SUPPLEMENTARY INFORMATION

Ultra-Rapid Rates of Water Splitting for Biohydrogen Gas Production through *in vitro* Artificial Enzymatic Pathways

Eui-Jin Kim,[1,†] Jae-Eung Kim,[1] and Yi-Heng P. Job Zhang[2*]

[1] Biological Systems Engineering Department, Virginia Tech, Blacksburg, Virginia 24061, USA

[2] Current address: Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Avenue, Tianjin Airport Economic Area, Tianjin 300308, China

[†] Current address: Department of Life Science and Basic Science Institute, Sogang University, Seoul 04107, Republic of Korea

* Corresponding Author: Y.-H. Zhang (ORCID ID: 0000-0002-4010-2250); email: yhjob_zhang@outlook.com, Tel: 001-540-235-4299.

TABLE OF CONTENTS

Materials and Methods

Chemicals and strains

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Waltham, MA), unless noted otherwise. The enzymes for DNA manipulations were purchased from New England Biolabs (NEB, Ipswich, MA). *Escherichia coli* Top10 (Invitrogen, Carlsbad, CA) was used to conduct DNA manipulations as well as the high-throughput screening for NAD^+-preferring glucose 6-phosphate dehydrogenase (mG6PDH). *E. coli* BL21 (DE3) (Invitrogen) was used to conduct the expressions of recombinant enzymes used in this study (**Table 1**). A purified over-expressed soluble [NiFe]-hydrogenase from *Pyrococcus furiosus* (SHI) was prepared as described elsewhere (Chandrayan et al. 2015).

In silico structure analysis of enzymes

The three-dimensional (3D) molecular structure for TmG6PDH (glucose 6-phosphate dehydrogenase from *Thermotoga maritima*) (Hansen et al. 2002) was built by SWISS-MODEL based on the known 3D-structure of *Leuconostoc mesenteroides* G6PDH (PDB: 1DPG) as a template (sequence identity: 35%, similarity: 52%). Also, 3D-structure modeling for DI (a noble diaphorase (DI) from *T. maritima*) was carried out with *P. furiosus* NADH oxidase/nitrite reductase (PDB: 1XHC) as a template (sequence identity: 53%, similarity: 71%). The structures of $NADP^+$ and NAD^+ were built by using Chemdraw (PerkinElmer, Waltham, MA). The conformation space of the corresponding coenzyme binding area was analyzed using Autodock program (Scripps Research Institute, La Jolla, CA) as described previously (Chen et al. 2016).

Construction of plasmids

Plasmid pET28a-P_{tac}-*g6pdh* (**Figure S1**) was constructed as follows. The gene coding for TmG6PDH (G6PDH from *T. maritima*) (Hansen et al. 2002) was amplified from pET20b-*g6pdh* by using a pair of primers G6PDH-F (5′-AT**G AAT TC**A TGA AGT GCA GTC TGG GAT TGG-3′: *Eco*RI site is in boldface) and G6PDH-R (5′-AT**C TCG AG**C AGT TTT CTC CAT TTT CTA CCG-3′: *Xho*I site is in boldface). The resulting 1.5-kb DNA fragment was digested with *Eco*RI/*Xho*I and then cloned into the *Eco*RI/*Xho*I sites of a pET28a-based *tac* promoter containing vector pET28a-P_{tac}-*6pgdh* (Huang et al. 2016) to yield pET28a-P_{tac}-*g6pdh*, in which the gene coding for TmG6PDH lies under the control of both T7 promoter and *tac* promoter.

The gene (KEGG, TM0754) encoding a putative diaphorase from *T. maritima* (DI) was amplified from the genomic DNA of *T. maritima* MSB8 by PCR with a pair of primers TmDI-F (5′-AT**C TCG AG**G TGA AAG TAG TGA TCG TTG GAA AC-3′: *Xho*I site is in boldface) and TmDI-R (5′-AT**C TGC AG**T CAT CGC GTG CTT CTT AGT CTT TCC-3′: *Pst*I site is in boldface). The resulting 1.1-kb DNA fragments were digested with *Xho*I/*Pst*I and then cloned into the *Xho*I/*Pst*I sites of a pTWIN1 (NEB, Ipswich, MA) based CBM (cellulose binding module)-intein containing vector pCIG (Hong et al. 2008) to yield pCI-TmDI, in which the N-terminus region of DI is fused to the C-terminus region of CBM-intein fusion domains.

Construction of a G6PDH mutant libraries for coenzyme engineering

The saturation mutagenesis libraries for S33, A64, and R65T66 positions of TmG6PDH were constructed by using the NEB Phusion site-directed mutagenesis kit (Ipswich, MA). In order to generate

a mutant library for S33 site, a pair of degenerate primers S33-F (5′-TGA TCT TCG GTG CTN NKG GTG ACC TCA CAA AAA GAA AGC-3′: randomized positions are underlined, unless noted otherwise) and S33-R (5′-CAA TTC CGA AAG GTT GTT CGA TC-3′) were used for PCR amplification. Similarly, following degenerate primer sets A64-F (5′-GTT CTC GGT GCG NNK CGA ACG AAA ATG GAC GAT AAG -3′) / ART-R (5′-GAA AAA GCG TTC GGG CAG AAT ACC-3′) and R65T66-F (5′-GTT CTC GGT GCG GCA NNK NNK AAA ATG GAC GAT AAG AAA TTC AG-3′) / ART-R were used for PCR reactions to generate the mutant libraries for A64 and R65T66 sites, respectively. PCR reactions (50 μL) were conducted with 1 ng of template plasmid (pET28a-P*tac*-*g6pdh*) under the following conditions: 98°C denaturation for 2 min; 20 cycles of 98°C denaturation for 30 s, 60°C annealing for 30 s and 72°C extension for 2.5 min; and final 72°C extension for 5 min. The resulting PCR products were digested with *Dpn*I followed by the purification with agarose gel electrophoresis and Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). The purified mutant libraries were transformed into *E. coli* TOP10 for further screening steps described in the next section.

High-throughput screening for NAD-dependent G6PDH

The double-layer screening of TmG6PDH mutants on NAD^+ was carried out by the screening method developed previously (Huang et al. 2016) with minor modifications. After transformation of the mutant plasmid libraries, *E. coli* TOP10 cells were spread onto 1.5% agar Luria-Bertani (LB) medium containing 50 μg/mL kanamycin with an expected colony number of 200-300 per Petri dish. After incubation at 37°C overnight, the dishes were heat-treated at 80°C for 1 h to kill the *E. coli* cells, deactivate *E. coli* mesophilic enzymes, and degrade metabolites as well as intracellular coenzymes (Huang et al. 2016). 10 mL of a 0.5% (wt/vol) melted agarose solution (60°C) containing 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 50 μM tetranitroblue tetrazolium chloride (TNBT), 1.0 μM phenzine methosulfate (PMS), 2 mM glucose-6-phosphate, and 0.1 mM NAD^+ was carefully poured on the heat-treated colonies. After incubation at room temperature for 30-60 min, positive colonies were identified based on the formation of black haloes. The identified colonies embedded in the agarose gel were taken out with a sterile toothstick and then retrieved by using Zymo ZR Plasmid Miniprep kit (Zymo Research, Irvine, CA). The selected plasmids extracted from the agarose gel were transformed into *E. coli* TOP10 for further analysis.

Purification of enzymes

All recombinant enzymes used in this study are listed in **Table 1**. *E. coli* BL21 (DE3) strains transformed with the corresponding expression plasmid were grown in 3 mL LB medium containing 100 μg/mL ampicillin or 50 μg/mL kanamycin at 37°C overnight. The seed cultures were inoculated into a flask containing 200 mL of LB medium with the appropriate antibiotic and incubated in a rotary shaking rate of 250 rpm at 37 °C until A_{600} reached 0.6-0.8, at which point recombinant protein expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.01-0.1 mM. The cell cultures were incubated at 37 °C for 6 h or at 18°C for 16-20 h for protein expression. After their cultivation, cells were harvested by centrifugation at 4°C and then washed twice with the Tris-HCl buffer (pH 7.5) containing 50 mM NaCl. The cell pellets were resuspended in a 0.1 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer (pH 7.5) containing 50 mM NaCl and 10 mM imidazole. The cell lysates, generated by sonication and centrifugation, were loaded onto a column packed with HisPur Ni-NTA resin (Fisher Scientific). Lastly, the His-tagged proteins

were eluted with a 0.1 M HEPES buffer (pH 7.5) containing 50 mM NaCl and 250 mM imidazole. Protein concentration was determined by the Bradford assay using bovine serum albumin (BSA) as a standard.

Purification and reconstitution of diaphorases

Expressions of the recombinant diaphorase enzymes, $NADP^+$-dependent rubredoxin oxidoreductase from *Pyrococcus furiosus* (NROR) and a noble diaphorase from *Thermotoga maritima* (DI), were carried out using *E. coli* BL21 (DE3) (Invitrogen) in LB medium containing 100 μg/mL ampicillin. Cells harboring expression plasmid pCI-NROR (Kim et al. 2016) or pCI-TmDI (See the "Construction of plasmids" section) were cultivated in a rotary shaker with a shaking rate of 250 rpm at 37°C until A_{600} reached 0.4-0.6, at which point 0.1 mM IPTG was added to induce protein expressions at 18°C for 16-20 h. CBM-intein-PfuNROR and CBM-intein-TmDI fusion proteins were purified by using a regenerated amorphous cellulose (RAC) (Zhang et al. 2006) as described previously (Hong et al. 2008). RAC absorption was conducted by mixing of a cell lysate in 50 mM Tris-HCl buffer (pH 8.5) with 0.1 mg of RAC at room temperature for 15 min. The RAC pellets were centrifuged and then resuspended in the same buffer to remove impure proteins in the RAC matrix. Self-cleavage by intein that linked between CBM and target protein was induced by an acidic incubation of the proteins as described previously (Hong et al. 2008). To induce the self-cleavage as well as the reconstitution with a flavin, RAC pellets harboring the absorbed CBM tagged proteins were suspended with 10 mL of a 50 mM HEPES buffer (pH 6.5) containing 0.1 mM flavin adenine dinucleotide (FAD) and 0.5 M NaCl and then incubated at 37°C for 9 h. Reconstituted NROR or DI were recovered from the incubated RAC slurry mixture to a supernatant after centrifugation at 6000 rpm for 15 min at 4°C. The enzymes were further purified and concentrated in 0.1 M HEPES buffer (pH 7.5) by using the Amicon Ultra centrifugal filter with 10 kDa cut-offs (10,000 MWCO, Millipore, Billerica, MA).

Preparation of BV-conjugated DI (BCV-DI)

The BCV was synthesized by two chemical reactions (**Fig. 4a).** Benzyl iodide reacted with 4,4'-bipyridine, yielding monobenzyl viologen (MBV). MBV was further reacted with 4-(bromomethyl)phenylacetic acid**,** yielding BCV. BCV has a complete BV moiety as well as the carboxylate group required for the formation of an amide linkage with amine groups on the DI. Like BV, the resulting BV-conjugated DI was readily reduced by sodium dithionite (E_m < - 500 mV) as indicated by the color conversion from colorless to blue.

Benzyl (4-carboxymethyl)benzyl viologen (BCV) was prepared in a two-step process as elucidated in **Figure 4a**. BCV synthesis was carried out based on a method generated for the organic synthesis of ethyl carboxyethyl viologen (ECV) as reported previously (Dinh et al. 2016) with minor modifications. Benzyl iodide (28 mmol) was added dropwise to a 100 mL of acetone containing 12 mmol 4,4′-bipyridine and then mixed well at 60°C for 20 h. The resulting precipitate was filtered by a filter paper and then washed with acetone several times to remove the residual impurity (Hansen et al. 2012). The purified synthetic intermediate, monobenzyl viologen (MBV) (3 mmol), was dissolved in a 250 mL of acetonitrile containing 4-(bromomethyl)phenylacetic acid (30 mmol), and the mixture was refluxed at 90°C for 48 h. The resulting precipitate was washed with toluene several times by using a filter paper and then dried by using N_2 flow. The final synthetic product BCV was stored at 4°C.

To activate the carboxyl group of BCV, 0.2 M BCV dissolved in 0.2 M phosphate buffer (pH 7.5) was mixed gently with the same volume (2.5 mL) of 0.1 M MES buffer (pH 5.0) containing 0.2 M 1-

ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) (Dinh et al. 2016). After this activation reaction at room temperature for 1 h, NaOH solution was added dropwise to the reaction mixture until the pH was adjusted to 7.5. To generate benzyl viologen (BV) conjugated enzyme, 20 mg of DI in 2.5 mL of 0.2 M phosphate buffer (pH 7.5) was added to the above mixture and stirred slowly at room temperature overnight. The resulting BV-conjugated DI was washed with 0.1 M HEPES buffer (pH 7.5) several times and then concentrated by ultrafiltration with 10,000 MWCO Amicon Ultra centrifugal filter.

The relative molecular ratio of BV moiety and DI enzyme in the resulting BV-conjugated DI was determined as follows. The amount of BV moiety in the BV-conjugated DI was determined by measuring the absorbance change by the BV reduction in the presence of 10 mM sodium dithionite ($Na_2S_2O_4$). The reaction of BV reduction was carried out in an anaerobic screw-cap IR quartz cuvette (Reflex Analytical Co., Ridgewood, NJ) with a degassed 0.1 M HEPES buffer (pH 7.5). The amount of functional BV moiety in the BV-conjugated DI was calculated with a known millimolar extinction coefficients (ε) for reduced benzyl viologen (BV^+) at a wavelength of 578 nm, 8.65 $mM^{-1}cm^{-1}$. On the other hand, the concentration of DI was determined by the Bradford assay using bovine serum albumin (BSA) as a standard.

Preparation of NAD-conjugated dehydrogenases

The covalently linked NAD-PEG complex, which was synthesized with N^6-(2-carboxyethyl)-NAD (Kishimoto et al. 1991; Sakamoto et al. 1986), diamino PEG (MW ~3,000), and a cross-linking reagent EDC (Katayama et al. 1983; Kishimoto et al. 1991), was conjugated to the N-terminus region of dehydrogenases by generating an activated NAD-PEG intermediate (Eguchi et al. 1986).

A synthetic intermediate, 1-(2-carboxyethyl)-NAD was prepared by the following procedure as described previously (Kishimoto et al. 1991). To generate the alkylated NAD, 2 g of NAD was dissolved in 100 mL of 3-iodopropionate solution (0.1 M). The pH of the reaction mixture was adjusted to 6.5 using 2 M LiOH. 3-iodopropionic acid (0.4 g) was added to the mixture every 24 h, and the pH of mixture was readjusted to 6.5 using 2 M LiOH. The mixture was stirred in the dark condition for 120 h at room temperature. The mixture was gently poured into 2 L of cold acetone/ethanol (4:1). The resulting precipitate was collected by centrifugation, and washed with acetone/ethanol (4:1) then acetone, and finally dried by using N_2 flow. The crude 1-(2-carboxyethyl)-NAD was purified by DEAE-Sephadex A-25 column chromatography as described previously (Sakamoto et al. 1986). The dried precipitate dissolved in water was applied on the column (2.0 × 50 cm) equilibrated with 20 mM lithium acetate buffer (pH 5.0), and the nucleotides were eluted with a linear gradient of 0-0.5 M LiCl (total 1 L) by monitoring the absorbance at 260 nm. The fractions exhibited the first peak were collected, precipitated with the acetone/ethanol (4:1), washed with acetone, and dried with N_2 flow. The amount of 1-(2-carboxyethyl)-NAD was estimated with a known millimolar extinction coefficients (ε) at 260 nm (18.0 $mM^{-1}cm^{-1}$) for the nucleotides (Sakamoto et al. 1986). The 1-(2-carboxyethyl)-NAD was then converted to N^6-(2-carboxyethyl)-NAD as described previously (Sakamoto et al. 1986). The 1-(2-carboxyethyl)-NAD was dissolved in a degassed 1.3% $NaHCO_3$ solution, reduced with sodium dithionite, then heated at 70°C with the pH kept at 11. After the rearrangement, the reaction mixture was put on a QAE-Sephadex A-25 column (2.0 × 50 cm) equilibrated with 20 mM Li_2CO_3/HCl buffer (pH 9.5) to remove sulfite, and the nucleotides were eluted with a linear gradient of 0-0.5 M LiC1 (total 1 L) by monitoring the absorbance at 266 nm. The fractions containing N^6-(2-carboxyethyl)-NAD were

collected, precipitated, washed, and dried as described above. The precipitate was dissolved in 200 mL of a 20 mM lithium acetate containing 0.8% acetaldehyde, and the pH was adjusted to 7. The N^6-(2-carboxyethyl)-NAD was enzymatically oxidized with yeast alcohol dehydrogenase (1,000 U) for 2 h, and then purified by DEAE-Sephadex A-25 chromatography as described above. The fractions exhibited the main peak were collected, precipitated, washed, and dried as described above.

Poly(ethylene glycol) linked NAD (PEG-NAD) was generated by the reaction between diamino PEG (MW 3,000) and N^6-(2-carboxyethyl)-NAD in the presence of a cross-linking reagent 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) as described previously (Katayama et al. 1983; Kishimoto et al. 1991). N^6-(2-carboxyethyl)-NAD (60 mg) and EDC (0.2 g) were added to a 2 mL of 66% diamino PEG solution (pH 6.0), and the pH was adjusted to 4.8 with HCl solution (2 M). The reaction mixture was kept overnight at 4°C. A known millimolar extinction coefficients (ε) of PEG-NAD at 266 nm (25.0 $mM^{-1}cm^{-1}$) was used to estimate the amount of PEG-NAD (Eguchi et al. 1986). Then, lyophilized PEG-NAD (15 μmol) was dissolved in CH_2Cl_2 (2 mL) containing 54 mM trimethylamine. 50 μmol of a bifunctional reagent 3,3'-(1,6-dioxo-1,6-hexanediyl)bis-2-thiazolidinethione (DHBT) was synthesized (Eguchi et al. 1986), and dissolved in CH_2Cl_2 (1 mL) then mixed with the PEG-NAD. After 1 h reaction at room temperature, the solvent CH_2Cl_2 was evaporated by N_2 flow. The residual compounds were resuspended in a 1 mL of phosphate buffer (pH 6.5) containing 0.1 M NaCl, and the solution was filtered by using a filter paper to remove the remaining precipitate. The NAD^+-preferring glucose 6-phosphate dehydrogenase from *Thermotoga maritima* (mG6PDH) and the NAD^+-preferring 6-phosphogluconate dehydrogenase from *T. maritima* (m6PGDH) were purified (See the "Purification of enzymes" section) to construct NAD-conjugated mG6PDH and NAD-conjugated mG6PDH, respectively. The NAD-conjugated dehydrogenases were constructed as described previously (Kishimoto et al. 1991; Nakamura et al. 1986). The activated PEG-NAD was mixed with mG6PDH or m6PGDH (0.5 μmol) dissolved in a 1 mL of phosphate buffer (pH 6.5) containing 0.1 M NaCl. After 10 h at room temperature, the reaction mixture was applied to a DEAE-Sephadex A-50 column (2.0 × 20 cm) equilibrated with a 20 mM phosphate buffer (pH 6.5) containing 0.1 M NaCl. The column was washed with same buffer, and then the resulting NAD-conjugated dehydrogenases were eluted with an NaCl gradient of 0.1-1.0 M (total 250 mL). Fractions showed the absorbance at both 266 and 280 nm were collected, washed with 0.1 M HEPES buffer (pH 7.5), then concentrated by ultrafiltration with 10,000 MWCO Amicon Ultra centrifugal filter. The amount of NAD-conjugated dehydrogenases was determined by using the known millimolar extinction coefficients (ε) of PEG-NAD at 266 nm (25.0 $mM^{-1}cm^{-1}$). On the other hand, the enzyme concentrations were determined by the Bradford assay using bovine serum albumin (BSA) as a standard.

G6PDH activity assay

Glucose 6-phosphate dehydrogenase (G6PDH) activities of wild-type G6PDH, its mutated enzymes, and NAD-conjugated mG6PDH were determined in 0.1 M HEPES buffer (pH 7.5) containing 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 2 mM G6P, 2 mM $NAD(P)^+$ at 80°C for 5 min. The NAD(P)H formation rates were measured at 340 nm with Biomate 3 UV-visible spectrophotometer (Thermo Scientific, Waltham, MA). The enzyme unit was defined as 1 μmole of NAD(P)H produced per min. To determine kinetic parameters on coenzymes, the enzyme activity was also measured with varying concentrations of $NAD(P)^+$ (2 μM to 10 mM). K_m and k_{cat} values for $NAD(P)^+$ were calculated by nonlinear regression

fitting to the Michaelis-Menten equation. The results shown represent the means from triplicate independent experiments.

6PGDH activity assay

6-Phosphogluconate dehydrogenase (6PGDH) activities for m6PGDH (NAD^+-preferring 6PGDH from *T. maritima*) and NAD-conjugated m6PGDH were determined as described previously (Chen et al. 2016).

NAD(P)H-dependent BV oxidoreductase activity assay

NAD(P)H-dependent BV oxidoreductase activities of NROR ($NADP^+$-dependent rubredoxin oxidoreductase from *P. furiosus*) and a new DI from *T. maritima* were determined at 50°C by measuring the absorbance increase at a wavelength of 578 nm with a UV-visible spectrophotometer. The diaphorase reactions were carried out in an anaerobic screw-cap IR quartz cuvette with a degassed 0.1 M HEPES buffer (pH 7.5) containing NROR or DI (0.1-5 μg), 0.5 mM NAD(P)H, and oxidized benzyl viologen (BV^{2+}, 0.5 mM of each). The specific activity was calculated with a known millimolar extinction coefficient (ε) for reduced benzyl viologen (BV^+) at A_{578}, 8.65 $mM^{-1}cm^{-1}$. The activity was expressed as unit (U) per mg enzyme, while 1 U was defined as 1 μmole BV reduced per min. The results shown represent the means with standard deviations from triplicate experiments.

In vitro synthetic enzymatic biosystem-based reactions for H_2 generation

The enzymatic reactions to examine maximum H_2 generation rates from glucose 6-phosphate (G6P) were carried out in a bioreactor at 50 or 80°C. Because all purified enzymes used for H_2 generation were stored in 50% (wt/wt) glycerol at -80°C, appropriate volumes of each enzyme were combined and diluted to less than 0.1% glycerol by the addition of 0.1 M HEPES buffer (pH 7.5) and finally concentrated with 10,000 MWCO Amicon Ultra centrifugal filter. For the first comparisons (**Figure 3d**), reaction mixtures (2 mL) were adjusted to a composition of 0.1 M HEPES buffer (pH 7.5), 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 8 mM NAD^+, 0.1 M G6P, 2.5 g/L (5.8 μM) SHI, and 1.0 g/L (17 μM) of the selected TmG6PDH mutant exhibits dual coenzyme preference (mG6PDH, the S33E/R65M/T66S mutant). In addition, 4.0 g/L DI (0.1 mM) and 0.6 mM BV were added to the reaction mixture to examine an effect of biomimetic electron transport chain on the H_2 productivity. Alternatively, 4.0 g/L (0.1 mM) BV-conjugated DI (harboring 0.6 mM of BV moiety) was added to replace non-conjugated DI and free BV. For high-speed H_2 production with G6P (**Figure 4b**), the reaction temperature was increased to 80°C and the composition of reaction mixture (2 mL) was changed to 0.1 M HEPES buffer (pH 7.5), 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 8 mM NAD^+, 0.2 M G6P, 2.5 g/L SHI, 1.0 g/L mG6PDH, 1.0 g/L BV-conjugated DI (25 μM DI, 0.15 mM of BV moiety). Alternatively, 1.0 g/L (18 μM) NAD^+-preferring Tm6PGDH (m6PGDH) (Chen et al. 2016) and 0.1 g/L (3.7 μM) 6PGL were additionally added to the reaction mixture to generate two moles of NADH from one mole of G6P.

For starch-powered rapid H_2 production at 80°C via the entire enzymatic pathway illustrated in **Figure 5a**, a kind of soluble starch, maltodextrin (0.1 M or 0.4 M in glucose equivalent) was used as a carbohydrate source. The final reaction mixture (0.5 mL) with concentrated enzymes and maltodextrin was adjusted to 100 mM HEPES buffer (pH 7.5) containing 50 mM sodium phosphate (pH 7.5), 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 8 mM NAD^+, and 0.5 mM thiamine pyrophosphate. The amounts of enzymes added for these reactions are listed in **Table 1**.

In order to examine the effect of conjugate formation between NAD and dehydrogenase, the composition of reaction mixture was further changed. To measure H_2 productivities under varying NAD^+ concentrations (**Figure S4**), mG6PDH or NAD-conjugated G6PDH (0.1 mM) was added to 100 mM HEPES buffer (pH 7.5) containing 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 0.1 M G6P, 2.5 g/L SHI, 2.0 g/L BV-conjugated DI (50 μM DI, 0.3 mM of BV moiety), and varying concentrations of free NAD^+. To determine a total turnover number (TTN) of the NAD-based enzymatic pathways (**Figure 6**), long-term H_2 production experiments were conducted as follows. All concentrated enzymes plus two NAD-conjugated dehydrogenases (0.1 mM of each) were added to a standard reaction mixture (0.5 mL) of 100 mM HEPES buffer (pH 7.5) containing 50 mM sodium phosphate (pH 7.5), 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 0.5 mM thiamine pyrophosphate (**Table S4**). The substrate, maltodextrin (0.4 M in glucose equivalent) was added several times to the reaction mixture during the fed-batch reaction. As a control reaction, free NAD^+ (0.1 mM) and non-conjugated dehydrogenases (mG6PDH and m6PGDH, 50 U of each) were added to replace two NAD-conjugated enzymes.

Systems for continuous H_2 detection

Continuous H_2 measurement was conducted in a continuous flow system purged with 50 mL/min ultrapure N_2 gas (Airgas) as described previously (Rollin et al. 2015; Rollin et al. 2016) H_2 evolution was detected with a tin oxide thermal conductivity H_2 sensor (TGS 821, Figaro USA Inc., Arlington Heights, IL) equipped with a gas-tight flexible gas line (Rollin et al. 2015; Rollin et al. 2016). Temperature of bioreactor and condenser were controlled by refrigerated/heated circulating baths (NESLAB RTE7, Thermo Scientific; Isotemp 3016D, Fisher Scientific). To conduct H_2 generation experiments at 50°C, the temperature of refrigerated/heated circulating baths for bioreactor and condenser were adjusted to 50.5°C and 21°C, respectively. For the H_2 production at 80°C, the temperature setting of bioreactor and condenser were changed to 81°C and 18°C, respectively. Data acquisition was carried out with a USB-6210 device (National Instruments, Austin, TX) and analyzed by Lab-View SignalExpress 2009 (National Instruments). The H_2 signals were calibrated by using in-line mass flow controllers (Aalborg) and ultrapure H_2 gas (Airgas), as described previously (Rollin et al. 2015; Rollin et al. 2016).

Cost analysis of enzymatic hydrogen production

A detailed economic analysis was conducted by using the H2A model based on DOE Version 3.1. According to the distributed hydrogen production H2A model, the plant will have a H_2 production capacity of 1,500 kg of H_2 per day. The mass balance of H_2 production has been updated: 1,500 kg H_2 will be produced from 10,125 kg of starch and 7,875 kg of water catalyzed by 15 kg of enzymes needed.

The production costs of H_2 include three major parts: (1) carbohydrate costs, (2) enzyme costs and (3) coenzyme costs as well as other expenditures – CapEx. Several key inputs are (1) carbohydrate costs ($0.30/kg starch, current price), (2) prices of enzyme cocktail of approximately $200/kg (predicted with experimental data supported), (3) one kg of enzyme can produce 100 kg of H_2 (predicted with experimental data supported). Therefore, one kg of hydrogen production costs includes $2.025 from starch and $2.00 from enzymes.

Table S3 presents NADP, NAD, NMN and NR cost as well as their TTN values based on experimental data (5,000 for NADP and NAD at 80°C; 110,000 for NAD-conjugate at 80°C) and

predicted data (200,000 for NMN and NR). As a result, coenzyme cost for the production of one kg of hydrogen is estimated to $335, $99.5, $4.82, $0.42 and $0.13 via NADP, NAD, NAD-conjugate, NMN and NR, respectively. This analysis strongly suggests that the most important work for decreasing the *in vitro* biohydrogen production costs is coenzyme engineering when issues of enzyme production costs and stability are addressed. Considering several-fold higher TTN values achieved later (by decreasing reaction temperature or reaction condition optimization), we estimate that *in vitro* biohydrogen production costs including substrate, enzymes and coenzymes could be less than $5.35/kg H_2 when NAD-conjugate, NMN, or NR was used by FY2020. If we used NADP, the production cost would be too costly.

Supplementary Tables and Figures

Table S1. Kinetic parameter for G6PDH activity at 80°C

	NAD^+		$NADP^+$		$(k_{cat}/K_m)^{NAD}$ / $(k_{cat}/K_m)^{NADP}$
	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	
WT	3.6 ± 0.3	9.5 ± 0.4	0.056 ± 0.007	18.5 ± 0.6	0.008 ± 0.001
S33E	0.53 ± 0.11	7.5 ± 0.7	0.19 ± 0.04	15.7 ± 0.8	0.19 ± 0.08
R65M/T66S	0.27 ± 0.05	9.2 ± 0.5	0.13 ± 0.03	18.6 ± 0.4	0.23 ± 0.02
Best Mutant (mG6PDH) (S33E/R65M/T66S)	0.12 ± 0.02	8.4 ± 0.6	0.43 ± 0.08	14.0 ± 0.9	2.1 ± 0.3

Table S2. Specific activity of BCV-DI conjugate at 50°C

Electron donor (Concentration)	**Enzyme / Electron mediator** (Concentration)	**Specific activity** (U/mg protein)
NADH (2 mM)	DI (25 μM) / BV (153 μM)	8.4 ± 1.3
	BCV-DI conjugate (25/153 μM*)	22.5 ± 2.9
NADPH (2 mM)	DI (25 μM) / BV (153 μM)	0.7 ± 0.2
	BCV-DI conjugate (25/153 μM*)	3.1 ± 0.8

*The molar concentrations of DI and BV in the conjugate were normalized by the Bradford protein assay and the extinction coefficient of fully reduced BV (with dithionite), respectively.

Table S3. Production cost analysis of enzymatic H2 production due to coenzyme costs

Coenzyme cost influence							Estimated cost (FY20)
Name	Cost	MW	TTN	Coenzyme	TTN	Coenzyme	
	($/kg)	(g/mol)	(now)	($/kg H2)	(future)	($/kg H2)	($/kg H2)
NADP	4500	744.4	5000	334.98	20000	83.75	$ 87.77
NAD	1500	663.4	5000	99.51	20000	24.88	$ 28.90
NAD-conjugate	1600	663.4	110000	4.82	400000	1.33	$ 5.35
NMN	500	334.2	200000	0.42	400000	0.21	$ 4.23
NR	200	255.2	200000	0.13	400000	0.06	$ 4.09

According to **Table 1**, it is expected that in vitro biohydrogen production cost could be decreased to below \$10/kg H_2, including other expenditures such as CapEx (i.e., bioreactors, compressors, CO_2/H_2 separation membrane) and operation costs (not including consumables). The H2A excel table was used to calculate current enzymatic H_2 production costs and future production costs.

Table S4. Enzymes used for the long-term H_2 production (Figure 6).

Enzyme (Abbreviation)	E.C. #	Gene Source	Sp. Act. at 50°C (U/mg)	Enzyme Loading (U/mL)	Reference
α-glucan phosphorylase (αGP)	2.4.1.1	*Thermotoga maritima*	20[1]	5	(You et al. 2017)
Phosphoglucomutase (PGM)	5.4.2.2	*Thermococcus kodakarensis*	100[1]	5	(You et al. 2017)
NAD-conjugated mG6PDH	1.1.1.49	*T. maritima*	**12**[2]	**50**[6]	This study
6-phosphogluconolactonase (6PGL)	3.1.1.31	*T. maritima*	230[3]	5	(Zhu and Zhang 2017)
NAD-conjugated m6PGDH	1.1.1.44	*T. maritima*	**30**[1]	**50**[6]	This study
Ribose 5-phosphate isomerase (RPI)	5.3.1.6	*T. maritima*	300	2	(Zhu and Zhang 2017)
Ribulose 5-phosphate 3-epimerase (RuPE)	5.1.3.1	*T. maritima*	66	2	(Zhu and Zhang 2017)
Transketolase (TK)	2.2.1.1	*Thermus thermophilus*	5.3	2	(Zhu and Zhang 2017)
Transaldolase (TAL)	2.2.1.2	*T. maritima*	3.9	2	(Zhu and Zhang 2017)
Triose-phosphate isomerase (TIM)	5.3.1.1	*T. thermophilus*	450	1	(You and Zhang 2014)
Aldolase (ALD)	4.1.2.13	*T. thermophilus*	16	1.5	(You and Zhang 2014)
Fructose 1,6-bisphosphatase (FBP)	3.1.3.11	*T. maritima*	6	2	(Zhu and Zhang 2017)
Phosphoglucose isomerase (PGI)	5.3.1.9	*T. thermophilus*	190[1]	2	(Ninh et al. 2015)
NAD-dependent BV-oxidoreductase (DI)	1.18.1.4	*T. maritima*	20[4]	20	This study
[NiFe]-Hydrogenase (SHI)	1.12.1.3	*Pyrococcus furiosus*	121[5]	300	(Chandrayan et al. 2015)

[1]Specific activity was measured at 70°C

[2]Specific activity was measured at 80°C

[3]Specific activity was measured at 23°C

[4]Specific activity of BV-conjugated DI (1 g/L of BV-conjugated DI contains 0.15 mM BV)

[5]Specific activity of SHI was 121 U/mg based on reduced methyl viologen at 80°C (Chandrayan et al. 2015b).

[6]Approximately 0.1 mM of total conjugated NAD^+ was used by adding 4.2 g/L (0.7 mM) NAD-conjugated mG6PDH and 1.65 g/L (0.3 mM) NAD-conjugated m6PGDH.

Figure S1.

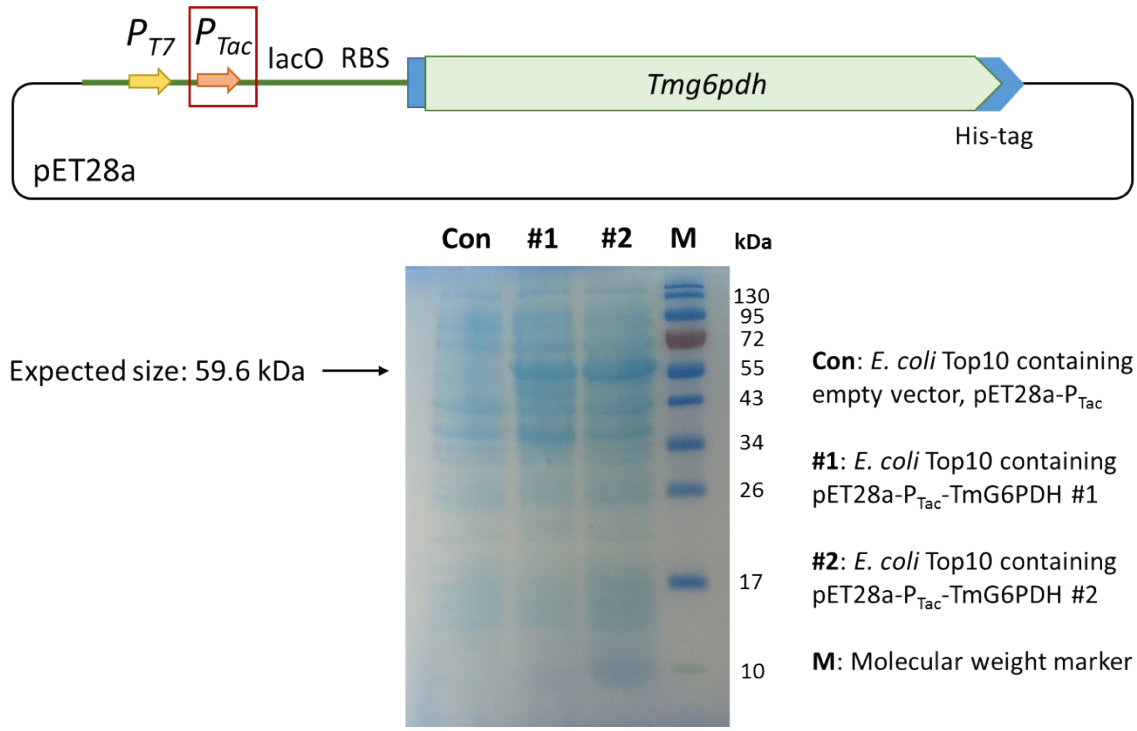

Figure S1. Dual-promoter controlling TmG6PDH expression in *E. coli* TOP10 and BL21(DE3). This plasmid could be utilized for the purpose of screening and purification without sub-cloning procedure.

Figure S2.

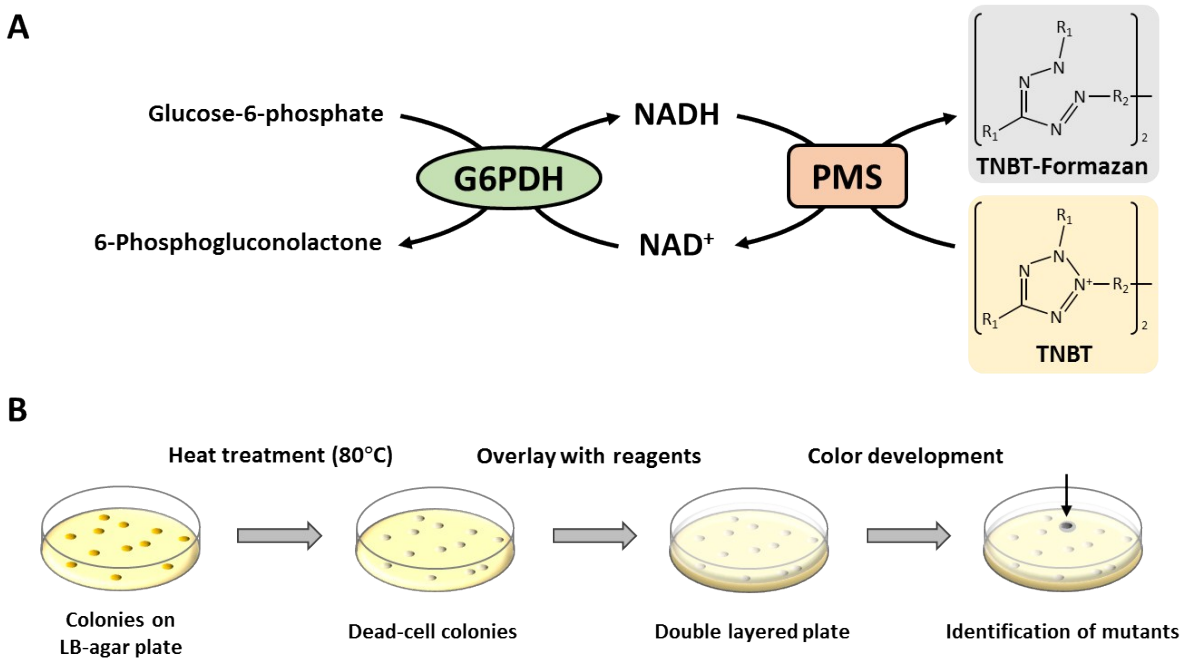

Figure S2. Scheme of the high-throughput screening method for the identification of NAD-preferring G6PDH mutants. TNBT, tetranitroblue tetrazolium chloride (R_1 = 4-nitrobenzeyl, R_2 = 2-methoxybenzyl); PMS, phenzine methosulfate.

Figure S3.

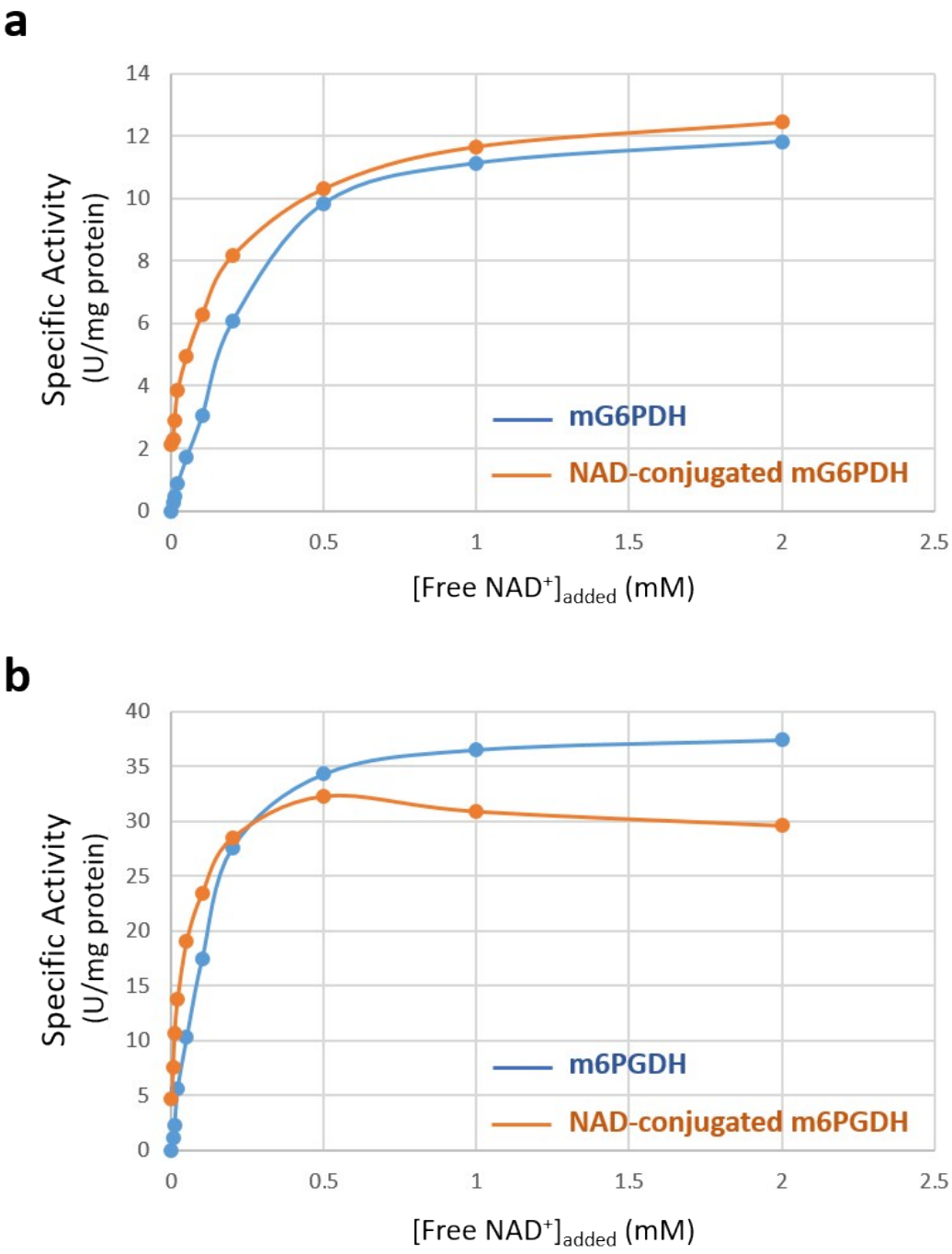

Figure S3. Specific activities for mG6PDH and NAD-conjugated mG6PDH (a) and m6PGDH and NAD-conjugated m6PGDH (b) determined with 0.02 mM of enzyme (or NAD-conjugated enzyme) in the presence of varying concentrations of free NAD^+.

Figure S4.

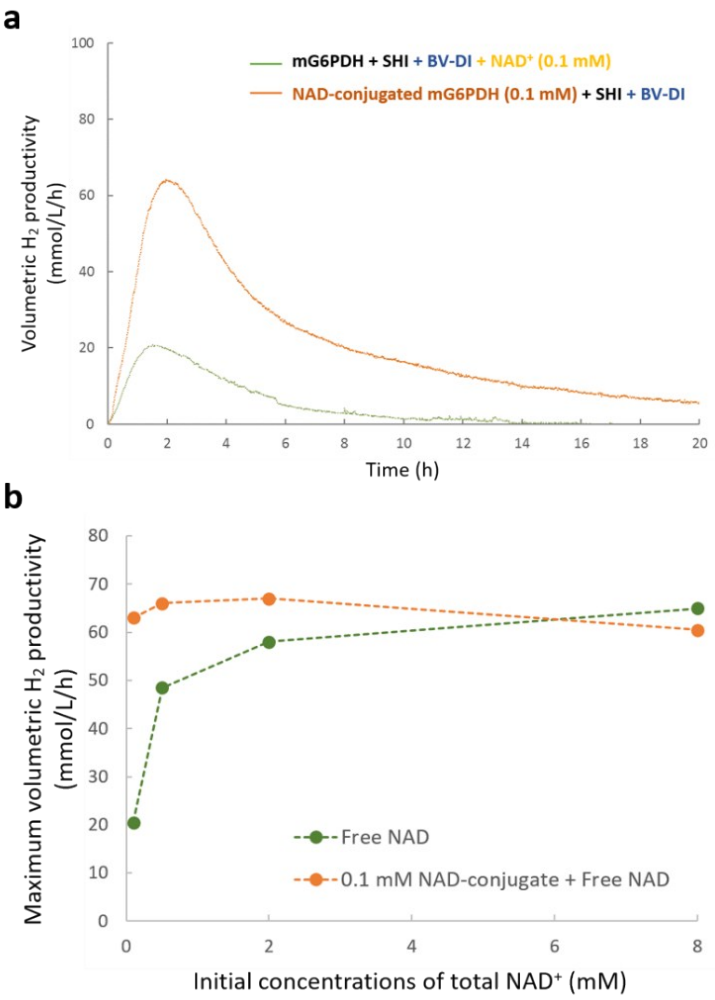

Figure S4. H_2 generation profiles of mG6PDH and NAD-conjugated mG6PDH at the low NAD^+ concentration (0.1 mM) (a) and their maximum H_2 productivities under the varying NAD^+ concentrations (b). H_2 productivity was measured with a partial pathway comprised of mG6PDH (or NAD-conjugated mG6PDH), BV-conjugated DI, and SHI using 0.1 M G6P and varying concentrations of free NAD^+ (0.1 ~ 8 mM). The reaction using 0.1 mM of NAD-conjugated mG6PDH exhibited around three times higher H_2 productivity than that of the reaction using free NAD^+ at the same concentration (**Fig. S4a**). Based on a product cost analysis, the estimated H_2 production cost could be lowered nearly three times when $NADP^+$ is replaced to NAD^+ in the enzymatic pathway (**Table S3**). The maximum volumetric H_2 productivities of the control reactions resulted in an increasing pattern according to the increase of initial NAD^+ concentrations (**Fig. S4b**). In contrast, when NAD-conjugated mG6PDH was applied to replace mG6PDH, a relatively constant level of H_2 productivity was observed irrespective of the NAD^+ concentrations added (**Fig. S4b**).

References

Chandrayan SK, Wu C-H, McTernan PM, Adams MWW. 2015. High yield purification of a tagged cytoplasmic [NiFe]-hydrogenase and a catalytically-active nickel-free intermediate form. Protein Expr. Purif. 107:90-94.

Chen H, Zhu ZG, Huang R, Zhang Y-H. 2016. Coenzyme engineering of a hyperthermophilic 6-phosphogluconate dehydrogenase from $NADP^+$ to NAD^+ with its application to biobatteries. Sci. Rep. 6:36311.

Dinh TH, Lee SC, Hou CY, Won K. 2016. Diaphorase-viologen conjugates as bioelectrocatalysts for NADH regeneration. Journal of The Electrochemical Society 163(6):H440-H444.

Eguchi T, Iizuka T, Kagotani T, Lee JH, Urabe I, Okada H. 1986. Covalent linking of poly (ethyleneglycol)-bound NAD with *Thermus thermophilus* malate dehydrogenase. The FEBS Journal 155(2):415-421.

Hansen DC, Pan Y, Stockton J, Pitt WG, Wheeler DR. 2012. Cyclic voltammetry investigation of organic species considered for use as catalysts in direct-carbohydrate fuel cells. J. Electrochem. Soc. 159(11):H834-H841.

Hansen T, Schlichting B, Schonheit P. 2002. Glucose-6-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: expression of the g6pd gene and characterization of an extremely thermophilic enzyme. FEMS Microbiol. Lett. 216(2):249-253.

Hong J, Wang Y, Ye X, Zhang Y-HP. 2008. Simple protein purification through affinity adsorption on regenerated amorphous cellulose followed by intein self-cleavage. J. Chromatogr. A 1194(2):150-154.

Huang R, Chen H, Zhong C, Kim JE, Zhang Y-HP. 2016. High-throughput screening of coenzyme preference change of thermophilic 6-phosphogluconate dehydrogenase from $NADP^+$ to NAD^+. Sci. Rep. 6:32644.

Katayama N, Urabe I, Okada H. 1983. Steady-State Kinetics of Coupled Two-Enzyme Reactor with Recycling of Ploy (ethylene glycol)-Bound NAD. The FEBS Journal 132(2):403-409.

Kim E-J, Adams M, Wu C-H, Zhang YHP. 2016. Exceptionally high rates of biological hydrogen production by biomimetic *in vitro* synthetic enzymatic pathways. Chemistry 22:16047–16051.

Kishimoto T, Itami M, Yomo T, Urabe I, Yamada Y, Okada H. 1991. Improved methods for the preparation of N6-(2-carboxyethyl)-NAD and poly(ethylene glycol)-bound NAD(H). Journal of Fermentation and Bioengineering 71(6):447-449.

Nakamura A, Urabe I, Okada H. 1986. Anchimeric assistance in the intramolecular reaction of glucose-dehydrogenase-polyethylene glycol NAD conjugate. Journal of Biological Chemistry 261(36):16792-16794.

Ninh PH, Honda K, Sakai T, Okano K, Ohtake H. 2015. Assembly and multiple gene expression of thermophilic enzymes in *Escherichia coli* for *in vitro* metabolic engineering. Biotechnol. Bioeng. 112:189-196.

Rollin JA, Martin del Campo J, Myung S, Sun F, You C, Bakovic A, Castro R, Chandrayan SK, Wu C-H, Adams MWW and others. 2015. High-yield hydrogen production from biomass by *in vitro* metabolic engineering: Mixed sugars coutilization and kinetic modeling. Proc. Nat. Acad. Sci. USA 112:4964-4969.

Rollin JA, Ye XH, Martin dCJS, Adams MWW, Zhang Y-HP. 2016. Novel hydrogen detection apparatus along with bioreactor Systems. Adv. Biochem. Eng. Biotechnol. 152:35-51.

Sakamoto H, Nakamura A, Urabe I, Yamada Y, Okada H. 1986. Analysis of improvement of the synthesis of N6-(2-carboxyethyl)-NAD. Journal of Fermentation Technology 64(6):511-516.

You C, Shi T, Li Y, Han P, Zhou X, Zhang Y-HP. 2017. An *in vitro* synthetic biology platform for the industrial biomanufacturing of myo-inositol from starch. Biotechnol. Bioeng.:Accepted for publication.

You C, Zhang YHP. 2014. Annexation of a high-activity enzyme in a synthetic three-enzyme complex greatly decreases the degree of substrate channeling. ACS Syn. Biol. 3:380-386.

Zhang Y-HP, Cui J, Lynd LR, Kuang LR. 2006. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules 7(2):644-648.

Zhu ZG, Zhang Y-HP. 2017. *In vitro* metabolic engineering of bioelectricity generation for the complete oxidation of glucose. Metab. Eng. 39:110-116.