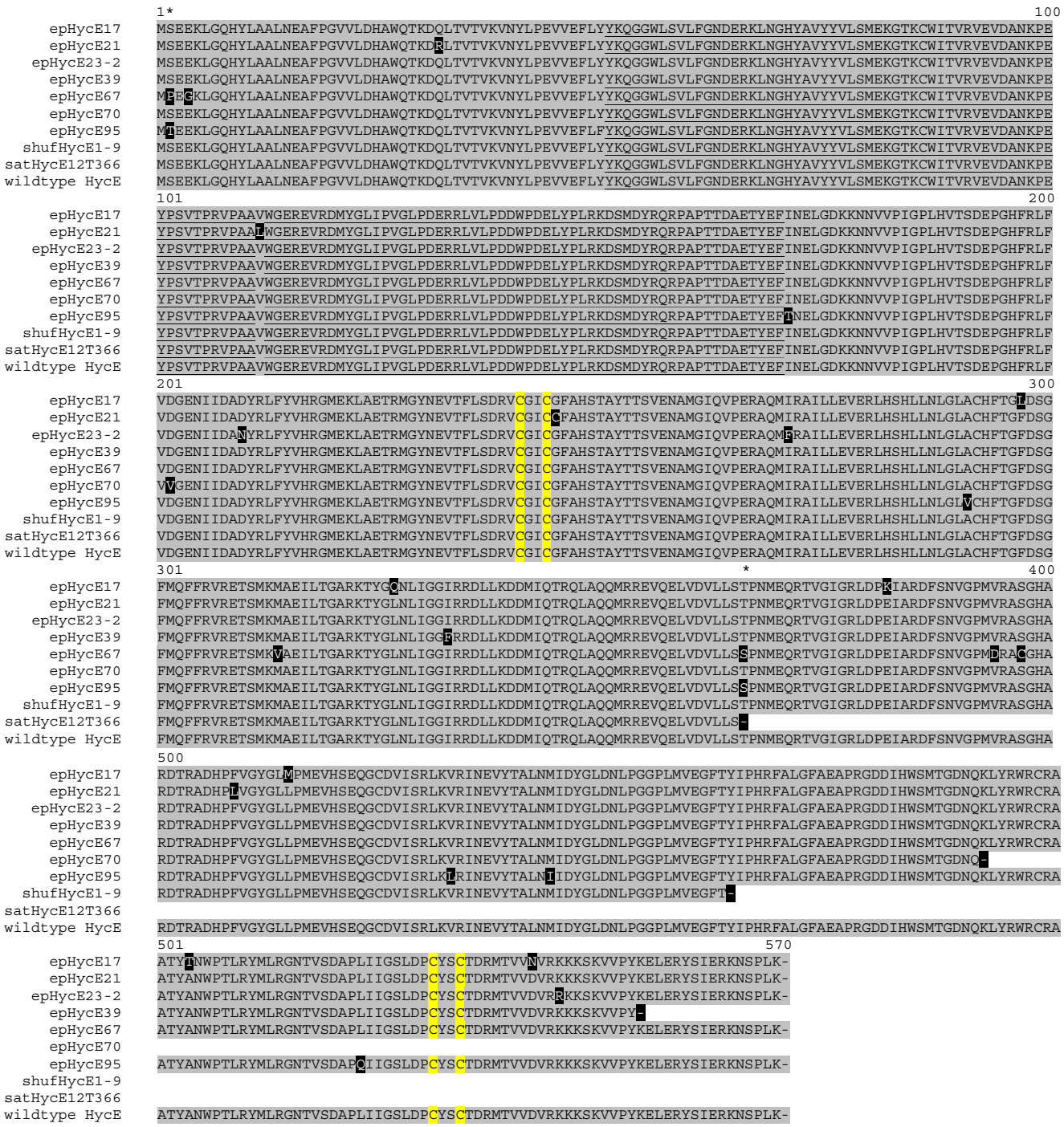

Fig. 1 Protein sequences of the hydrogenase 3 epPCR, DNA shuffling, and saturation mutagenesis variants. Amino acid changes are indicated by *black highlight*, and conserved Ni–Fe hydrogenase cysteines are indicated by *no shading* (Cys_{241}, Cys_{244}, Cys_{531}, and Cys_{534}). *Underline* indicates amino acid positions 51 to 111 and 113 to 170 in which there are no mutations, and two *asterisks* indicate amino acid positions 2 and 366 in which amino acids were changed in both epHycE67 and epHycE95

only the TAG stop codon is possible). After re-electroporating the plasmid into the original host BW25113 *hyaB hybC hycE* to confirm the phenotype was due to plasmid-based HycE, hydrogen production in two of the saturation mutagenesis mutants, satHycE12T366 and satHycE19T366, was assayed using the closed hydrogen assay; these variants produced 18% to 32% more hydrogen than the best DNA-shuffling mutant, shufHycE1-9 (Table 1); hence, the 204-amino acid truncation was consistently found and consistently led to enhanced hydrogen production. The 204-amino acid HycE truncation of the saturation mutagenesis variant was verified using SDS-PAGE (Fig. 2).

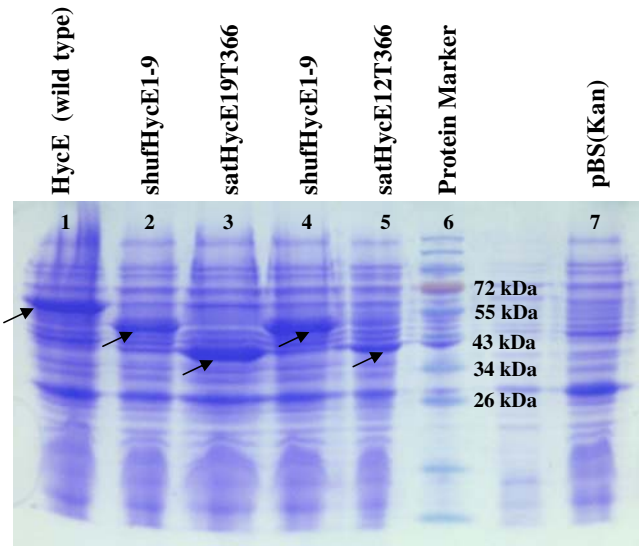

Fig. 2 Expression of the truncated HycE variants shufHycE1-9, satHycE19T366, and satHycE12T366 along with wild-type HycE in *E. coli hyaB hybC hycE*/pBS(Kan)HycE as visualized by SDS-PAGE. *Arrows* indicate bands corresponding to HycE and HycE variants. *E. coli hyaB hybC hycE*/pBS(Kan) (empty vector) was used as a negative control (no expression of HycE)

Protein modeling of HycE To evaluate where the beneficial mutations lie, the MOE software was used to create a homology model based on the large subunit of the *D. gigas* [Ni–Fe]-hydrogenase (Volbeda et al. 1995); these proteins share 14% identity. Seven of the eight mutations in epHycE95 are shown in Fig. 3.

Discussion

In this study, we constructed nine HycE variants (Table 1) that can produce up to 30-fold higher hydrogen than BW25113 *hyaB hybC hycE*/pBS(Kan)HycE through ep-PCR, DNA shuffling, 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 (Woods 1936). These results may be improved by eliminating formate dehydrogenase-N and formate dehydrogenase-O (Maeda et al. 2008).

Unlike many engineered proteins that are a single polypeptide, the large subunit of hydrogenase 3, HycE, is part of the membrane-bound FHL complex (FdhH-HycB-HycC-HycD-HycF-HycG-HycE) and binds HycG, the small subunit of hydrogenase 3 (Bagramyan and Trchounian 2003). Hence, the improvements in hydrogen production may be due to enhanced catalytic reactions by the large subunit, enhanced maturation, or improved FHL complex formation. For example, epHycE70 and shufHycE1-9 should produce hydrogen without requiring HycI because these HycE alleles have unexpected termination codons at position 492 and 464, respectively (note that HycI cleaves wild-type HycE at position 537 [Rossmann et al. 1994]). Furthermore, via saturation mutagenesis, we also found a variant satHycE12T366 with a truncation at codon T366 of HycE, which produces more hydrogen than the DNA shuffling variant shufHycE1-9. All three of these truncations cause the loss of two important cysteine residues at C_{531} and C_{534} that are used by the wild-type HycE for Ni and Fe binding as part of a $DPCX_2CX_2(H/R)$ motif; previously, the replacement of these cysteine residues with alanine residues led to a significant decrease in hydrogenase activity (Magalon and Böck 2000). Our surprising truncation results (Fig. 1) found through three independent protein-engineering methods (DNA shuffling, epPCR, and saturation mutagenesis) suggest that perhaps some new interaction with HycG is required to incorporate Ni and Fe after the beneficial truncations; this hydrogenase small subunit has seven cysteines in its 255 amino acids and may be a suitable target for additional mutagenesis. Furthermore, HycE may have two distinct functions: The N-terminal region may be most significant for producing hydrogen, and the C-terminal region may repress hydrogen production; hence, deleting the C terminus by truncation triggers enhanced hydrogen production. Other possibilities for the increased hydrogen production include alterations

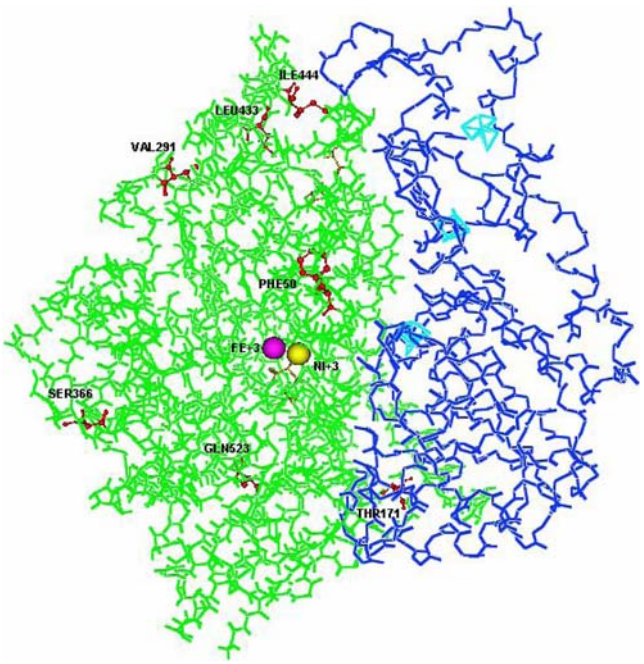

Fig. 3 Protein modeling of the large subunit of hydrogenase 3, HycE (*green*), along with the small subunit HycG (*blue*). Mutations in epHycE95 are displayed in *red* (*PHE50* for Y50F, *THR171* for I171T, *VAL291* for A291V, *SER366* for T366S, *LEU433* for V433L, *ILE444* for M444I, and *GLN523* for L523Q; S2T not shown). Metal cofactors Ni^{+} (*yellow*), Fe^{+3} (*pink*) and Fe–S clusters (*blue*) are indicated

in membrane insertion of the proteins and changes in allosteric regulation.

From the eight HycE variants (seven epPCR variants and one shuffling variant), there are no mutations from amino acid positions 51 to 111 and 113 to 170; hence, these regions appear important for hydrogenase activity. Furthermore, epHycE67 and epHycE95 have amino acid changes at the same positions (2 and 366) indicating that the amino acid changes S2P/S2T and T366S may be important for improved activity (note at position 366, a hydroxyl appears to be required since serine was substituted for threonine). Therefore, these two positions may be significant for enhanced hydrogen production and may be good targets for saturation mutagenesis.

E. coli hydrogenase 3 and *D. gigas* hydrogenase are [Ni–Fe]-type hydrogenases (Drapal and Böck 1998; Volbeda et al. 1995); however, our bioinformatics analysis (Vardar-Schara et al. 2008) indicates that HycE of hydrogenase 3 does not have all six [Ni–Fe]-binding motifs of the *D. gigas* hydrogenase and the four present motifs are not in the same order in the primary sequence (Burgdorf et al. 2002), whereas *E. coli* hydrogenase 1 and hydrogenase 2 (hydrogen uptake activity) (Maeda et al. 2007a) contain all six motifs in the same order as the *D. gigas* hydrogenase (Vardar-Schara et al. 2008). Furthermore, the large subunits of hydrogenases 1 and 2 have 40% and 42% identity, respectively, compared to that of *D. gigas*, whereas HycE has only 14% identity. This indicates that *E. coli* hydrogenase 3 differs significantly from that of *D. gigas*.

Given its low identity, the homology modeling for HycE is just a starting point for identifying the relevance of the beneficial mutations found here. With this caveat, the modeling suggests that the mutations in the HycE variants are positioned primarily on the surface opposite the small subunit of hydrogenase such as A291V, T366S, and V433L shown for the epHycE95 variant (Fig. 3). This position was also observed for the beneficial mutations in the epHycE17, epHycE21, epHycE23-2, epHycE39, and epHycE67 variants. In contrast, D202V of the epHycE70 variant and Y50F/I171T/M441I of the epHycE95 variant were adjacent to the small subunit of the hydrogenase; these mutations may strengthen the interaction between the large subunit and small subunit and thereby enhance the flow of electrons generated from formate by formate dehydrogenase-H. Note that the electrons created by formate dehydrogenase-H flow to HycB to HycF to HycG to HycE (Bagramyan and Trchounian 2003). Since metabolic engineering (rational pathway engineering approach) has been used to enhance hydrogen production with *E. coli* (Bisaillon et al. 2006; Maeda et al. 2007b, 2008; Penfold et al. 2006; Yoshida et al. 2005, 2006), the HycE 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 (Patnaik et al. 2002) and DNA microarray analysis (Maeda et al. 2007c) to discern where the random chromosomal mutations lie. Hence, using *E. coli* as a reference system for producing hydrogen has many advantages.

Acknowledgments The authors are grateful for the Keio clones sent by the National of Institute of Genetics (Japan), for the chemochromic membrane sensor provided by Dr. Hilton G. Pryce Lewis of GVD, and for protein modeling provided by Dr. Thammajun Leungsakul. This research was supported by DARPA (HR0011-06-1-0001).

References

Andrews SC, Berks BC, McClay J, Ambler A, Quail MA, Golby P, Guest JR (1997) A 12-cistron *Escherichia coli* operon (*hyf*) encoding a putative proton-translocating formate hydrogenlyase system. Microbiology 143:3633–3647

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

Ballantine SP, Boxer DH (1986) Isolation and characterisation of a soluble active fragment of hydrogenase isoenzyme 2 from the membranes of anaerobically grown *Escherichia coli*. Eur J Biochem 156:277–284

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 3rd, 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

Blokesch M, Paschos A, Theodoratou E, Bauer A, Hube M, Huth S, Böck A (2002) Metal insertion into NiFe-hydrogenases. Biochem Soc Trans 30:674–680

Blokesch M, Albracht SP, Matzanke BF, Drapal NM, Jacobi A, Böck A (2004a) The complex between hydrogenase-maturation proteins HypC and HypD is an intermediate in the supply of cyanide to the active site iron of [NiFe]-hydrogenases. J Mol Biol 344: 155–167

Blokesch M, Paschos A, Bauer A, Reissmann S, Drapal N, Böck A (2004b) Analysis of the transcarbamoylation–dehydration reaction catalyzed by the hydrogenase maturation proteins HypF and HypE. Eur J Biochem 271:3428–3436

Blokesch M, Böck A (2006) Properties of the [NiFe]-hydrogenase maturation protein HypD. FEBS Lett 580:4065–4068

Boyington JC, Gladyshev VN, Khangulov SV, Stadtman TC, Sun PD (1997) Crystal structure of formate dehydrogenase H: catalysis

involving Mo, molybdopterin, selenocysteine, and an Fe4S4 cluster. Science 275:1305–1308

Burgdorf T, De Lacey AL, Friedrich B (2002) Functional analysis by site-directed mutagenesis of the NAD^+-reducing hydrogenase from *Ralstonia eutropha*. J Bacteriol 184:6280–6288

Cadwell RC, Joyce GF (1992) Randomization of genes by PCR mutagenesis. PCR Methods Appl 2:28–33

Canada KA, Iwashita S, Shim H, Wood TK (2002) Directed evolution of toluene *ortho*-monooxygenase for enhanced 1-naphthol synthesis and chlorinated ethene degradation. J Bacteriol 184:344–349

Carr R, Alexeeva M, Enright A, Eve TS, Dawson MJ, Turner NJ (2003) Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity. Angew Chem Int Ed Engl 42:4807–4810

Castle LA, Siehl DL, Gorton R, Patten PA, Chen YH, Bertain S, Cho HJ, Duck N, Wong J, Liu D, Lassner MW (2004) Discovery and directed evolution of a glyphosate tolerance gene. Science 304:1151–1154

Das D, Veziroglu TN (2001) Hydrogen production by biological processes: a survey of literature. Int J Hydrogen Energy 26: 13–28

Drapal N, Böck A (1998) Interaction of the hydrogenase accessory protein HypC with HycE, the large subunit of *Escherichia coli* hydrogenase 3 during enzyme maturation. Biochemistry 37: 2941–2948

Fishman A, Tao Y, Rui L, Wood TK (2005) Controlling the regiospecific oxidation of aromatics via active site engineering of toluene *para*-monooxygenase of *Ralstonia pickettii* PKO1. J Biol Chem 280:506–514

Forzi L, Hellwig P, Thauer RK, Sawers RG (2007) The CO and CN^- ligands to the active site Fe in [NiFe]-hydrogenase of *Escherichia coli* have different metabolic origins. FEBS Lett 581:3317–3321

Glieder A, Farinas ET, Arnold FH (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. Nat Biotechnol 20:1135–1139

Hüttenhofer A, Heider J, Böck A (1996) Interaction of the *Escherichia coli fdhF* mRNA hairpin promoting selenocysteine incorporation with the ribosome. Nucleic Acids Res 24:3903–3910

King PW, Przybyla AE (1999) Response of *hya* expression to external pH in *Escherichia coli*. J Bacteriol 181:5250–5256

Klibanov AM, Alberti BN, Zale SE (1982) Enzymatic synthesis of formic acid from H_2 and CO_2 and production of hydrogen from formic acid. Biotechnol Bioeng 24:25–36

Lenz O, Zebger I, Hamann J, Hildebrandt P, Friedrich B (2007) Carbamoylphosphate serves as the source of CN^-, but not of the intrinsic CO in the active site of the regulatory [NiFe]-hydrogenase from *Ralstonia eutropha*. FEBS Lett 581:3322–3326

Leungsakul T, Keenan BG, Yin H, Smets BF, Wood TK (2005) Saturation mutagenesis of 2,4-DNT dioxygenase of *Burkholderia* sp. strain DNT for enhanced dinitrotoluene degradation. Biotechnol Bioeng 92:416–426

Leungsakul T, Johnson GR, Wood TK (2006) Protein engineering of the 4-methyl-5-nitrocatechol monooxygenase from *Burkholderia* sp. strain DNT for enhanced degradation of nitroaromatics. Appl Environ Microbiol 72:3933–3939

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

Maeda T, Sanchez-Torres V, Wood TK (2007b) Enhanced hydrogen production from glucose by metabolically engineered *Escherichia coli*. Appl Microbiol Biotechnol 77:879–890

Maeda T, Vardar G, Self WT, Wood TK (2007c) Inhibition of hydrogen uptake in *Escherichia coli* by expressing the hydrogenase from the cyanobacterium *Synechocystis* sp. PCC 6803. BMC Biotechnol 7:25

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

Maeda T, Wood TK (2008) Formate detection by potassium permanganate for enhanced hydrogen production in *Escherichia coli*. Int J Hydrogen Energy (in press)

Magalon A, Böck A (2000) Analysis of the HypC–hycE complex, a key intermediate in the assembly of the metal center of the *Escherichia coli* hydrogenase 3. J Biol Chem 275:21114–21120

Maier T, Jacobi A, Sauter M, Böck A (1993) The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. J Bacteriol 175:630–635

Maier T, Lottspeich F, Böck A (1995) GTP hydrolysis by HypB is essential for nickel insertion into hydrogenases of *Escherichia coli*. Eur J Biochem 230:133–138

Nagy LE, Meuser JE, Plummer S, Seibert M, Ghirardi ML, King PW, Ahmann D, Posewitz MC (2007) Application of gene-shuffling for the rapid generation of novel [FeFe]-hydrogenase libraries. Biotechnol Lett 29:421–430

Paschos A, Glass RS, Böck A (2001) Carbamoylphosphate requirement for synthesis of the active center of [NiFe]-hydrogenases. FEBS Lett 488:9–12

Patnaik R, Louie S, Gavrilovic V, Perry K, Stemmer WP, Ryan CM, del Cardayré S (2002) Genome shuffling of *Lactobacillus* for improved acid tolerance. Nat Biotechnol 20:707–712

Penfold DW, Sargent F, Macaskie LE (2006) Inactivation of the *Escherichia coli* K-12 twin-arginine translocation system promotes increased hydrogen production. FEMS Microbiol Lett 262:135–137

Richard DJ, Sawers G, Sargent F, McWalter L, Boxer DH (1999) Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli*. Microbiology 145(Pt 10):2903–2912

Rossmann R, Sauter M, Lottspeich F, Böck A (1994) Maturation of the large subunit (HYCE) of *Escherichia coli* hydrogenase 3 requires nickel incorporation followed by C-terminal processing at Arg537. Eur J Biochem 220:377–384

Rui L, Kwon Y-M, Fishman A, Reardon KF, Wood TK (2004) Saturation mutagenesis of toluene ortho-monooxygenase of *Burkholderia cepacia* G4 for enhanced 1-naphthol synthesis and chloroform degradation. Appl Environ Microbiol 70: 3246–3252

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

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

Sauter M, Böhm R, Böck A (1992) Mutational analysis of the operon (*hyc*) determining hydrogenase 3 formation in *Escherichia coli*. Mol Microbiol 6:1523–1532

Sawers RG, Ballantine SP, Boxer DH (1985) Differential expression of hydrogenase isoenzymes in *Escherichia coli* K-12: evidence for a third isoenzyme. J Bacteriol 164:1324–1331

Seibert M, Flynn T, Benson D, Tracy E, Ghirardi M (1998) Biohydrogen. Plenum, New York, pp 227–234

Self WT, Hasona A, Shanmugam KT (2004) Expression and regulation of a silent operon, *hyf*, coding for hydrogenase 4 isoenzyme in *Escherichia coli*. J Bacteriol 186:580–587

van Loo B, Spelberg JH, Kingma J, Sonke T, Wubbolts MG, Janssen DB (2004) Directed evolution of epoxide hydrolase from *A. radiobacter* toward higher enantioselectivity by error-prone PCR and DNA shuffling. Chem Biol 11:981–990

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

Vardar G, Wood TK (2005) Protein engineering of toluene-o-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for enhanced chlorinated ethene degradation and o-xylene oxidation. Appl Microbiol Biotechnol 68:510–517

Volbeda A, Charon MH, Piras C, Hatchikian EC, Frey M, Fontecilla-Camps JC (1995) Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. Nature 373:580–587

Woods DD (1936) Hydrogenlyases: the synthesis of formic acid by bacteria. Biochem J 30:515–527

Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H (2005) Enhanced hydrogen production from formic acid by formate hydrogen lyase-overexpressing *Escherichia coli* strains. Appl Environ Microbiol 71:6762–6768

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