In Vitro Synthetic Biology Platform and Protein Engineering for Biorefinery

Jae-Eung Kim

Dissertation submitted to the faculty of
the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Biological Systems Engineering

Yi Heng Percival Zhang, Chair
Xueyang Feng, Co-Chair
Jianyong Li
Ryan S. Senger
Warren C. Ruder

May 5th, 2017
Blacksburg, Virginia

Keywords: biomanufacturing 4.0 , directed evolution, D-xylose, epimerase, hydrogen production, *in vitro* synthetic biology, L-arabinose, protein engineering, starch phosphorylation, zero-calorie sweetener

Abstract

In Vitro Synthetic Biology Platform and Protein Engineering for Biorefinery

Jae-Eung Kim

In order to decrease our dependence on non-renewable petrochemical resources, it is urgently required to establish sustainable biomass-based biorefineries. Replacing fossil fuels with renewable biomass as a raw feedstock for the production of chemicals and biofuels is a main driving force of biorefinering. Almost all kinds of biomass can be converted to biochemicals, biomaterials and biofuels via continuing advances on conversion technologies. In vitro synthetic biology is an emergent biomanufacturing platform that circumvents whole cell's constraints, so that it can implement some biotransformations better than whole-cell fermentation spending a significant fraction of energy and carbon sources for cellular duplication and side-product formation. In this work, the in vitro synthetic (enzymatic) biosystem is used to produce a future carbon-neutral transportation fuel, hydrogen gas, and two high-value chemicals, a sugar phosphate and a highly marketable sweetener, representing a new portfolio for new biorefineries.

Hydrogen gas is a promising energy carrier as a transportation fuel, offering a high energy conversion efficiency via fuel cells, nearly zero pollutants produced to end users, and high mass-specific and volumetric energy densities compared to rechargeable batteries. Distributed production of cost-competitive green hydrogen will be vital to the hydrogen economy. Substrate costs account for a major portion of the production cost for low-value bulk biocommodities, such as hydrogen. The reconstitution of 17 thermophilic enzymes enabled to construct an artificial enzymatic pathway converting all glucose units of starch, regardless of the branched and linear contents, to hydrogen gas at a theoretic yield (i.e., 12 H_2 per glucose), three times of the theoretical yield of dark microbial fermentation. Using a biomimetic electron transport chain, a maximum volumetric productivity was increased by more than 200-fold to 90.2 mmol of H_2/L/h at a high starch concentration from the original study in 2007.

In order to promote economics of biorefineries, the production of a sugar phosphate and a fourth-generation sweetener is under development. D-xylulose 5-phosphate (Xu5P), which cannot be prepared efficiently by regular fermentation due to the negatively charged and hydrophilic phosphate groups, was synthesized from D-xylose and polyphosphate via a minimized two-enzyme system using a promiscuous activity of xylulose kinase. Under optimized conditions, 32 mM Xu5P was produced from 50 mM xylose and polyphosphate, achieving a 64% conversion yield, after 36 h at 45 °C. L-arabinose, a FDA-approved zero-calorie sweetener, can be produced from D-xylose via a novel enzymatic pathway consisting of xylose isomerase, L-arabinose isomerase and xylulose 4-epimerase (Xu4E), a hypothetical monosaccharide 4-epimerase that can convert D-xylulose to L-ribulose. Xu4E activities were discovered for the first time from substrate promiscuity of some natural 4-epimerases. We applied three-round directed evolution to increase the catalytic function of carbon 4-epimerization on D-xylulose by more than 29-fold from the wild-type enzyme. Together, these results demonstrate that the in vitro synthetic biosystem as a feasible biomanufacturing platform has great engineering, and can be used to

convert renewable biomass resources to a spectrum of marketable products and renewable energy.

Remaining challenges of the in vitro synthetic biosystems are being addressed, for example, decreasing enzyme production costs, prolonging enzyme lifetime, engineering biomimetic coenzymes to replace natural coenzymes, and so on. This in vitro synthetic biology platform would become a cornerstone technology for biorefinery industries and advanced biomanufacturing (Biomanufacturing 4.0).

General Audience Abstract

In Vitro Synthetic Biology Platform and Protein Engineering for Biorefinery

Jae-Eung Kim

The carbon cycle is the circulation and transformation of carbon back and forth between living things and the environment. With the fixed amount of carbon dioxide in the atmosphere, the carbon cycle has been in the balance of exchanges between living things and the environment. As we evolve with increasing demand on crude oil, however, significant amounts of carbon are being released into the atmosphere much faster than they would have been released naturally. This rapid release is the primary cause of currently observed global warming. In order to decrease our dependence on petrochemical products, the biorefinery was introduced as the sustainable processing of biomass into a spectrum of alternatives to products from petrochemical refineries. Almost all kinds of biomass can be converted to biochemicals, biomaterials and biofuels via continuing advances on conversion technologies. In vitro synthetic biology is an emergent biomanufacturing platform that circumvents whole cell's constraints, so that it can implement some biotransformations better than whole-cell fermentation spending a significant fraction of energy and carbon sources for cellular duplication and side-product formation. In this work, the in vitro synthetic (enzymatic) biosystem is used to produce a future carbon-neutral transportation fuel, hydrogen gas, and two high-value chemicals, a sugar phosphate and a highly marketable sweetener, representing a new portfolio for new biorefineries.

Hydrogen gas is a promising energy carrier as a transportation fuel, offering a high energy conversion efficiency via fuel cells, nearly zero pollutants produced to end users, and high mass-specific and volumetric energy densities compared to rechargeable batteries. Distributed production of cost-competitive green hydrogen will be vital to the hydrogen economy. We demonstrated an in vitro 17-thermophilic enzyme pathway that can convert all glucose units of starch to hydrogen a theoretic yield, which is three times of the theoretical yield from dark microbial fermentation. D-xylulose 5-phosphate (Xu5P), which cannot be prepared efficiently by regular fermentation due to the negatively charged and hydrophilic phosphate groups, was synthesized from D-xylose and polyphosphate via a minimized two-enzyme system using a promiscuous activity of xylulose kinase. This minimal in vitro enzymatic pathway was optimized for improved conversion yield and productivity. L-arabinose, a FDA-approved zero-calorie sweetener, was also produced from D-xylose via a novel enzymatic pathway consisting of xylose isomerase, L-arabinose isomerase and hypothetical enzyme xylulose 4-epimerase (Xu4E), a monosaccharide 4-epimerase that can convert D-xylulose to L-ribulose. Xu4E activities due to substrate promiscuity of some natural 4-epimerases were discovered for the first time. Three rounds of directed evolution have been conducted to increase the catalytic function of carbon 4-epimerization on D-xylulose. As the result, the catalytic activity of Xu4E was improved by more than 29-fold from the wild-type enzyme.

Acknowledgement

First of all, I would like to express my special appreciation and thanks to my advisor Professor Dr. Y.-H. Percival Zhang, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been priceless.

I would also like to thank my committee members, Professor Dr. Jianyong Li, Professor Dr. Ryan S. Senger, Professor Dr. Warren C. Ruder, and Professor Dr. Xueyang Feng for serving as my committee members. I also want to thank you for letting my defense be a valuable and enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

A special thanks to my family. Words cannot express how grateful I am to my mother and father for all of the sacrifices that you've made on my behalf. Your prayer and supports for me were what sustained me thus far. I would also like to thank all of my friends, especially Dr. Eui-Jin Kim, Dr. Chen Hui, Huang Rui, SeungBeum Chris Suh, Seungmo Kim, Dongseok Shin, Seunghwan Kim and Aram Lee, who supported me in researching and writing, and incented me to strive towards my goal.

Finally, I would like to acknowledge the financial support I received at Virginia Tech and from graduate research assistantships.

Table of Contents

List of Figures

Chapter 2

Chapter 2-1

Chapter 2-2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendix a

List of Tables

Chapter 2

Chapter 2-1

Chapter 2-2

Chapter 6

Appendix a

Chapter 1. Introduction

The need for alternatives to products from petrochemical refineries is a well-established consensus to decrease environmental impacts and pursue sustainability. The development of economically viable biorefineries is the most important step toward a sustainable bioeconomy. As the analogue of a petroleum refinery where crude oil is processed to produce various products, a biorefinery is a process using biomass as a feedstock to produce biofuels, biochemicals, biomaterials, food and feed (Cherubini, 2010). In 2009, the International Energy Agency Bioenergy Task 42 defined the biorefinery as "the sustainable processing of biomass into a spectrum of marketable products and energy" (Jong et al., 2009). As its definition implies, two different driving factors are motivating biorefineries: energy-driven biorefineries, highlighting the sustainability aspect; and product-driven biorefineries, highlighting the establishment of marketable products to meet economic goals (Bozell and Petersen, 2010). The energy-driven biorefineries have been addressed by the effort of biofuel production, mainly focusing on (corn or sugarcane) ethanol or biodiesel fermentation over the last decades. The aim of biofuel production is to displace a portion of the current transportation fuels, such as gasoline and diesel. However, fuels are a low-value product and biofuel production itself cannot provide high-enough revenues to justify investment in biorefineries, even though its market size is large (Bozell and Petersen, 2010; Rollin et al., 2013). Consequently, the production of high-value biocommodities is required to leverage continuous biofuel productions and sustain biorefinery operations. In 2004, the US Department of Energy (DOE) released a report listing the most promising bio-based products from carbohydrates for biorefinery industries (Werpy et al., 2004). Since the publication of DOE's primary report, industrial evaluation and technical evolution have been continuously making revisions on the list of chemicals for biorefineries (Bozell and Petersen,

2010; Choi et al., 2015; Zhang et al., 2016). With many of these products already commercialized for industrial applications, biorefineries need continuously to identify new valuable products with large market potentials, and develop new conversion technologies and processes.

In vitro synthetic biology as an emergent biomanufacturing platform is the construction of purified enzymes that self-assembles synthetic enzymatic pathways which avoid cellular constraints, such as cell membranes, complicated cellular regulation, and bioenergetic limits (Rollin et al., 2013). Some constraints of the whole-cell fermentation, e.g., a loss of energy and carbon source due to cell duplication and side-product formation, can be relieved by reconstituting purified enzymes into *in vitro* synthetic (enzymatic) biosystems. In addition, the in vitro synthetic biosystem features several biomanufacturing advantages over whole-cell factories. These advantages include fast volumetric productivities or biomanufacturing rates (Hodgman and Jewett, 2012), high product yields and energy efficiencies (Opgenorth et al., 2014; Rieckenberg et al., 2014), easy product separation (Hodgman and Jewett, 2012), tolerance to toxic substrates or products (Guterl et al., 2012), broad reaction conditions (Panke et al., 2004), implementation of non-natural reaction (You et al., 2013), and great engineering flexibility in terms of using enzymes as highly exchangeable building biobricks (Zhang et al., 2010).

The advantages of in vitro synthetic biosystem as an emerging biomanufacturing platform are especially highlighted in producing chemicals that remain difficult to produce by whole-cell fermentation, e.g., sugar phosphates (Honda et al., 2010; Meyer et al., 2007). Due to the phosphate group, negatively charged and hydrophilic sugar phosphates cannot pass through

intact cell membranes, so that they cannot be fermented efficiently by regular microorganisms. D-xylulose 5-phosphate (Xu5P), a key sugar phosphate intermediate of the pentose phosphate pathway, is used for an essential substrate of enzyme assays (Lee et al., 2008; Wood, 1973) and the study of metabolic diseases (Iizuka and Horikawa, 2008; Shaeri et al., 2008). Despite of its importance, the commercial availability of Xu5P is limited to few vendors with the list selling price of more than $2,500 for 10 mg. The synthetic enzymatic pathway comprised of xylose isomerase and xylulokinase was designed to produce Xu5P, along with a third enzyme, polyphosphate kinase, responsible for in site ATP regeneration. Under the optimized condition, approximately 32mM Xu5P was produced from 50mM xylose, achieving 64% conversion yield, with polyphosphate after 36 h at 45 °C. Biosynthesis of less costly Xu5P could be highly feasible via this enzymatic pathway under the broad reaction conditions.

Along with the in vitro synthetic biology, a protein engineering approach, directed evolution, was used to produce L-arabinose from D-xylose via an enzymatic pathway involving a monosaccharide 4-epimerase discovered for the first time in this work. Protein engineering is a powerful tool to redesign proteins by improving or changing their properties (Renata et al., 2015). Over past decades, a dramatic increase in the amount of added sucrose and high fructose corn syrup (HFCS) has been correlated with increasing number of people in diabetes, obesity, cancer, and cardiovascular disease (Goran et al., 2013; Malik et al., 2006). Alternative zero-calorie sweeteners are highly required to replace sucrose and HFCS in the global market amounted for over $100 billion. L-arabinose is an US Food and Drug Administration (FDA)-approved zero-calorie natural sweetener with 50% sweetness of sucrose and the similar taste (Hao et al., 2015). In addition to the non-caloric sweetener, L-arabinose has dual functions of a sucrose neutralizer

and a potential prebiotic compound. Numerous epimerases were exploited to discover their promiscuous activities on D-xylulose as a substrate, then named it as D-xylulose 4-epimerase (Xu4E). Directed evolution was used to improve the specific activity of Xu4E by developing a high-throughput screening method based on a transcription-factor-driven biosensor. A novel three-enzyme pathway comprised of xylose isomerase (XI), Xu4E, and L-arabinose isomerase (AI) was constructed to enable the upgrade of D-xylose to L-arabinose.

Production of ethanol and biodiesel through the fermentation process has been the major stream of biofuel production for biorefineries. From the recent unprecedented success of electric vehicles in the automotive industry, it is expected to have a significant mechanism change initiated from combustion engines to electric motors to drive vehicles. Biorefineries producing sustainable energy for transportation fuels also need to re-draw their portfolio according to the impending changes in the powertrain of current vehicles. Hydrogen is a promising future energy carrier to replace fossil fuel-derived liquid fuels that are primarily used for transportation (Zhang, 2009; Zhang and Mielenz, 2011). Hydrogen offers a high energy conversion efficiency via fuel cells, nearly zero pollutants, and high mass-specific energy density compared to rechargeable batteries. Distributed production of cost-competitive green hydrogen gas from renewable energy sources will be vital to the hydrogen economy (Zhang, 2010).

This work emphasizes the use of in vitro synthetic biosystem as a biomanufacturing platform, and protein engineering approach for the production of marketable products and energy for future biorefineries. In the second chapter of this dissertation, the feasibility of enzymatic hydrogen production from carbohydrates is reviewed, including a discussion of future hydrogen

economy. Chapter 3 describes the complete enzymatic phosphorylation of starch with application to green hydrogen gas production at a theoretical yield. In chapter 4, a two-step enzymatic pathway for a sugar phosphate production is described. Chapter 5 introduces a facile method of random gene mutagenesis library construction for directed evolution. In chapter 6, a novel enzymatic pathway to produce a zero calorie sweetener, L-arabinose, from xylose is described with the protein engineering approach creating a new enzyme by expanding the confines of engineering biology. Chapter 7 includes a summary of this work and suggestions for future research directions.

References

Bozell JJ, Petersen GR. 2010. Technology development for the production of biobased products from biorefinery carbohydrates-the US Department of Energy's "Top 10" revisited. Green Chem 12(4):539-554.

Cherubini F. 2010. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. Energy Convers Manage 51(7):1412-1421.

Choi S, Song CW, Shin JH, Lee SY. 2015. Biorefineries for the production of top building block chemicals and their derivatives. Metab Eng 28:223-239.

Goran MI, Dumke K, Bouret SG, Kayser B, Walker RW, Blumberg B. 2013. The obesogenic effect of high fructose exposure during early development. Nat Rev Endocrinol 9(8):494-500.

Guterl JK, Garbe D, Carsten J, Steffler F, Sommer B, Reisse S, Philipp A, Haack M, Ruhmann B, Koltermann A and others. 2012. Cell-free metabolic engineering: Production of chemicals by minimized reaction cascades. ChemSusChem 5(11):2165-2172.

Hao L, Lu XL, Sun M, Li K, Shen LM, Wu T. 2015. Protective effects of L-arabinose in high-carbohydrate, high-fat diet-induced metabolic syndrome in rats. Food Nutr Res 59:28886.

Hodgman CE, Jewett MC. 2012. Cell-free synthetic biology: Thinking outside the cell. Metab Eng 14(3):261-269.

Honda K, Maya S, Omasa T, Hirota R, Kuroda A, Ohtake H. 2010. Production of 2-deoxyribose 5-phosphate from fructose to demonstrate a potential of artificial bio-synthetic pathway using thermophilic enzymes. J Biotechnol 148(4):204-207.

Iizuka K, Horikawa Y. 2008. ChREBP: A glucose-activated transcription factor involved in the development of metabolic syndrome. Endocr J 55(4):617-624.

Jong Ed, Langeveld H, Ree Rv. 2009. IEA Bioenergy Task 42 Biorefinery. Available at: www. iea-bioenergy. task42.

Lee JY, Cheong DE, Kim GJ. 2008. A novel assay system for the measurement of transketolase activity using xylulokinase from *Saccharomyces cerevisiae*. Biotechnol Lett 30(5):899-904.

Malik VS, Schulze MB, Hu FB. 2006. Intake of sugar-sweetened beverages and weight gain: A systematic review. Am J Clin Nutr 84(2):274-288.

Meyer A, Pellaux R, Panke S. 2007. Bioengineering novel *in vitro* metabolic pathways using synthetic biology. Curr Opin Microbiol 10:246-253.

Opgenorth PH, Korman TP, Bowie JU. 2014. A synthetic biochemistry molecular purge valve module that maintains redox balance. Nat Commun 5:4113.

Panke S, Held M, Wubbolts M. 2004. Trends and innovations in industrial biocatalysis for the production of fine chemicals. Curr Opin Biotechnol 15(4):272-279.

Renata H, Wang ZJ, Arnold FH. 2015. Expanding the enzyme universe: Accessing non-natural reactions by mechanism-guided directed evolution. Angew Chem Int Edit 54(11):3351-3367.

Rieckenberg F, Ardao I, Rujananon R, Zeng A-P. 2014. Cell-free synthesis of 1,3-propanediol from glycerol with a high yield. Eng Life Sci 14:380-386.

Rollin JA, Tam TK, Zhang Y-HP. 2013. New biotechnology paradigm: cell-free biosystems for biomanufacturing. Green Chem 15(7):1708-1719.

Shaeri J, Wright I, Rathbone EB, Wohlgemuth R, Woodley JM. 2008. Characterization of enzymatic D-xylulose 5-phosphate synthesis. Biotechnol Bioeng 101(4):761-767.
Werpy T, Petersen G, Aden A, Bozell J, Holladay J, White J, Manheim A, Eliot D, Lasure L, Jones S. 2004. Top value added chemicals from biomass: Volume 1 - Results of screening for potential candidates from sugars and synthesis gas. National Renewable Energy Laboratory - NREL.
Wood T. 1973. The preparation of xylulose 5-phosphate using transketolase. Prep Biochem Biotechnol 3(6):509-515.
You C, Chen H, Myung S, Sathitsuksanoh N, Ma H, Zhang X-Z, Li J, Zhang Y-HP. 2013. Enzymatic transformation of nonfood biomass to starch. Proc Natl Acad Sci USA 110(18):7182-7187.
Zhang Y-HP. 2009. A sweet out-of-the-box solution to the hydrogen economy: is the sugar-powered car science fiction? Energy Environ Sci 2(3):272-282.
Zhang Y-HP. 2010. Renewable carbohydrates are a potential high-density hydrogen carrier. Int J Hydrogen Energy 35(19):10334-10342.
Zhang Y-HP, Mielenz JR. 2011. Renewable Hydrogen Carrier - Carbohydrate: Constructing the Carbon-Neutral Carbohydrate Economy. Energies 4(2):254-275.
Zhang Y-HP, Sun J, Ma Y. 2016. Biomanufacturing: history and perspective. J Ind Microbiol Biotechnol:1-12.
Zhang Y-HP, Sun JB, Zhong JJ. 2010. Biofuel production by in vitro synthetic enzymatic pathway biotransformation. Curr Opin Biotechnol 21(5):663-669.

Chapter 2. Hydrogen production from carbohydrates

Chapter 2-1. High-yield production of biohydrogen from carbohydrates and water based on in vitro synthetic (Enzymatic) pathways

Kim, J.-E., and Zhang, Y.H.P. (2015). Chapter 3. High-yield production of biohydrogen from carbohydrates and water based on in vitro synthetic (Enzymatic) pathways. In Book "Hydrogen production from renewable resources", Springer Book Series - Biofuels and Biorefineries, Volume 5, pp. 77-94. Reproduced by permission of Springer.

Chapter 3
High-Yield Production of Biohydrogen from Carbohydrates and Water Based on In Vitro Synthetic (Enzymatic) Pathways

Jae-Eung Kim and Yi-Heng Percival Zhang

Abstract Distributed production of green and low-cost hydrogen from renewable energy sources is necessary to develop the hydrogen economy. Carbohydrates, such as cellulose, hemicellulose, starch, sucrose, glucose, and xylose, are abundant renewable bioresources and can provide the source of hydrogen. In this chapter, in vitro synthetic (enzymatic) pathways that overcome the limiting yields of hydrogen-producing microorganisms are discussed. These in vitro synthetic pathways produce hydrogen with theoretical yields from polymeric and monomeric hexoses or xylose with water of 2 mol of hydrogen per carbon molecule of carbohydrate. In the past years, hydrogen production rate of in vitro synthetic enzymatic pathways has been improved to 150 mmol/L/h by 750-fold through systematic optimization. All of the thermostable enzymes used in the pathways have been recombinantly produced in *E. coli*, and some of them are immobilized for enhanced stability and simple recycling. Redox enzymes are being engineered to work on low-cost and highly stable biomimetic coenzymes. It is expected that low-cost green hydrogen can be produced at $2.00/kg hydrogen in small-sized atmospheric pressure bioreactors in the future.

Keywords Carbohydrate • Hydrogen production • Synthetic enzymatic pathway • In vitro synthetic biosystem • Innovative biomanufacturing • Systems biocatalysis

J.-E. Kim
Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall, Blacksburg, VA 24061, USA

Y.-H.P. Zhang (✉)
Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall, Blacksburg, VA 24061, USA

Cell Free Bioinnovations Inc., 2200 Kraft Drive, Suite 1200B, Blacksburg, VA 24060, USA

Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Avenue, Tianjin Airport Economic Area, Tianjin 300308, China
e-mail: ypzhang@vt.edu

Z. Fang et al. (eds.), *Production of Hydrogen from Renewable Resources*, Biofuels and Biorefineries 5, DOI 10.1007/978-94-017-7330-0_3

3.1 Introduction

3.1.1 Hydrogen

Dihydrogen gas is a colorless, odorless, but very flammable diatomic molecule. Later on throughout this chapter, hydrogen is used to mean dihydrogen gas for convenience. Hydrogen atoms widely exist in natural inorganic and organic matters, such as water, hydrocarbons, and carbohydrates. Hydrogen has been proposed as a future alternative fuel to reduce our demand on traditional fossil fuel-based energy consumption due to its clean energy property, enhanced energy conversion efficiency, and high specific energy density (J/kg) [7, 9, 54]. The combustion or electrochemical conversion of hydrogen produces only water as a final by-product. Vehicles equipped with hydrogen fuel cells are far more energy efficient than traditional internal combustion engine-based vehicles [53]. For example, Toyota will start producing a large number of affordable hydrogen fuel cell vehicles at selling prices of $~50,000 in 2016. Hydrogen also has a higher specific mass energy density than any other fuel sources including gasoline and diesel. However, the low volumetric energy density of hydrogen is currently the biggest issue for its practical use. Thus, the development of new types of green hydrogen production and storage technology remains as challenges.

3.1.2 Hydrogen Production Approaches

Currently most hydrogen is produced from natural gas or coals through a reforming process or gasification followed by water shifting, respectively. These processes are not environment-friendly, releasing CO_2, and the resources are not sustainable. Therefore, the use of hydrogen produced in such ways would have little impact on reducing our demand on traditional fossil fuel-based energy sources [43]. Alternatively, hydrogen can be generated by splitting water molecules with high-temperature thermal energy sources or electricity. Direct splitting of water molecules requires high-temperature thermal energy over 2,000 °C [43]. Sulfur-based thermochemical decomposition of water with heterogeneous metal oxide catalysts can take place at much lower temperatures than direct thermal decomposition, but it still requires at least 750 °C or higher [14]. The thermal energy required for water splitting can be derived from solar energy. Water splitting by solar energy to generate hydrogen, however, is a very slow and inefficient process due to photocatalysts using limited range of visible light and low insolation flux (e.g., ~200 W/m^2) [30, 58]. Most photocatalytic water-splitting processes have shown their production rate less than 1 mmol H_2/L/h [23]. Much higher rates of hydrogen production can be achieved with electricity. A stationary electrolyzer with current technology has been reported to be able to generate hydrogen from water at the rate of about 40 mol/L/h [22]. In spite of such high hydrogen production rates and purity

of hydrogen generation by water electrolysis, its practical applicability is limited due to its high production cost (e.g., ~0.05 US dollars per kWh of electricity). Water electrolysis cannot be free from environmental issues as long as the electricity for electrolyzers is generated from coal or natural gas-powered electric generators [17].

Solar energy and electricity can also be applied to microorganisms to produce hydrogen through photo-fermentation and microbial electrolysis cells (MECs), respectively. Dark fermentation has a theoretical maximum hydrogen production yield of 4 mol of hydrogen per mole of glucose (~33 % efficiency) because of the Thauer limit [42]. Practical efficiency of dark fermentation would be lower than the Thauer limit [16, 21]. Microbial electrolysis cells can achieve much higher efficiencies of about 80 %, but high costs of apparatus and slow production rates (~5.4 mmol H_2/L/h) are the biggest challenges for large-scale hydrogen production from MECs [17, 28]. Overall, microbial fermentation is not an efficient way to produce hydrogen because of the microbial basal metabolism that competes with hydrogen production and eventually reducing the overall product yield.

3.1.3 In Vitro (Cell-Free) Enzymatic Pathways for Water Splitting

In vitro synthetic biosystems for water splitting can produce high-purity (i.e., zero CO production) hydrogen with high yields and rates [59]. In vitro synthetic biosystems are a new cell-free platform that assembles a number of (purified) enzymes and cofactors into different in vitro synthetic (enzymatic) pathways for implantation of various desired biochemical reactions [2, 10, 59, 63]. The optimal reaction condition for numerous enzymes can be found by examining different buffers, such as HEPES, Tris and PBS, a broad range of pH, and cofactor concentrations, and also multi-metal ions, such as Mg^{2+} and Mn^{2+}, to meet trade-off needs of different enzymes. If all enzymes are thermostable, high reaction temperature could be chosen as an optimal condition. These synthetic pathways can utilize different carbohydrates and ambient thermal energy as energy inputs to overcome thermodynamically unfavorable water-splitting reactions to produce hydrogen at mild conditions (below 100 °C). Near theoretical yields of 12 mol of hydrogen per mole of glucose unit consumed have been achieved [50, 60]. More important feature of these synthetic pathways is that they are able to produce hydrogen by absorbing low-temperature waste heat [63]. As the result, endothermic water-splitting reactions can generate more output of chemical energy in the form of hydrogen than input of chemical energy from carbohydrates and water (i.e., 122 % energy efficiency in terms of higher heating values). It is possible to achieve the energy efficiency over 100 % because the water-splitting reaction is a very unique entropy-driven chemical reaction [60]. The in vitro synthetic biosystem for water splitting is more advantageous for hydrogen production than microbial fermentation and photocatalytic water-splitting systems, because of the absence of cellular

membranes and microbial complexity that lower mass transfer and increase biocatalyst density [37]. The highest hydrogen production rate achieved by an in vitro synthetic pathway is about 150 mmol/L/h [38], while microbial fermentation has been reported to produce hydrogen with the rate of 1.96 mol/L/h [52].

3.2 Design of In Vitro Synthetic Enzymatic Pathways

In vitro synthetic pathways for water splitting powered by carbohydrates are reconstituted nonnatural catabolic pathways consisting of more than ten enzymes in four modules: (1) generation of phosphorylated sugars from poly- or monosaccharides without ATP, (2) NADPH generation via the oxidative pentose phosphate pathway (PPP), (3) hydrogen generation, and (4) G6P regeneration via the non-oxidative PPP and gluconeogenesis (Fig. 3.1 and Table 3.1). From the first module, poly- or monosaccharides are converted to their phosphorylated sugar units by phosphorylases or kinases without the use of ATP. For example, the phosphorylation of starch to glucose 1-phosphate units is catalyzed by starch phosphorylase, and different sugar substrates are catalyzed by different enzymes, cellodextrins and cellobiose by cellodextrin and cellobiose phosphorylases, sucrose by sucrose phosphorylase, glucose by polyphosphate glucokinase, and xylulose by polyphosphate xylulokinase. All poly- and oligosaccharides are phosphorylated to be glucose 1-phosphate, which is converted to glucose 6-phosphate by phosphoglucomutase, and then enter the second module reactions to generate NADPH. Xylose takes a different pathway to that catalyzed by xylose isomerase and polyphosphate xylulokinase to produce xylulose 5-phosphate. During the NADPH generation, glucose 6-phosphate (six-carbon sugar) enters the oxidative pentose phosphate pathway and becomes ribulose 5-phosphate (five-carbon sugar) by generating 2 mol of NADPH, releasing 1 mol of CO_2, and absorbing 1 mol of water molecule. When 1 mol of six-carbon sugar is completely consumed, 6 mol of CO_2 are released with the generation of 12 mol of NADPH. Each mole of NADPH is equivalent to 1 mol of hydrogen production catalyzed by NADPH-dependent hydrogenase. Thus, in the hydrogen generation module, 12 mol of hydrogen can be produced when 1 mol of six-carbon sugar is consumed for water splitting. In the non-oxidative pentose phosphate pathway, ribulose 5-phosphates are converted to fructose 6-phosphates and glyceraldehyde 3-phosphates via a carbon skeleton rearrangement. During the G6P regeneration module, a pair of glyceraldehyde 3-phosphates and dihydroxyacetone phosphate is combined to form a fructose 6-phosphate by multiple enzymes via the gluconeogenesis pathway. At this step, one extra mole of water molecule is absorbed by fructose 1,6-bisphosphatase. Finally, all fructose 6-phosphates are catalyzed to regenerate equal moles of glucose 6-phosphate which enter back to the NADPH generation module by completing the cofactor-balanced synthetic pathway for water splitting. The overall

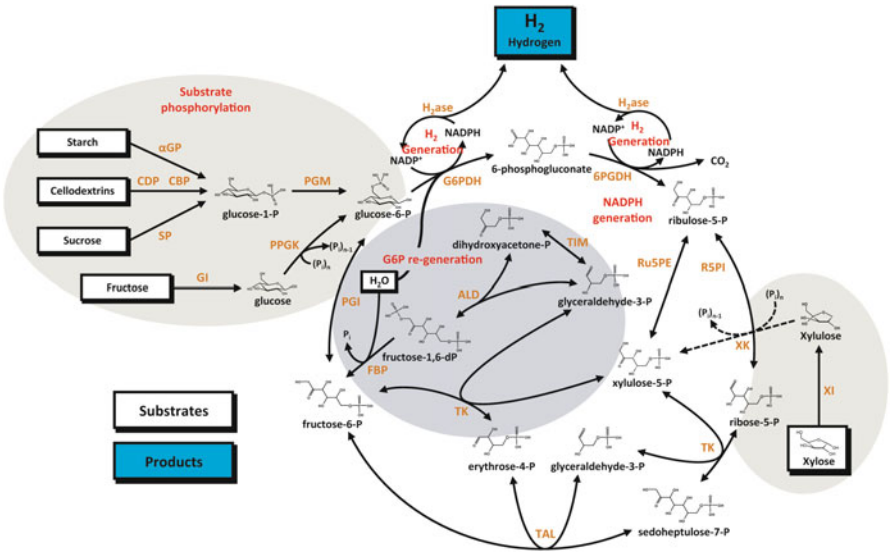

Fig. 3.1 ATP-free cofactor-balanced pathway for hydrogen generation from various types of carbohydrates with water

Table 3.1 Net reactions of each module and the overall reaction of in vitro synthetic biosystem for water splitting

Modules	Net reactions
Module #1	$C_6H_{10}O_5 + H_3PO_4\ (= P_i) \rightarrow$ glucose-6-P
Module #2 (oxidative phase) $\times$ 6	6 $\times$ (glucose-6-P + 2 $NADP^+$ + $\mathbf{H_2O} \rightarrow$ ribulose-5-P + 2 NADPH + 2 H^+ + CO_2)
Module #2 (non-oxidative phase) $\times$ 2	2 $\times$ (3 ribulose-5-P $\rightarrow$ 2 fructose-6-P + glyceraldehyde-3-P)
Module #3 $\times$ 12	12 $\times$ (NADPH + $H^+ \rightarrow NADP^+ + H_2$)
Module #4	2 glyceraldehyde-3-P + $\mathbf{H_2O} \rightarrow$ fructose-6-P + P_i
Module #4 $\times$ 5	5 $\times$ (fructose-6-P $\rightarrow$ glucose-6-P)
Final reaction equation:	$C_6H_{10}O_5 + \mathbf{7H_2O} \rightarrow 12H_2 + 6CO_2$

reaction of this in vitro synthetic pathway for water splitting can be summarized by the equation (Eq. 3.1) as shown in Table 3.1. All of the enzymes used for this hydrogen-producing pathway are listed in Table 3.2.

$$CH_2O(aq) + H_2O(l) \rightarrow 2H_2(g) + CO_2(g) \tag{3.1}$$

Table 3.2 Enzymes used for in vitro synthetic biosystem for water splitting

Modules	E.C. #	Enzyme names (abbreviations)	Reactions
Substrate phosphorylation	2.4.1.1	Glycogen phosphorylase (αGP)	$Glycogen_{(n)} + P_i \rightarrow glucose\text{-}1\text{-}P + glycogen_{(n-1)}$
	2.4.1.49	Cellodextrin phosphorylase (CDP)	$Cellodextrin_{(n)} + P_i \rightarrow glucose\text{-}1\text{-}P + cellodextrin_{(n-1)}$
	2.4.1.20	Cellobiose phosphorylase (CBP)	$Cellobiose + P_i \rightarrow glucose\text{-}1\text{-}P + glucose$
	2.4.1.7	Sucrose phosphorylase (SP)	$Sucrose + P_i \rightarrow glucose\text{-}1\text{-}P + fructose$
	5.3.1.5	Glucose isomerase (GI)	$Fructose \rightarrow glucose$
	2.7.1.63	Polyphosphate glucokinase (PPGK)	$Glucose + (P_i)_n \rightarrow glucose\text{-}1\text{-}P + (P_i)_{n-1}$
	5.4.2.2	Phosphoglucomutase (PGM)	$Glucose\text{-}1\text{-}P \rightarrow glucose\text{-}6\text{-}P$
	5.3.1.5	Xylose isomerase (XI)	$Xylose \rightarrow xylulose$
	2.7.1.17	Polyphosphate xylulokinase (PPXK)	$Xylulose + (P_i)_n \rightarrow xylulose\text{-}5\text{-}P + (P_i)_{n-1}$
NADPH generation	1.1.1.49	Glucose-6-phosphate dehydrogenase (G6PDH)	$Glucose\text{-}6\text{-}P + NADP^+ + H_2O \rightarrow 6\text{-}phosphogluconate + NAHPH + H^+$
	1.1.1.44	6-phosphogluconic dehydrogenase (6PGDH)	$6\text{-}phosphogluconate + NADP^+ \rightarrow ribulose\text{-}5\text{-}P + NADPH + CO_2$
	5.3.1.6	Ribose 5-phosphate isomerase (RPI)	$Ribulose\text{-}5\text{-}P \rightarrow ribose\text{-}5\text{-}P$
	5.1.3.1	Ribulose-5-phosphate 3-epimerase (RPE)	$Ribulose\text{-}5\text{-}P \rightarrow xylulose\text{-}5\text{-}P$
	2.2.1.1	Transketolase (TK)	$Xylulose\text{-}5\text{-}P + ribose\text{-}5\text{-}P \rightarrow sedoheptulose\text{-}7\text{-}P + glyceraldehyde\text{-}3\text{-}P$
	2.2.1.1	Transketolase (TK)	$Xylulose\text{-}5\text{-}P + erythrose\text{-}4\text{-}P \rightarrow fructose\text{-}6\text{-}P + glyceraldehyde\text{-}3\text{-}P$
	2.2.1.2	Transaldolase (TAL)	$Sedoheptulose\text{-}7\text{-}P + glyceraldehyde\text{-}3\text{-}P \rightarrow fructose\text{-}6\text{-}P + erythrose\text{-}4\text{-}P$
Hydrogenation	1.12.1.3	Hydrogenase (H_2ase)	$NADPH + H^+ \rightarrow NADP^+ + H_2$
G6P regeneration	5.3.1.1	Triose-phosphate isomerase (TIM)	$Glyceraldehyde\text{-}3\text{-}P \rightarrow dihydroxyacetone\ phosphate$
	4.1.2.13	Aldolase (ALD)	$Glyceraldehyde\text{-}3\text{-}P + dihydroxyacetone\ phosphate \rightarrow fructose\text{-}1,6\text{-}bisphosphate$
	3.1.3.11	Fructose-1,6-bisphosphatase (FBP)	$Fructose\text{-}1,6\text{-}bisphosphate + H_2O \rightarrow fructose\text{-}6\text{-}P + P_i$
	5.3.1.9	Phosphoglucose isomerase	$Fructose\text{-}6\text{-}P \rightarrow glucose\text{-}6\text{-}P$

3.3 Examples of Hydrogen Production from Carbohydrates

3.3.1 Hydrogen Production from Starch and Cellodextrins

In vitro enzymatic pathways to produce hydrogen from carbohydrates were first demonstrated by Woodward and colleagues using glucose, xylose, or sucrose with only a couple of enzymes [47, 48]. These simple enzymatic pathways could achieve less than 10 % of the theoretical yields of sugars, due to only one NADPH generation per hexose or pentose. To complete the oxidation, the oxidative pentose phosphate cycle was coupled with hydrogenase to produce hydrogen from glucose 6-phosphate, resulting in about 96 % of the theoretical yield [49]. High cost of glucose 6-phosphate prevents its practical application. These in vitro synthetic pathways were further improved by Zhang and collaborators to demonstrate hydrogen production from different types of carbohydrates: starch [60], cellulosic materials [50], xylose [31], and sucrose [32]. The first hydrogen production from starch by an in vitro synthetic enzymatic pathway proved its feasibility to produce low-cost hydrogen from inexpensive starch without ATP by achieving a high production yield surpassing the theoretical production yield of dark fermentation [60]. Most enzymes used for this proof-of-principle experiment producing hydrogen from starch were off-the-shelf enzymes, and the enzymatic pathway exhibited a production rate of 0.4 mmol/L/h with 43 % yield (5.2 mol H_2/mol glucose consumed) (Table 3.3). A couple of years later, the hydrogen production rates and yields were enhanced to 0.5 mmol/L/h with 93 % yield and 3.9 mmol/L/h with 68 % yield when cellobiose or cellodextrins were used as a substrate, respectively, and through minor optimizations including increased substrate concentration, reaction temperature, and more rate-limiting enzyme loadings [50] (Table 3.3). The hydrogen production rate using cellodextrins was greatly improved by increasing hydrogenase loading and substrate concentration to 8 mM. However, low-yield hydrogen production was observed due to the incomplete reaction. Cellobiose and cellodextrins were prepared as hydrolytic products of cellulose through incomplete enzymatic or mixed acid hydrolysis, respectively.

Table 3.3 Comparison of hydrogen production rates and yields from different carbohydrates

Substrate	Conc. (mM)	Reaction temperature (°C)	H_2 production rates (mmol H_2/L/h)	Yield (%)	Ref.
Starch	1	30	0.4	43	[60]
Cellobiose	2	32	0.5	93	[50]
Cellodextrins (cellopentaose)	8	32	3.9	68	[50]
Xylose	2	50	2.2	96	[31]
Sucrose	2	37	3.0	97	[32]
G6P	100	60	150	–	[38]
Biomass sugar	3.19	40	2.3	100	[36]
G6P	100	50	54	–	[36]

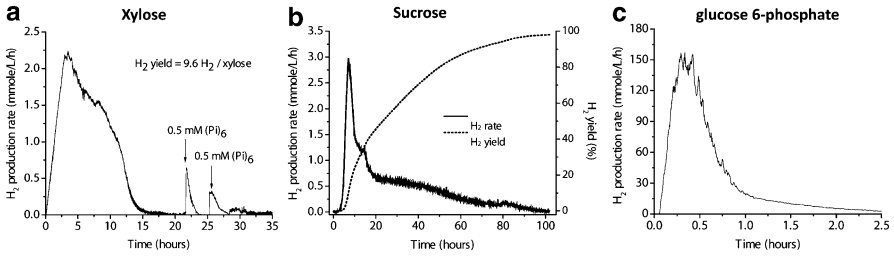

Fig. 3.2 Hydrogen production profiles from xylose (**a**), sucrose (**b**), and glucose 6-phosphate (**c**). (Reproduced with permission from [31, 32, 36])

3.3.2 Hydrogen Production from Xylose

Hemicellulose is another major component of lignocellulosic biomass besides cellulose. Xylose is the most abundant pentose and the major component of hemicellulose. Xylose composes about 20–30 % of lignocellulosic biomass by weight [31]. Thus, it is essential to use not only cellulosic materials but also the major portion of hemicellulosic materials to produce low-cost hydrogen from renewable carbohydrates. One mole of xylose can theoretically produce 10 mol of hydrogen by splitting 5 mol of water molecule when completely oxidized to carbon dioxide (Eq. 3.2). Since xylose is a five-carbon sugar, the in vitro synthetic pathway was modified by replacing hexose-phosphorylating enzymes with xylose isomerase and polyphosphate xylulokinase (Fig. 3.1). The hydrogen production from xylose exhibited the production rate of 2.2 mmol/L/h with 96 % yield (Table 3.3). Such high-yield hydrogen production was achieved with the addition of extra polyphosphate driving the reaction to completion (Fig. 3.2a), and the relatively high production rate was achieved with increased reaction temperature to 50 °C:

$$C_5H_{10}O_5(aq) + 5H_2O(l) \rightarrow 10H_2(g) + 5CO_2(g) \tag{3.2}$$

3.3.3 Hydrogen Production from Sucrose

Sucrose, also known as table sugar, is one of the cheapest carbohydrates because of its simple production process and abundance of cultivated sugar crops: sugarcane and sugar beets. Sucrose is a disaccharide composed of glucose and fructose. Therefore, two additional enzymes are required to convert fructose to equal moles of glucose 6-phosphate. An in vitro synthetic enzymatic pathway consisting of total 15 enzymes was designed to catalyze water-splitting process producing hydrogen powered by sucrose [32]. The hydrogen production from sucrose achieved the production rate of 3.0 mmol/L/h with 97 % yield (Fig. 3.2b). When

increased sucrose concentration, the maximum hydrogen production rate was as high as 9.7 mmol/L/h [32].

3.3.4 Hydrogen Production from Biomass Sugars

The complete conversion of glucose and xylose from plant biomass to hydrogen and carbon dioxide has been achieved via an in vitro synthetic enzymatic pathway. Pretreated biomass was hydrolyzed to glucose and xylose by using a commercial cellulase. Glucose and xylose were simultaneously converted to hydrogen with the theoretical yield of 2 mol of hydrogen per each carbon molecule [36]. A genetic algorithm was used to find the best fitting parameters of a nonlinear kinetic model with experimental data. Global sensitivity analysis was used to identify the key enzymes that have the greatest impact on reaction rate and yield. After optimization of enzyme loadings using computational modeling and data analysis methods, the hydrogen production rate could be increased to 32 mmol/L/h. The production rate was further enhanced to 54 mmol/L/h by increasing reaction temperature, substrate, and enzyme concentrations. The production of hydrogen from locally produced biomass is a promising means to achieve global green hydrogen production.

3.3.5 High-Rate Hydrogen Production from Glucose 6-Phosphate

High-yield hydrogen production from these different carbohydrates has opened up a new way to produce low-cost hydrogen from renewable biomass. In the past years, the hydrogen production rate has been increased to 150 mmol/L/h by 750-fold when glucose 6-phosaphte is used as a substrate by an in vitro synthetic enzymatic pathway (Fig. 3.2c) [38, 49]. All enzymes used in the pathway have been replaced with recombinant thermostable enzymes produced in *E. coli*. Some of these enzymes are immobilized to enhance their stability. Hydrogen production rates and yields from different carbohydrates are summarized in Table 3.3.

3.4 Technical Obstacles to Low-Cost H_2 Production

To achieve low-cost hydrogen production from carbohydrates, a few obstacles have to be overcome. The ultimate hydrogen production costs are strongly related to the following factors: costs of substrate, enzyme cost, cofactor cost, and product-related downstream processing cost especially for product separation and purification [61]. It has been shown that the in vitro synthetic biosystem can produce

hydrogen from various types of carbohydrates in near theoretical yields. Therefore, the use of inexpensive, abundant carbohydrate sources, such as cellulosic and hemicellulosic materials, to produce hydrogen can solve one of the major obstacles. Gaseous products can be easily separated and purified from the aqueous phase enzymatic reaction, leaving the other two obstacles unsolved. Hydrogen production rate is also an important factor, because it determines potential implementation of low-cost hydrogen production techniques from carbohydrates mainly related to capital investment.

3.4.1 Enzyme Cost and Stability

Enzyme costs are highly related to their production costs, turnover number (TTN), and stability. The current production costs for enzymes produced as recombinant proteins in *E. coli* at lab scales are high, for example, $~1,000,000 per kg of dry protein, but it is expected that their production costs will decrease to the level of industrial bulk enzyme production costs of around $5–100 per kg of dry protein [57]. Increasing TTNs can decrease enzyme costs exponentially through increasing enzyme efficiency [55]. It is estimated that TTNs over 10^8–10^9 are required to reduce enzyme costs low enough for industrial-scale use [11, 12, 51]. High TTNs can be achieved by using thermostable enzymes from thermophilic microbes, or through enzyme immobilization [33, 44, 54].

The use of thermostable enzymes can decrease their production costs by decreasing enzyme purification costs. Thermostable enzymes cloned from thermophilic microbes, such as *Thermotoga maritima*, can be stable at 60–70 °C, which is higher than the temperature where most of the other enzymes cloned from mesophilic hosts are stable. Therefore, purification procedures retaining only soluble thermostable enzymes after treating at high temperature for relatively short times (10–30 min) make it a simple and cost-effective way to purify target enzymes from cell lysates. Simple enzyme purification will eventually lower the overall production costs.

Enzyme immobilization is a technique in a relatively mature stage, and various immobilization techniques have been introduced and used to improve TTNs, enzyme stability, and catalytic efficiency [24, 33, 41]. Among many techniques, the cellulose-binding module (CBM) tagged protein immobilization combines enzyme purification and immobilization into one step [18]. This simple one-step enzyme purification and immobilization technique showed about an 80-fold enhanced half-life time of phosphoglucose isomerase (PGI) when the enzyme with CBM tag was immobilized on regenerated amorphous cellulose (RAC) (Fig. 3.3) [33]. Enzyme immobilization also enables enzymes to be recyclable. Enzymes immobilized on magnetic nanoparticles (MNPs) have shown to improve enzymatic reactions rates and also be able to recycle simply by using a magnetic force [34]. Green fluorescent protein (CBM-TGC) was used to demonstrate simply selective recycling of CBM-tagged GFP immobilized on Avicel-containing MNPs

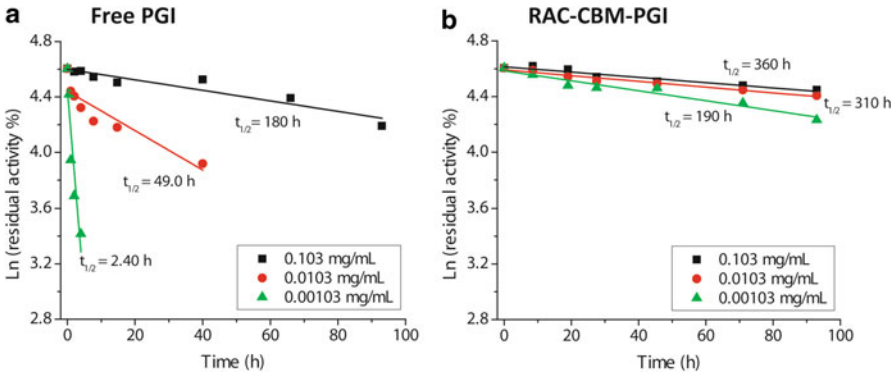

Fig. 3.3 Comparison of thermal stability between free PGI (**a**) and immobilized PGI (**b**) in different concentrations (Reproduced with permission from Myung et al. [33])

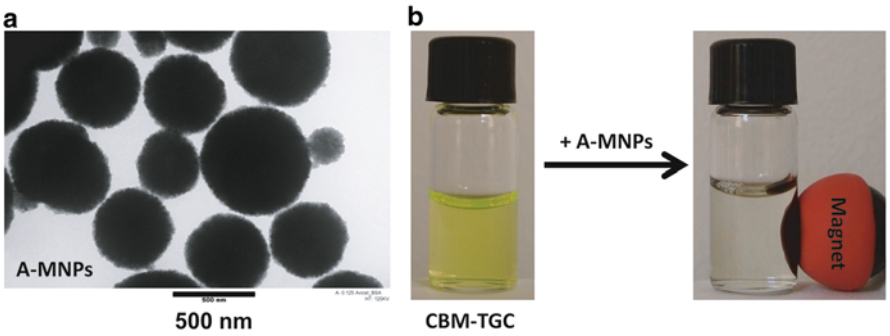

Fig. 3.4 Selective recycling of enzymes immobilized on Avicel-containing magnetic nanoparticles (A-MNPs) (**a**). The simple process of collecting CBM-tagged green fluorescent proteins (CBM-TGC) adsorbed on A-MNPs by a magnetic force (**b**) (Reproduced with permission from Myung et al. [34])

(A-MNPs) (Fig. 3.4). As the result of selective recycling, enzyme-related costs can be greatly reduced.

Protein protection additives, such as ligands and salts, have been studied as a simple approach to enhance enzyme storage and reaction stability. DMSO and glycerol are popular additives as cryoprotectants, stabilizing proteins during multi-freezing-and-thawing cycles. Various polyethylene glycols have been used as thermoprotectants. Thermostability of trypsin was increased from 49 to 93 °C without deteriorating its catalytic properties in the presence of glycol chitosan [13]. Amines, polyethylene glycol, and glycerol as additives improved catalase storage stability as well as its enzymatic performance in high temperature and alkaline pH [6]. These additives may, however, be potential inhibitors to the reaction system [19]. Therefore, additional dialysis or ultrafiltration may be required before the reactions.

3.4.2 Enzymatic Reaction Rates

Reaction rate is an important factor for determining potential applications of the enzymatic reaction and capital investment. Currently, the fastest hydrogen production rate achieved by an in vitro synthetic enzymatic pathway is about 150 mmol/L/h (=0.3 g H_2/L/h) [38]. This current rate is the fastest enzymatic hydrogen production from sugars and fast enough for distributed hydrogen generators, but it is slower than the need of on-demand hydrogen production for hydrogen fuel cell vehicles. It is anticipated that the production rate can be improved to 20 g H_2/L/h because microbial fermentation, which is usually slower than cell-free enzymatic reactions, has already achieved a production rate of over 20 g H_2/L/h [5, 54, 62]. Approaches made to increase enzymatic hydrogen production rate include elevated reaction temperature, optimized enzyme ratio, high substrate concentration, high enzyme loading, and substrate channeling among cascade enzymes. The substrate channeling was done by co-immobilizing multiple cascade enzymes, so that the local enzyme concentration is increased and susceptible intermediates have less time exposed to reaction solutions [20]. The cluster of multi-enzymes held by co-immobilizing scaffoldins results in increased reaction rates. Three important cascade enzymes, TIM, ALD, and FBP, in the enzymatic hydrogen production were expressed with dockerin domain which can self-assemble with mini-scaffoldin for co-immobilization on A-MNPs (Fig. 3.5a) [34]. In comparison with non-immobilized enzymes, the co-immobilized enzyme complex results in about 4.6 times increased reaction rate (Fig. 3.5b).

3.4.3 Cofactor Cost and Stability

Cofactors are chemical compounds required for enzymatic reactions, such as ATP and NAD(P)H. The issues from cofactor costs and stability have been addressed by various approaches including cofactor recycling systems [25, 46] and use of low-cost stable biomimetic cofactors [1, 29]. Although the regeneration of natural cofactors through recycling systems is economically beneficial to most current enzymatic reactions, the most farsighted solution would be to replace native cofactors with low-cost stable biomimetic cofactors. The structures of natural cofactors and biomimetic cofactors are shown in Fig. 3.6 with their estimated prices (USD, 2015). By sharing the nicotinamide moiety as a universal binding site for electron carriers, these natural and biomimetic cofactors vary in the chemical structures bound to the nitrogen atom in pyridine (Fig. 3.6). The alternative natural cofactors and biomimetic cofactors with simpler structures than NADP or NAD are estimated to have lower costs. However, most wild-type redox enzymes have no activities with such biomimetic cofactors. A number of studies have been done to change the cofactor specificity or preference through cofactor engineering. Cofactor engineering has three major types of approaches: rational design, directed

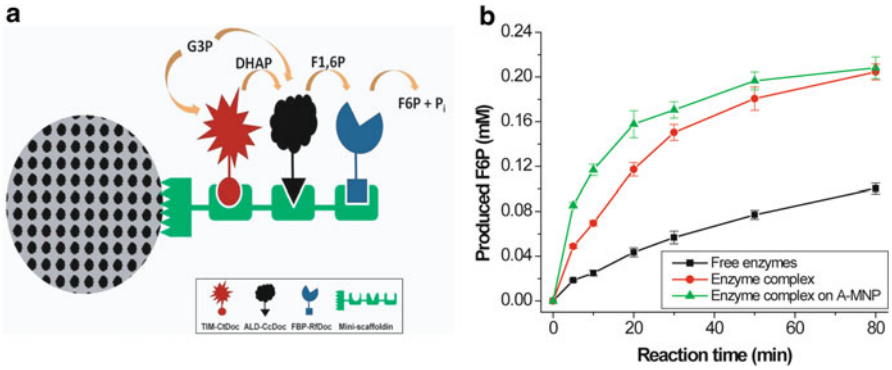

Fig. 3.5 The cascade enzymatic reaction among co-immobilized enzymes on MNPs for substrate channeling (**a**). Comparison of the reaction rates among free enzymes, non-immobilized enzyme complex, and the enzyme complex immobilized on A-MNP (**b**) (Reproduced with permission from Myung et al. [34])

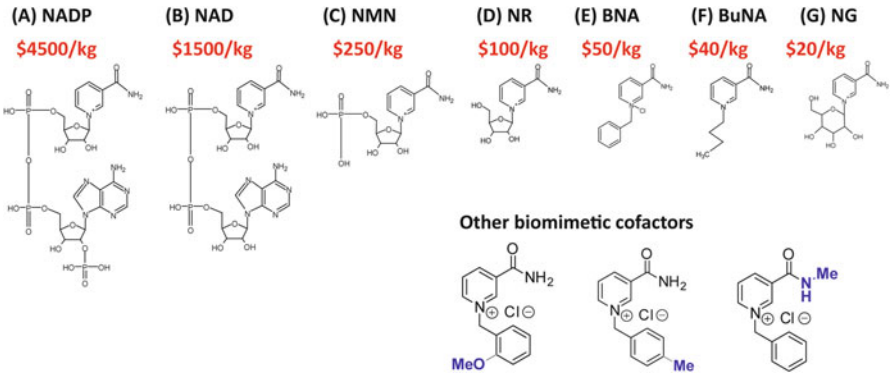

Fig. 3.6 Structures and estimated prices of natural cofactors (**A**, **B**), alternative natural cofactors (**C**, **D**), and biomimetic cofactors (**E–G**)

evolution, or swapping modules. Using these approaches, a number of studies have shown their redox enzymes with changed cofactor preferences from NADP to NAD [3, 8, 39], or from NAD to NADP [15, 45]. In 2012, Scott et al. discovered that their engineered alcohol dehydrogenase for broadened cofactor specificity and improved activity with NAD can utilize a minimal natural cofactor, NMN [4]. Fish et al. proposed the use of 1-benzyl-3-carbamoyl-pyridinium (BNA) as a biomimetic cofactor to replace NAD(P)H and discovered that two wild-type enzymes, horse liver alcohol dehydrogenase and monooxygenase, can actually utilize this cofactor [26, 27]. Clark collaborated with Fish and demonstrated that engineered P450 with two amino acid mutations can utilize BNA as a cofactor [40]. Recently, a large international collaborative group synthetized another biomimetic cofactor, 1-butyl-3-carbamoyl-pyridinium (BuNA), and demonstrated that wild-type enoate reductase can utilize it as a cofactor [35]. In most cases, such changes will decrease

apparent activities greatly. The best example may be the engineered P450. The mutant P450 (W1064S/R966D) exhibited its activity on biomimetics up to seven times of that of wild-type enzyme on NADH [40]. Cofactor engineering is in its early stage, but its success will greatly influence enzymatic synthesis of organic chemicals and in vitro synthetic biosystems for biocommodity production.

3.5 Conceptual Obstacles to Enzymatic H_2 Production

In vitro synthetic biosystems consist of numerous enzymes as building bricks, enzyme complexes as building modules, and/or (biomimetic) coenzymes. These many components are assembled into in vitro synthetic pathways for implementing complicated bioreactions. They emerge as an alternative solution for accomplishing desired biotransformation without concerns of cell proliferation, complicated cellular regulation, and side-product formation. In addition to the capability of achieving high product yields as the most important advantage, in vitro synthetic biosystems feature several other biomanufacturing advantages, such as fast reaction rates, easy product separation, open process control, broad reaction conditions, and tolerance to toxic substrates or products.

The largest obstacle to the enzymatic hydrogen production is conceptual change. Microbe-based fermentation has been used by human beings for more than 10,000 years. As a result, most biotechnologists believe that living microbes are the best biocatalysts because they can duplicate themselves. Indeed, the primary goal of living microbes is their proliferation while bioconversions are side effects. Consequently, the success of several examples is needed to convince biotechnologists of accepting a new paradigm of in vitro synthetic biosystems.

3.6 Conclusions and Future Outlook

Hydrogen is an important commodity chemical with a global market size of approximately 100 billion US dollars. High-yield hydrogen has been produced from carbohydrates and water catalyzed by in vitro synthetic enzymatic pathways. The overall hydrogen production costs are mainly proportional to carbohydrate cost [59]. Assuming a substrate cost contributes a major portion of hydrogen product cost, it is expected that hydrogen can be produced at costs less than $2.00/kg hydrogen when the technology is well developed [59] (Fig. 3.7). Carbohydrates are the most abundant renewable natural resources on earth, and they are estimated to cost about $1.50/kg accounting for 75 % of the prospective hydrogen production cost [59] (Fig. 3.7). The next biggest portion of the hydrogen production cost is from the initial capital investment and operating expense accounting for about 17 %. The sum of capital and operating expenses would be about $0.35/kg hydrogen according to the similar expenses based on anaerobic digestion. Enzyme and

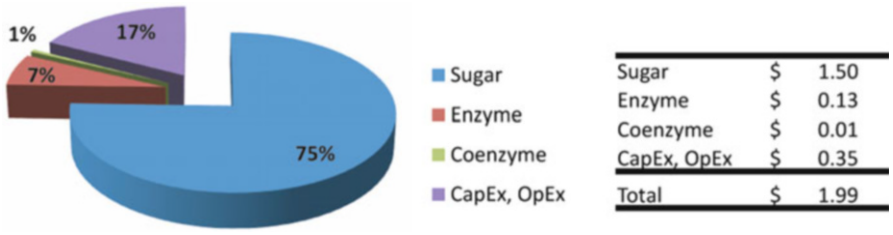

Fig. 3.7 Prospective H_2 production cost and its contribution factors: substrate costs, enzyme costs, cofactor costs, and initial capital expense with operating expense

cofactor costs account for 7 % and 1 % of the enzymatic hydrogen production cost, respectively, when their TTN values are more than 10^9 and 10^7, respectively.

High-yield hydrogen production from carbohydrates will open up several potential applications for the hydrogen economy, where hydrogen is an alternative transportation fuel or a short-term electricity storage carrier. The potential applications include from small-sized distributed hydrogen refueling stations (e.g., 1–2 kg per day for a single house, or 50–200 kg for village) to the most ambitious application, such as a sugar fuel cell vehicles [53, 56].

The potential global market size for hydrogen as a future energy carrier replacing gasoline and diesel could be trillions of dollars. Such great potential, along with a bright future featuring enhanced energy conversion efficiency, nearly zero pollutants, and zero greenhouse gas emissions, will motivate the world to solve the remaining obstacles within next decades.

Acknowledgments YPZ was supported by the Virginia Tech Biological Systems Engineering Department, and subcontracts from NSF STTR I (IIP-1321528) and SBIR II (IIP-1353266) awards and Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences. In addition, funding for this work was provided in part by the Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, US Department of Agriculture.

References

1. Ansell RJ, Lowe CR. Artificial redox coenzymes: biomimetic analogues of NAD(+). Appl Microbiol Biotechnol. 1999;51(6):703–10.
2. Ardao I, Hwang ET, Zeng AP. In vitro multienzymatic reaction systems for biosynthesis. Fundam Appl New Bioprod Syst. 2013;137:153–84.
3. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S. Optimizing an artificial metabolic pathway: engineering the cofactor specificity of Corynebacterium 2,5-diketo-D-gluconic acid reductase for use in vitamin C biosynthesis. Biochem-Us. 2002;41(20):6226–36.
4. Campbell E, Meredith M, Minteer SD, Banta S. Enzymatic biofuel cells utilizing a biomimetic cofactor. Chem Commun. 2012;48(13):1898–900.
5. Cooney MJ, Svoboda V, Lau C, Martin G, Minteer SD. Enzyme catalysed biofuel cells. Energy Environ Sci. 2008;1(3):320–37.
6. Costa SA, Tzanov T, Carneiro AF, Paar A, Gübitz GM, Cavaco-Paulo A. Studies of stabilization of native catalase using additives. Enzyme Microb Technol. 2002;30(3):387–91.

7. Crabtree GW, Dresselhaus MS, Buchanan MV. The hydrogen economy. Phys Today. 2004;57 (12):39–44.
8. Dohr O, Paine MJI, Friedberg T, Roberts GCK, Wolf CR. Engineering of a functional human NADH-dependent cytochrome P450 system. Proc Natl Acad Sci U S A. 2001;98(1):81–6.
9. Dresselhaus MS. Basic research needs for the hydrogen economy. Abstr Pap Am Chem S. 2004;227:U1084–5.
10. Dudley QM, Karim AS, Jewett MC. Cell-free metabolic engineering: biomanufacturing beyond the cell. Biotechnol J. 2015;10(1):69–82.
11. Eijsink VGH, Bjork A, Gaseidnes S, Sirevag R, Synstad B, van den Burg B, Vriend G. Rational engineering of enzyme stability. J Biotechnol. 2004;113(1–3):105–20.
12. Eijsink VGH, Gaseidnes S, Borchert TV, van den Burg B. Directed evolution of enzyme stability. Biomol Eng. 2005;22(1–3):21–30.
13. Fernández L, Gómez L, Ramírez HL, Villalonga ML, Villalonga R. Thermal stabilization of trypsin with glycol chitosan. J Mol Catal B: Enzym. 2005;34(1):14–7.
14. Ginosar DM, Petkovic LM, Glenn AW, Burch KC. Stability of supported platinum sulfuric acid decomposition catalysts for use in thermochemical water splitting cycles. Int J Hydrogen Energy. 2007;32(4):482–8.
15. Glykys DJ, Banta S. Metabolic control analysis of an enzymatic biofuel cell. Biotechnol Bioeng. 2009;102(6):1624–35.
16. Hallenbeck PC, Benemann JR. Biological hydrogen production; fundamentals and limiting processes. Int J Hydrogen Energy. 2002;27(11):1185–93.
17. Holladay JD, Hu J, King DL, Wang Y. An overview of hydrogen production technologies. Catal Today. 2009;139(4):244–60.
18. Hong J, Wang YR, Ye XH, Zhang Y-HP. Simple protein purification through affinity adsorption on regenerated amorphous cellulose followed by intein self-cleavage. J Chromatogr A. 2008;1194(2):150–4.
19. Iyer PV, Ananthanarayan L. Enzyme stability and stabilization—aqueous and non-aqueous environment. Process Biochem. 2008;43(10):1019–32.
20. Jia F, Narasimhan B, Mallapragada S. Materials-based strategies for multi-enzyme immobilization and co-localization: a review. Biotechnol Bioeng. 2014;111(2):209–22.
21. Kleerebezem R, van Loosdrecht M. Mixed culture biotechnology for bioenergy production. Curr Opin Biotechnol. 2007;18(3):207–12.
22. Levin DB, Pitt L, Love M. Biohydrogen production: prospects and limitations to practical application. Int J Hydrogen Energy. 2004;29(2):173–85.
23. Liao C-H, Huang C-W, Wu J. Hydrogen production from semiconductor-based photocatalysis via water splitting. Catalysts. 2012;2(4):490–516.
24. Liese A, Hilterhaus L. Evaluation of immobilized enzymes for industrial applications. Chem Soc Rev. 2013;42(15):6236–49.
25. Liu WF, Wang P. Cofactor regeneration for sustainable enzymatic biosynthesis. Biotechnol Adv. 2007;25(4):369–84.
26. Lo HC, Fish RH. Biomimetic NAD(+) models for tandem cofactor regeneration, horse liver alcohol dehydrogenase recognition of 1,4-NADH derivatives, and chiral synthesis. Angew Chem Int Edit. 2002;41(3):478–81.
27. Lo HC, Leiva C, Buriez O, Kerr JB, Olmstead MM, Fish RH. Bioorganometallic chemistry. 13. Regioselective reduction of NAD(+) models, 1-benzylnicotinamde triflate and beta-nicotinamide ribose-5′-methyl phosphate, with in situ generated [CP*Rh(Bpy)H](+): structure-activity relationships, kinetics, and mechanistic aspects in the formation of the 1,4-NADH derivatives. Inorg Chem. 2001;40(26):6705–16.
28. Logan BE, Call D, Cheng S, Hamelers HVM, Sleutels THJA, Jeremiasse AW, Rozendal RA. Microbial electrolysis cells for high yield hydrogen gas production from organic matter. Environ Sci Technol. 2008;42(23):8630–40.

29. Lutz J, Hollmann F, Ho TV, Schnyder A, Fish RH, Schmid A. Bioorganometallic chemistry: biocatalytic oxidation reactions with biomimetic NAD(+)/NADH co-factors and [Cp*Rh(bpy)H](+) for selective organic synthesis. J Organomet Chem. 2004;689(25):4783–90.
30. Maeda K, Domen K. Photocatalytic water splitting: recent progress and future challenges. J Phys Chem Lett. 2010;1(18):2655–61.
31. Martin del Campo JS, Rollin J, Myung S, Chun Y, Chandrayan S, Patino R, Adams MWW, Zhang Y-HP. High-yield production of dihydrogen from xylose by using a synthetic enzyme cascade in a cell-free system. Angew Chem Int Edit. 2013;52(17):4587–90.
32. Myung S, Rollin J, You C, Sun FF, Chandrayan S, Adams MWW, Zhang Y-HP. In vitro metabolic engineering of hydrogen production at theoretical yield from sucrose. Metab Eng. 2014;24:70–7.
33. Myung S, Zhang XZ, Zhang Y-HP. Ultra-stable phosphoglucose isomerase through immobilization of cellulose-binding module-tagged thermophilic enzyme on low-cost high-capacity cellulosic adsorbent. Biotechnol Prog. 2011;27(4):969–75.
34. Myung SW, You C, Zhang Y-HP. Recyclable cellulose-containing magnetic nanoparticles: immobilization of cellulose-binding module-tagged proteins and a synthetic metabolon featuring substrate channeling. J Mater Chem B. 2013;1(35):4419–27.
35. Paul CE, Gargiulo S, Opperman DJ, Lavandera I, Gotor-Fernández V, Gotor V, Taglieber A, Arends IW, Hollmann F. Mimicking nature: synthetic nicotinamide cofactors for C=C bioreduction using enoate reductases. Org Lett. 2012;15(1):180–3.
36. Rollin JA, Martin del Campo JS, Myung S, Sun FF, You C, Bakovic AE, Castro RL, Chandrayan S, Wu C-H, Adams MWW, Senger R, Zhang Y-HP. High-yield hydrogen production from biomass by in vitro metabolic engineering: mixed sugars coutilization and kinetic modelling. Proc Natl Acad Sci U S A. 2015;112:4964–9.
37. Rollin JA, Tam TK, Zhang Y-HP. New biotechnology paradigm: cell-free biosystems for biomanufacturing. Green Chem. 2013;15(7):1708–19.
38. Rollin JA, Ye X, del Campo JM, Adams MW, Zhang Y-HP. Novel hydrogen bioreactor and detection apparatus. Adv Biochem Eng Biotechnol. 2014;274:1–17.
39. Rosell A, Valencia E, Ochoa WF, Fita I, Pares X, Farres J. Complete reversal of coenzyme specificity by concerted mutation of three consecutive residues in alcohol dehydrogenase. J Biol Chem. 2003;278(42):40573–80.
40. Ryan JD, Fish RH, Clark DS. Engineering cytochrome P450 enzymes for improved activity towards biomimetic 1,4-NADH cofactors. Chembiochem. 2008;9(16):2579–82.
41. Sheldon RA, Sorgedrager M, Janssen MHA. Use of cross-linked enzyme aggregates (CLEAs) for performing biotransformations. Chim Oggi. 2007;25(1):62–7.
42. Thauer RK, Jungermann K, Decker K. Energy conservation in chemotrophic anaerobic bacteria. Bacteriol Rev. 1977;41(1):100.
43. Turner JA. Sustainable hydrogen production. Science. 2004;305(5686):972–4.
44. Wang YR, Zhang Y-HP. A highly active phosphoglucomutase from Clostridium thermocellum: cloning, purification, characterization and enhanced thermostability. J Appl Microbiol. 2010;108(1):39–46.
45. Watanabe S, Kodaki T, Makino K. Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. J Biol Chem. 2005;280(11):10340–9.
46. Weckbecker A, Groger H, Hummel W. Regeneration of nicotinamide coenzymes: principles and applications for the synthesis of chiral compounds. Adv Biochem Eng Biot. 2010;120:195–242.
47. Woodward J, Mattingly SM, Danson M, Hough D, Ward N, Adams M. In vitro hydrogen production by glucose dehydrogenase and hydrogenase. Nat Biotechnol. 1996;14(7):872–4.
48. Woodward J, Orr M. Enzymatic conversion of sucrose to hydrogen. Biotechnol Prog. 1998;14(6):897–902.
49. Woodward J, Orr M, Cordray K, Greenbaum E. Biotechnology: enzymatic production of biohydrogen. Nature. 2000;405(6790):1014–5.

50. Ye XH, Wang YR, Hopkins RC, Adams MWW, Evans BR, Mielenz JR, Zhang Y-HP. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. Chemsuschem. 2009;2(2):149–52.
51. Ye XH, Zhang CM, Zhang Y-HP. Engineering a large protein by combined rational and random approaches: stabilizing the Clostridium thermocellum cellobiose phosphorylase. Mol Biosyst. 2012;8(6):1815–23.
52. Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H. Enhanced hydrogen production from formic acid by formate hydrogen lyase-overexpressing Escherichia coli strains. Appl Environ Microbiol. 2005;71(11):6762–8.
53. Zhang Y-HP. A sweet out-of-the-box solution to the hydrogen economy: is the sugar-powered car science fiction? Energy Environ Sci. 2009;2(3):272–82.
54. Zhang Y-HP. Using extremophile enzymes to generate hydrogen for electricity. Microbe. 2009;4:560–5.
55. Zhang Y-HP. Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. Biotechnol Bioeng. 2010;105(4):663–77.
56. Zhang Y-HP. Renewable carbohydrates are a potential high-density hydrogen carrier. Int J Hydrogen Energy. 2010;35(19):10334–42.
57. Zhang Y-HP. Hydrogen production from carbohydrates: a mini-review. Acs Sym Ser. 2011;1067:203–16.
58. Zhang Y-HP. What is vital (and not vital) to advance economically-competitive biofuels production. Proc Biochem. 2011;46:2091–110.
59. Zhang Y-HP. Production of biofuels and biochemicals by in vitro synthetic biosystems: opportunities and challenges. Biotechnol Adv. 2014. doi:10.1016/j.biotechadv.2014.10.009.
60. Zhang Y-HP, Evans BR, Mielenz JR, Hopkins RC, Adams MWW. High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. PLoS One. 2007;2(5): e456. doi:10.1371/journal.pone.0000456.
61. Zhang Y-HP, Mielenz JR. Renewable hydrogen carrier – carbohydrate: constructing the carbon-neutral carbohydrate economy. Energies. 2011;4(2):254–75.
62. Zhu ZG, Tam TK, Sun FF, You C, Zhang Y-HP. A high-energy-density sugar biobattery based on a synthetic enzymatic pathway. Nat Commun. 2014;5:3026.
63. Zhu ZG, Tam TK, Zhang Y-HP. Cell-free biosystems in the production of electricity and bioenergy. Fundam Appl New Bioprod Syst. 2013;137:125–52.

Chapter 2-2. Use of carbohydrates for hydrogen storage

Kim, J.-E., and Zhang, Y.H.P. (2016). Chapter 9. Use of carbohydrates for hydrogen storage. In Compendium to Hydrogen Energy - Hydrogen storage, transmission, transportation and infrastructure, Volume 2, pp. 219-241. Reproduced by permission of Elsevier.

Use of carbohydrates for hydrogen storage

9

J.-E. Kim[1], *Y.-H. Percival Zhang*[1,2,3]
[1]Virginia Tech, Blacksburg, VA, USA; [2]Cell-Free Bioinnovations Inc., Blacksburg, VA, USA; [3]Institute for Critical Technology and Applied Science (ICTAS), Virginia Tech, Blacksburg, VA, USA

Abbreviations

ATP	adenosine triphosphate
MFC	microbial fuel cell
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
P_i	phosphate
SFCV	sugar fuel cell vehicle
SyPaB	synthetic pathway biotransformation

Enzymes of cell-free synthetic pathway biotransformation

6PGDH	6-phosphogluconate dehydrogenase
ALD	aldolase
FBP	fructose-1,6-bisphosphatase
G6PDH	glucose 6-phosphate dehydrogenase
GNP	glucan phosphorylase
H_2ase	hydrogenase
PGI	phosphoglucose isomerase
PGM	phosphoglucomutase
PPGK	polyphosphate glucokinase
R5PI	phosphoribose isomerase
Ru5PE	ribulose 5-phosphate epimerase
TAL	transaldolase
TIM	triosephosphate isomerase
TK	transketolase
XI	xylulose isomerase
XK	xylulokinase

Compendium of Hydrogen Energy. http://dx.doi.org/10.1016/B978-1-78242-362-1.00009-2

9.1 Introduction

9.1.1 Carbohydrates

Carbohydrates technically refer to any hydrates of carbons with the empirical formula of $C_m(H_2O)_n$, where m and n may be different. Although some chemical compounds conform to this formula, such as formaldehyde and glycolaldehyde, they are not truly considered as carbohydrates. The most accurate definition of carbohydrates is polyhydroxy aldehydes or polyhydroxy ketones with at least three carbon atoms (Carey, 2003; McNaught, 1996). Carbohydrates, simply called sugar, include all saccharides in different carbon lengths and monomeric units of them. A monosaccharide is the simplest unit of carbohydrates with three or more carbon atoms (McNaught, 1996). Glucose (Figure 9.1a) and xylose (Figure 9.1b) are the two most abundant monosaccharides on the planet because they are the key components of plant-cell walls. An oligosaccharide consists of 2–10 monosaccharides connected by glycosidic linkages. When more than 10 monosaccharides are connected to each other by glycosidic linkages, it is often called a polysaccharide, e.g., starch (Figure 9.1c) and cellulose (Figure 9.1d).

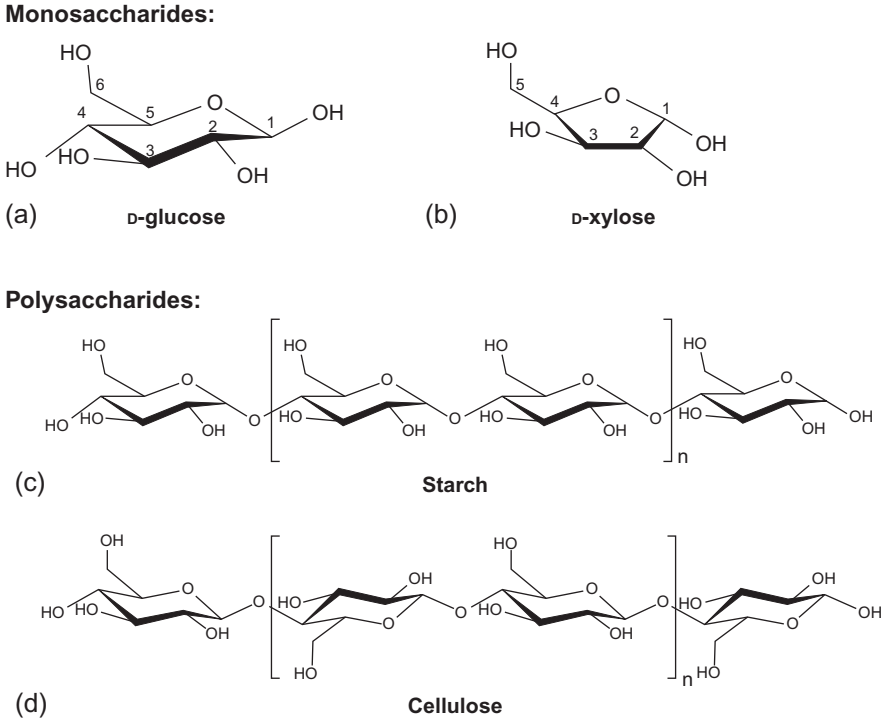

Figure 9.1 Different carbohydrates in their chain length or carbon contents. The most common monosaccharide, D-glucose (a) and the most abundant pentose, five-carbon sugar, D-xylose (b). The two most common polysaccharides are starch (c) and cellulose (d).

Starch and cellulose are different in their glycosidic linkages. Anhydroglucose units of starch are joined by α(1,4)-glycosidic linkages and α(1,6)-glycosidic linkages for branch points, while the same units of cellulose are joined by β(1,4)-glycosidic linkages. Different linkages contribute to their significantly different properties and their biological roles. α(1,4)-Glycosidic linkages in starch result in helical structures, and branched structures of starch make it easy to be hydrolyzed quickly to glucose for the energy metabolism. As the result, starch is a primary source of the energy storage compound for most living organisms (Carey, 2003). β(1,4)-Glycosidic linkages of glucan chains result in linear polysaccharide chains, forming layered sheets tightly held by highly ordered hydrogen bonds, and making cellulose insoluble in water, rigid, inflexible, and low accessible to water and cellulase (Gao et al., 2014; Kirk and Othmer, 2000). Therefore, cellulose serves as the major structural component in plant cell walls.

Carbohydrates produced by plant photosynthesis are the most abundant renewable organic compound on the earth, and its annual production is estimated to be more than 100 billion metric tons (Carey, 2003). Various carbohydrate sources include lignocellulose containing cellulose and hemicellulose, grains (plant seeds) containing starch, and juices of sugarcane and sugar beets containing sucrose. Although cultivated grains and sucrose-producing crops utilize ~30% of arable lands and consume ~70% refresh water withdrawal, their energy content utilized by human beings accounts for approximately 2.5% of the energy stored as lignocellulosic biomass (Zhang, 2013). Only a very small part of lignocellulose (e.g., 5%) is utilized directly or indirectly for meeting human needs, and most of the rest is degraded to CO_2 by the earth biosystem (Naik et al., 2010). Therefore, the utilization of all carbohydrate components in under-utilized lignocellulosic biomass for the production of hydrogen will be important to the bio and hydrogen economy without competing with food sources such as grains and sucrose.

9.1.2 Hydrogen economy and storage

Hydrogen is a colorless, odorless, but very flammable gas. Most hydrogen produced by the reforming of natural gas and coal (Dresselhaus, 2004) is used for petroleum-refining processes and ammonia synthesis (Kirk and Othmer, 2000). The hypothetical hydrogen economy consists of a series of four processes: producing hydrogen, storing hydrogen, distributing hydrogen, and converting the stored hydrogen into electrical energy at the site of end users (Crabtree et al., 2004; Zhang, 2009a). Hydrogen has been suggested as the best alternative fuel to replace fossil fuel-derived fuels primarily used for transportation, mainly due to its clean energy property, high-energy conversion efficiency, and high specific energy density (J/kg) (Crabtree et al., 2004; Dresselhaus, 2004; Zhang, 2009a). Although hydrogen is an attractive energy carrier, the hydrogen economy faces the following technical obstacles: (1) low-cost, scalable, and renewable hydrogen production; (2) a lack of high-density hydrogen storage approaches; (3) high-cost infrastructure for hydrogen distribution; (4) the production of affordable and durable fuel cells; and (5) safety concerns. In order to solve these problems, the use of carbohydrates, such as starch or cellulose, is suggested as an out-of-box solution (Zhang, 2009a; Zhang et al., 2007).

Typical ways to store hydrogen are compressed hydrogen in high-pressure gas cylinders or cryogenic liquid hydrogen (Schlapbach and Zuttel, 2001; Züttel, 2003). However, compressing and liquefying hydrogen processes are too energy-intensive, and their hydrogen storage densities are too low. Because onboard hydrogen storage for a vehicle must be compact, light, safe, and affordable (Schlapbach and Zuttel, 2001), these conventional storage methods cannot fulfill these goals well due to volumetric constrains, high costs, and safety issues (Sakintuna et al., 2007). The goal of novel hydrogen storage approaches for vehicles is to find the materials that can store hydrogen much higher than compressed or cryogenic hydrogen. Intensive research has been conducted to investigate and improve metal hydrides (Chao and Klebanoff, 2012; Imamura et al., 2005; Van Vucht et al., 1970), complex hydrides (Bogdanović et al., 2003; Keaton et al., 2007; Vajo et al., 2005), and carbon materials (Lee and Lee, 2000; Panella et al., 2005). Their relatively high storage capacity of hydrogen and moderate temperature and pressure conditions make them appealing compared to the methods of compressing or cryogenically liquefying hydrogen, but it is not easy to release hydrogen from those materials. The hydrogen storage densities via a number of methods are compared in Figure 9.2, with the hydrogen storage goal proposed by the U.S. Department of Energy (DOE).

Alternatively, some researchers have suggested the production of on-site or on-demand hydrogen via hydrogen carriers such as hydrocarbons (Ahmed and Krumpelt, 2001; Carrette et al., 2001), ethanol (Deluga et al., 2004; Mattos et al., 2012), methanol (Olah, 2005; Pérez-Hernández et al., 2012), ammonia (Klerke et al., 2008; Thomas and Parks, 2006), formic acid (Hull et al., 2012; Schuchmann and Muller, 2013; Yu and Pickup, 2008), and carbohydrates (Ye et al., 2009; Zhang et al., 2007). These different hydrogen carriers are compared based on their costs, typical operational conditions, reactor costs, necessity of product purification, and safety/toxicity (Table 9.1). Hydrocarbons, ethanol, and methanol can be

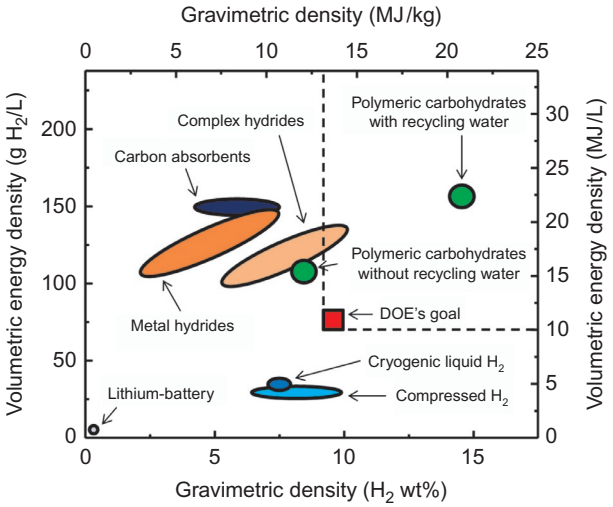

Figure 9.2 The available hydrogen storage means and DOE's goal for hydrogen storage.

Table 9.1 Comparison of hydrogen carriers

Hydrogen carrier	Carrier costs*	Reaction temperature	Reaction pressure	Reactor costs	Purification needed	Safety/ toxicity
Hydrocarbons	$24.9/GJ	≥700 °C	10–40 bar	High	High	Modest
Biodiesel	$39.1/GJ	≥1000 °C	30–80 bar	High	High	Modest
Ethanol	$22.0/GJ	500–700 °C	–	High	Modest	Low
Methanol	$16.8/GJ	200–500 °C	25–50 bar	Modest	Low	Modest
Ammonia	$32.3/GJ	300–700 °C	–	Modest	Modest	Modest
Formic acid	$150/GJ	30–50 °C	~1 bar	Modest	Low	High
Carbohydrates	$12.9/GJ	30–80 °C	~1 bar	Low	No needed	Very low

*These are based on the end users' prices of hydrocarbons (gasoline) for $3.50 per gallon, biodiesel for $5.00 per gallon, ethanol for $2.00 per gallon, methanol for $1.00 per gallon, ammonia for $600 per ton, formic acid for $800 per ton, and delivered carbohydrates for $0.22 per kilogram.

converted to hydrogen and carbon dioxide with carbon monoxide as by-products via several different reforming or oxidation reactions at elevated temperatures. Since a small amount of carbon monoxide is highly toxic to PEM fuel cells (Deluga et al., 2004), extra steps are necessary to remove trace amounts of carbon monoxide, resulting in complicated systems and decreased overall energy efficiencies. Therefore, the concept of hydrocarbons and ethanol as hydrogen carriers for small-size vehicles was abandoned long ago. Relatively low-temperature methanol reforming has been suggested to convert methanol with steam into hydrogen under 20 bars. A recent breakthrough has enabled this reaction to occur under ambient pressure using metal-based catalysts (Nielsen et al., 2013). However, the use of unstable catalysts and low rate of hydrogen production remain challenging. Ammonia is suggested as a hydrogen carrier because it is a carbon-free gas and easily liquefied. However, ammonia is not suitable for PEM fuel cells because trace amounts of ammonia can poison PEM fuel cells (Thomas and Parks, 2006). Recently, formic acid was proposed as another promising hydrogen carrier that can be synthesized from CO_2 and H_2 via either chemical catalysis by iridium metal (Hull et al., 2012) or biocatalysis by enzymes or microorganisms (Schuchmann and Muller, 2013) under mild temperature and pressure conditions. Formic acid can be easily converted back to hydrogen through direct formic acid fuel cells at low temperature (Yu and Pickup, 2008), which offers a great advantage over other hydrogen carriers such as methanol and ethanol requiring high-temperature (>200 °C) reforming processes. However, formic acid has incomparably low hydrogen density (4.3 H_2 wt% for gravimetric density and ~7.5 MJ/L for volumetric energy density) (Yu and Pickup, 2008). In addition to that, highly concentrated formic acid is toxic, very corrosive, and unstable (Reutemann and Kieczka, 1996).

We suggest the use of carbohydrates as a new hydrogen carrier. Carbohydrates have several advantages (Table 9.1). First, carbohydrates are the least costly natural resource and its utilization is a carbon neutral process. Second, the production of highly pure hydrogen is a reality via a new technology called synthetic pathway biotransformation (SyPaB). No product purification steps are required to obtain pure hydrogen for the use with PEM fuel cells. Third, the enzymatic reactions do not require any costly reactors bearing high temperature or pressure because they are conducted under modest reaction conditions (e.g., ~30–80 °C and ~1 bar). Fourth, carbohydrates are nontoxic and nearly inflammable. Therefore, this new biocatalysis breakthrough may make on-site and even onboard vehicle hydrogen production systems feasible.

9.2 Converting carbohydrates to hydrogen by SyPaB

9.2.1 Overview of hydrogen production from carbohydrates

Low-cost renewable carbohydrates are an attractive energy source for green hydrogen production. A number of carbohydrate-to-hydrogen technologies have been under development (Figure 9.3). These technologies can be classified into chemical catalysis, biological transformations, and their hybrids. The theoretical maximum yields of hydrogen production from carbohydrates with water are 12 moles of hydrogen

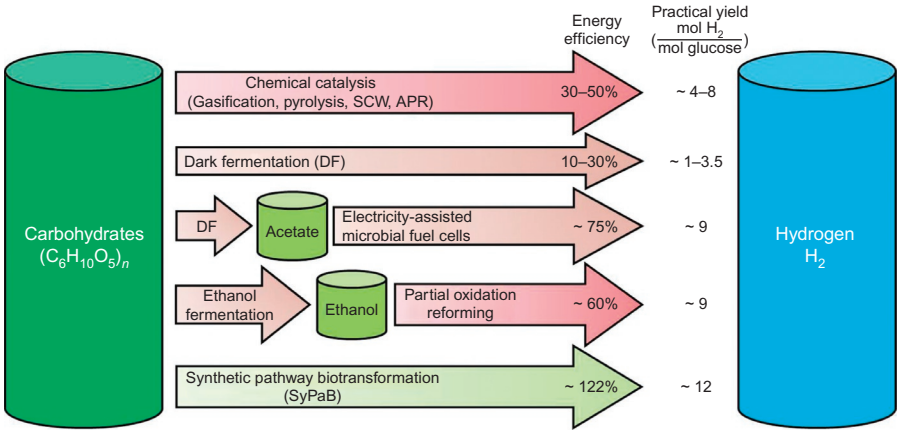

Figure 9.3 Comparison of carbohydrates-to-hydrogen technology.

per mole of hexose and 10 moles of hydrogen per mole of pentose. SyPaB is the only approach that can achieve these theoretical maximum yields. Chemical catalysis features fast reaction rates but harsh reaction conditions (Zhang et al., 2012), while biological transformations (except SyPaB) generally feature modest reaction conditions, easy product separation, and high-purity hydrogen production but low reaction rates and low product yields.

Chemical catalysis includes gasification, pyrolysis, gasification in supercritical water (SCW), and aqueous-phase reforming (APR). Gasification is a process that converts biomass carbohydrates into hydrogen and other products under high temperatures (above 1000 K) and the presence of oxygen and/or water (Navarro et al., 2007). Gasification involves an endothermic reaction of water molecules splitting into H_2 and O by absorbing outside thermal energy or partially combusting carbohydrates. The products of gasification contain some undesired impurities (Rezaiyan and Cheremisinoff, 2005). The regular gasification process has been investigated to produce hydrogen and increase overall yield through the integration with other exothermic processes such as an air-blown bubbling fluidized bed gasifier, a steam reformer, and a water–gas–shift membrane reactor (Ji et al., 2009; Lin et al., 2005). Several reports and reviews about biomass gasification for hydrogen production are also available (Babu, 2002; Cummer and Brown, 2002; Ni et al., 2006; Rezaiyan and Cheremisinoff, 2005; Stevens, 2001). Pyrolysis is a chemical decomposition of biomass at high temperatures (650–800 K) in the absence of oxygen (Rezaiyan and Cheremisinoff, 2005). Unlike gasification process, pyrolysis generally aims to produce liquid bio-oils and solid charcoal plus small amounts of gaseous products (e.g., H_2, CO_2, CO, CH_4) (Ni et al., 2006). The reaction temperature, heating rate, duration, particle size of biomass, and type of catalyst use are important to determine yields of multiple products and their composition (Ni et al., 2006; Probstein and Hicks, 2006; Zanzi et al., 2002). Gaseous products, especially hydrogen, become more favorable under high temperature, high heating rate, and long volatile phase residence time

conditions (Demirbaş, 2002). Other gaseous coproducts, such as methane, hydrocarbon vapors, and carbon monoxide, can be converted into more hydrogen through steam reforming or water–gas-shift reaction. In compared to gasification, pyrolysis is a less favorable process for hydrogen production due to much lower hydrogen yields (Ni et al., 2006). One of the major issues for both pyrolysis and gasification is the formation of tar from biomass, which often causes the formation of other undesired products and undergo polymerization, which creates more complex structures, consequentially affecting the production of hydrogen (Ni et al., 2006). Extensive studies on tar reduction have been reported (Corella et al., 1999; Milne et al., 1998; Narvaez et al., 1997; Simell et al., 1997, 1999; Sutton et al., 2001). When biomass contains high moisture content (above 35%), gasification of the biomass can be conducted in the SCW condition (above 647 K and 220 bars). Under this condition, wet biomass can be rapidly converted into gaseous products at a high gasification ratio of 100% and hydrogen volumetric ratio of 50% (Ni et al., 2006). APR is a chemical catalysis occurring in an aqueous phase under relatively low temperature (400–550 K). Unlike vapor-phase reforming processes, APR produces hydrogen without volatilizing water under high pressure (50–70 bars). It is easy to separate hydrogen gas from aqueous water, which means major energy savings. Moderate conditions of APR can minimize undesirable decomposition reactions and provide a favorable condition for the water–gas-shift reaction so that the process generates hydrogen with low amounts of carbon monoxide (Chheda et al., 2007). However, leaching and instability of catalyst components for APR still remain major disadvantages of the process. These thermochemical processes also accompany the production of other impurities (e.g., CO), which require extra steps of separation and purification to obtain pure hydrogen gas.

Biological transformations mediated by microorganisms or isolated enzymes have advantages over chemical catalysis, such as higher selectivity, lower energy input, possibly higher energy efficiency, and less costly bioreactors (Zhang et al., 2012). Biological transformations include dark fermentation, light fermentation, their combination (Argun et al., 2009), electrically assisted microbial fuel cells (MFCs) (Logan and Regan, 2006), microbial ethanol fermentation followed by ethanol partial oxidation reforming (Haryanto et al., 2005), and cell-free SyPaB (Ye et al., 2009; Zhang et al., 2007). In the dark fermentation, natural or engineered microorganisms can produce only 4 moles of hydrogen and 2 moles of acetate per mole of glucose (33% efficiency), which is called the Thauer limit (Thauer et al., 1977). The practical hydrogen yields are much lower than the Thauer limit (Hallenbeck and Benemann, 2002; Kleerebezem and van Loosdrecht, 2007). To increase overall hydrogen yields, the coproduced two moles of acetate can be further converted to 8 moles of hydrogen through the electrically assisted MFCs with extra energy input to overcome thermodynamically unfavorable hydrogen production from acetate (Logan and Regan, 2006). MFCs provide an overall hydrogen yield of approximately 9 moles of hydrogen produced per mole of glucose (75% efficiency). However, high capital investment required for MFCs and slow volumetric productivity are the greatest challenges for its large-scale application (Logan and Regan, 2006). Alternatively, organic acids produced during the dark fermentation can be utilized by photosynthetic bacteria for hydrogen production (Mathews and Wang, 2009; Miyake et al., 1984), but this process is impractical for potential applications because of its very slow rates of hydrogen

production and low energy concentration (called a nonpoint energy source, which is difficult for energy collection) (Zhang, 2011a). Ethanol fermentation can be conducted with a nearly theoretical yield, i.e., 2 moles of ethanol produced per mole of glucose by yeasts or bacteria (Lin and Tanaka, 2006). Ethanol can be converted to hydrogen by the partial oxidation-reforming process (Deluga et al., 2004; Haryanto et al., 2005). However, the products of such a reforming process still contain a small amount (though serious enough to damage fuel cells) of CO. The overall process of ethanol fermentation followed by partial oxidation reforming has a practical hydrogen production yield similar to that of MFCs, but lower overall energy efficiency than MFCs because of more energy losses during microbial fermentation, distillation, and chemical reforming (Zhang, 2011a). Overall, these microbe-based biological transformations suffer from low hydrogen yields because the microbial basal metabolism competes with hydrogen production. This competition eventually reduces the overall energy efficiency and product yields.

In vitro SyPaB is a new biomanufacturing platform that assembles a number of purified enzymes and cofactors from different sources *in vitro* for the implementation of complicated biochemical reactions. This biotransformation can achieve the theoretical maximum yield of hydrogen from carbohydrates, 12 moles of hydrogen per mole of hexose (Ye et al., 2009; Zhang et al., 2007). High-yield hydrogen is produced from starch by a reconstituted non-natural catabolic pathway consisting of 13 enzymes in four sub-modules: (1) substrate phosphorylation without adenosine triphosphate (ATP), (2) the oxidative pentose phosphate pathway, (3) hydrogen generation, and (4) partial pathways of glycolysis and gluconeogenesis (Figure 9.4). In the substrate phosphorylation catalyzed by phosphorylases, polysaccharides or oligosaccharides release phosphorylated glucose units without the use of ATP (e.g., glucose 1-phosphate), and they are converted into glucose 6-phosphate, entering in the following reactions. The six-carbon sugar glucose 6-phosphate becomes a five-carbon sugar via the oxidative pentose phosphate pathway after the release of CO_2 and 2 moles of nicotinamide adenine dinucleotide phosphate (NADPH). Two moles of hydrogen are produced from continuously regenerated NADPH. In the carbon rearrangement of the pentose phosphate pathway, phosphate pentoses (e.g., ribulose 5-phosphate) are converted into hexose sugars (e.g., fructose 6-phosphate) and three-carbon sugars (e.g., glyceraldehyde 3-phosphate). The three-carbon sugars are combined to form fructose 6-phosphates by multiple enzymes in the glycolysis and gluconeogenesis pathways. Fructose 6-phosphate is further catalyzed to regenerate glucose 6-phosphate for completing the sugar recycling loop. The overall carbohydrate-to-hydrogen reaction mediated by SyPaB can be summarized as $C_6H_{10}O_5(aq) + 7H_2O(l) \rightarrow 12H_2(g) + 6CO_2(g)$. Most of the reactions are reversible, but the formation of gaseous products makes it easy to separate them from the aqueous phase reaction. Thus, the removal of the gaseous products favors the unidirectional overall reaction and drives the reaction forward to completion. The first proof-of-principle experiment has been conducted and shown to be able to produce nearly 12 moles of hydrogen per glucose consumed at the rate of $\sim$0.4 mmol/h/L (Zhang et al., 2007). The reaction rates have been increased by nearly 10-fold through minor optimization of increasing substrate concentration and rate-limiting enzyme loadings (Ye et al., 2009).

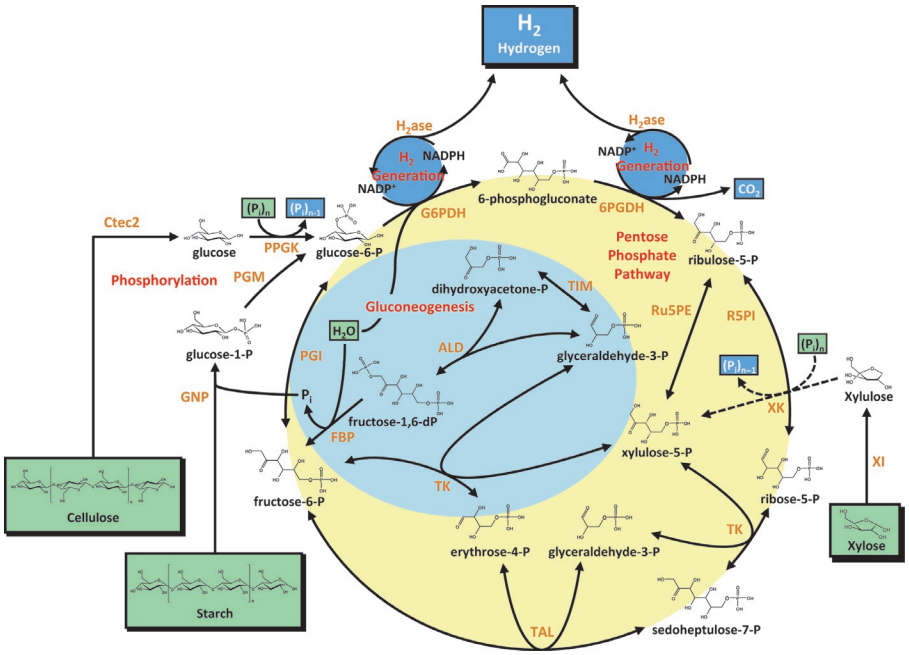

Figure 9.4 Enzymatic synthetic pathway biotransformation (SyPaB) to produce hydrogen from different carbohydrates.

9.2.2 Advantages of SyPaB

Generally speaking, biological transformations are advantageous over chemical catalysis in the modest reaction conditions, lower energy, and high initial capital investment required and relatively purer production of hydrogen gas (Zhang, 2009a). Among various biological transformations, SyPaB has several distinguishable advantages: the highest energy efficiency (122%; more output chemical energy than input chemical energy), the highest product yield, the fastest reaction rates, and the simplest products (only hydrogen and carbon dioxide).

SyPaB has the highest energy efficiency in converting carbohydrates to hydrogen. The overall carbohydrate-to-hydrogen reactions are spontaneous ($\Delta G^\circ = \sim -50$ kJ/mol) and endothermic ($\Delta H^\circ = +596$ kJ/mol), meaning that some of ambient thermal energy, which is available from any low-temperature waste heat sources such as refrigerators, air conditioners, and fuel cells, is absorbed and converted to chemical energy used for hydrogen production (Ye et al., 2009; Zhang et al., 2007). This thermodynamic analysis suggests that the overall reactions are driven by entropy gain from the aqueous-to-gaseous phase change, rather than enthalpy losses. These enzymatic reactions for hydrogen production are among rare entropy-driven chemical reactions involving phase changes from highly ordered substrates to poorly ordered products: for example, $N_2O_5(s) \rightarrow 2NO_2(g) + O_2(g)$ and $Ba(OH)_2{\cdot}8H_2O(s) + 2\,NH_4SCN\,(s) \rightarrow Ba(SCN)_2\,(aq) + 2NH_3\,(aq) + 10H_2O\,(l)$, where both reactions involve

great phase changes from solid substrates to gaseous, or aqueous and liquid products, respectively. As the result, these entropy-driven reactions can generate more output of chemical energy in the form of hydrogen (122%; based on combustion energy) than input of chemical energy in carbohydrates (Figure 9.3).

SyPaB has higher reaction rates than microbial fermentations because the absence of a cellular membrane and microbial complexity allows faster mass transfer and higher biocatalyst density per volume (Rollin et al., 2013; Zhang and Mielenz, 2011). For example, enzymatic fuel cells can usually generate the higher maximum power than microbial fuel cells due to the absence of cellular membranes limiting mass transfer and high enzyme loading achieved per unit volume or area without unnecessary biomolecules in the reaction solution (Cooney et al., 2008; Osman et al., 2011; Zhu et al., 2014). Today the highest enzymatic hydrogen production rate has been increased to 0.3 g H_2/L/h, compared to the highest microbial hydrogen fermentation on sugars and faster than electrically assisted MFCs.

SyPaB produces H_2 with CO_2 in a ratio of 67% to 33% (Ye et al., 2009; Zhang et al., 2007). The separation cost of hydrogen from carbon dioxide is estimated to be very low; the mixture of hydrogen and carbon dioxide can also be used for PEM fuel cells with only ~1% loss in the fuel cell efficiency because carbon dioxide is an inert gas (Huang and Zhang, 2011a). However, hydrogen production by chemical catalysis contains a small but critical amount of carbon monoxide for damaging fuel cells. Therefore, the high selectivity of SyPaB makes carbohydrates as a possible hydrogen carrier.

9.2.3 Unique advantages of carbohydrates as a hydrogen carrier

The major obstacle of the hydrogen economy is a lack of high-density hydrogen storage approaches at low costs. Clearly, carbohydrates as a hydrogen carrier have several advantages over other hydrogen carriers, such as hydrocarbons (Ahmed and Krumpelt, 2001; Carrette et al., 2001), ethanol (Deluga et al., 2004; Mattos et al., 2012), methanol (Olah, 2005; Pérez-Hernández et al., 2012), ammonia (Klerke et al., 2008; Thomas and Parks, 2006), and formic acid (Hull et al., 2012; Schuchmann and Muller, 2013; Yu and Pickup, 2008). Table 9.1 compares these different hydrogen carriers based on their costs, typical operation conditions, reaction requirements, and safety.

First, the use of carbohydrates is economically beneficial due to its abundant renewability in nature and low costs. Roughly more than 100 billion tons of dry cellulosic material is produced by terrestrial plants annually (Berner, 2003; Falkowski et al., 2000). It is estimated that about 150 billion gallons of gasoline can be replaced with hydrogen energy produced from 700 million tons of biomass via SyPaB (Huang and Zhang, 2011b). Therefore, only 10% utilization of lignocellulose biomass could greatly reduce our energy dependence on fossil fuel. In addition to the economic benefits from renewable carbohydrates as a hydrogen carrier, carbohydrates are nontoxic, nearly inflammable, and well distributed over the world (Zhang, 2013).

Second, carbohydrates have very high hydrogen storage density (Figure 9.2). Polysaccharides have the chemical formula of $(C_6H_{10}O_5)_n$. Gravimetric density of hydrogen in carbohydrates is 14.8 H_2 mass% (gram of hydrogen gas production per gram of

polysaccharide consumption = 24 g/162 g) when water molecules are recycled from PEM fuel cells. Volumetric energy density of polysaccharide as hydrogen carrier is much higher than 100 kg/m^3. Clearly, carbohydrates exceed the DOE's hydrogen storage requirement.

Third, carbohydrates are a carbon-neutral energy source in terms of the whole life cycle. The amount of CO_2 released during hydrogen production from carbohydrates via SyPaB would be equal to CO_2 consumed for growing carbohydrates used (Kirk and Othmer, 2000). Thus, the use of carbohydrates as a fuel to produce hydrogen can achieve nearly zero net CO_2 emissions.

Carbohydrates as an energy carrier with these advantages could be the ultimate solution for many energy sustainability challenges such as hydrogen and electricity storage, CO_2 fixation and long-term storage, and transportation fuel production (Zhang and Mielenz, 2011).

9.3 Challenges of carbohydrates as hydrogen storage and respective solutions

Carbohydrates become a possible hydrogen carrier mainly due to SyPaB, which can completely convert carbohydrates with water to highly pure hydrogen under modest reaction conditions. The hydrogen production costs from carbohydrates are strongly related to three major cost factors: carbohydrate costs, enzymes costs, and cofactors costs (Zhang and Mielenz, 2011). The hydrogen generation rates catalyzed by SyPaB determine its potential applications.

9.3.1 Enzyme costs and stability

The costs of enzymes used in SyPaB are influenced greatly by their production costs, i.e., US dollars per kilogram enzyme and their turn-over number (TTN, mole product per mole of enzyme). All enzymes are cytoplasmic enzymes instead of membrane proteins so that all of them can be produced as recombinant proteins in *Escherichia coli*. According to industrial bulk enzyme production knowledge, such bulk enzymes have production costs of ~$5–100/kg dry protein weight (Zhang, 2011a). Increasing TTN of enzymes exponentially decreases enzyme costs (Zhang, 2010a). Enzyme stability can be improved through enzyme immobilization (Wang and Zhang, 2010), the use of thermostable enzymes (Ye et al., 2010), and protein engineering (Eijsink et al., 2004, 2005; Ye et al., 2012). It is estimated that the enzyme costs would be low enough for industrial-scale use when their TTNs reach more than 10^7–10^8 (Wang et al., 2011; Zhang, 2010a; Zhang et al., 2010). Such high TTNs have been shown to be achievable with the use of thermostable enzymes from thermophilic microbes, such as *Clostridium thermocellum* and *Thermotoga maritima* (Myung et al., 2010, 2011; Wang and Zhang, 2009, 2010). The use of thermostable enzymes produced in mesophilic hosts can also decrease the enzyme purification cost. Enzymes cloned from thermophilic microbes are usually more stable at higher temperature (60–70 °C) (Zhang, 2009b).

Therefore, heat treatment can deactivate *E. coli* endogenous proteins; after centrifugation, only the soluble thermostable proteins remain in the supernatant. This heat precipitation provides a simple and less-costly purification method for thermostable enzymes. In contrast, classic enzyme purification methods in labs, such as chromatographic techniques, are costly and hard to scale up. Immobilization of enzymes not only increases their TTNs but also helps recycle enzymes from the reactants (Zhang, 2009b). Enzyme immobilization is a relatively mature technique, and low-cost enzyme immobilization techniques have been widely used, such as alginate entrapment, cross-linking enzyme aggregate (Sheldon et al., 2007). Furthermore, the cellulose-binding-module-tagged protein immobilization can integrate the purification and immobilization of enzymes in one step (Hong et al., 2008).

Stable enzymes could be stored for long periods, depending on the storage condition and its inherent properties. Proper storage conditions for some enzymes can keep them active up to several years. For example, the shelf life of the proteases used in detergents is several years. Immobilized thermostable glucose isomerase, widely used in the food industry, has a working lifetime of up to 2 years at 55 °C. Immobilized glucose oxidase, used in blood sugar testing strips, can be stored for more than 2 years at ambient temperature.

9.3.2 Cofactor costs and stability

Cofactors are nonprotein chemical compounds required for enzyme functions. Economic analysis suggests that TTN value of cofactors should be higher than 10^6 for the economically viable production of biocommodities (Zhang, 2010a). The issues from cofactor costs and stability can be addressed by the cofactor recycling systems (Liu and Wang, 2007) or the use of low-cost stable biomimetic cofactors (Ansell and Lowe, 1999; Lutz et al., 2004). Efficient regeneration of natural cofactors through recycling systems is economically beneficial to most enzymatic reactions, such as ATP and nicotinamide cofactors (NAD^+ and $NADP^+$) (Liu and Wang, 2007). Among several different $NAD(P)^+$ recycling systems, enzymatic cofactor regeneration is the most promising approach due to its high selectivity and high compatibility with other biological components of the reaction (Weckbecker et al., 2010). Cofactors can be recycled by either reaction-coupled regeneration where a second enzyme regenerates the cofactor by catalyzing a second substrate, or substrate-coupled regeneration where one single enzyme is utilized for both desired product formation and cofactor regeneration. Immobilization of the cofactor has been developed for more efficient cofactor regeneration (Liu and Wang, 2007).

The best solution would be the replacement of native $NAD(P)^+$ cofactors with biomimetic cofactors (mNADs) featuring low cost and high stability. However, most wild-type NAD(P)-dependent redox enzymes cannot work on mNADs, except for horse liver alcohol dehydrogenase (Lo and Fish, 2002), enoate reductase (Paul et al., 2013), and diaphorase (unpublished). Therefore, it is necessary to engineer cofactor preference of Rossmann-motif-based redox enzymes. Cofactor engineering is still in its developing stage, but several successful results have been reported recently (Bastian et al., 2011; Campbell et al., 2010, 2012; Döhr et al., 2001;

Huang et al., 2013; Paul et al., 2013). We also succeeded in converting the cofactor of NADP-dependent glucose 6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase to the biomimetic cofactor (data unpublished).

9.3.3 Reaction rate

The highest enzymatic hydrogen production rate achieved by us is about 150 mmol/h/L (data unpublished), which was increased by nearly 750 times in the past 10 years (Woodward et al., 2000). Such great enhancements are attributed to elevated reaction temperature, optimized enzyme ratio, high substrate concentration, and high enzyme loading. This reaction rate equals approximately 2 g glucose consumed/L/h. This enzymatic hydrogen production rate is comparative to the ethanol productivity of microbial ethanol fermentation, suggesting the technical feasibility of stationary hydrogen generation systems using SyPaB.

Given that a small-size hydrogen fuel cell vehicle needs to be powered with 1 kg H_2 per hour and the vehicle has a 50-liter on-board bioreactor, the expected hydrogen production rate required for the sugar fuel cell vehicle (SFCV) is estimated to be about 20 g H_2/L/h. This means we have to increase enzymatic hydrogen generation rates by another 60 times.

Table 9.2 shows an analysis of potential rate increases by several orders of magnitude through a combination of known technologies. Increasing reaction temperature from 30 to 80 °C (or even higher) may result in about 32-fold increased reaction rates, according to the Q10 effect where the reaction rate approximately doubles with every 10 °C increase in reaction temperature (Reyes et al., 2008). This trend is supported by

Table 9.2 Potential rate increases for hydrogen production from SyPaB (Zhang, 2011b)

Technology	Potential fold	References	Predicted fold
Increasing reaction temperature from 30 to 80 °C or higher	32	Reyes et al. (2008), Wang and Zhang (2009), and Ye et al. (2009)	4–20
Increasing the rate-limiting step enzyme loading	10	Ye et al. (2009)	2–5
Increasing overall enzyme loading	10	Yoshida et al. (2005)	5
Increasing substrate concentration by 50-fold	10	Ye et al. (2009)	5
Creating substrate channeling among enzymes	2–100	Myung et al. (2010), Srivastava and Bernhard (1986), and Zhang (2011c)	2
Improving catalytic efficiency of enzymes	~10		
Overall increasing rates	640,000–32,000,000		400–5000

the cases of thermostable enzymes (Ma et al., 1994; Wang and Zhang, 2009; Ye et al., 2009). Increasing the loading of enzymes in the rate-limiting steps can practically increase at least twofold of the reaction rate (Ye et al., 2009). Increasing overall enzyme concentration by 10-fold or higher can also achieve another fivefold increase (Yoshida et al., 2005). Increasing substrate concentration by 50-fold or higher is expected to be able to increase the reaction rate by fivefold. Creating substrate channeling among cascade enzymes has been shown to increase the reaction rate because cascade enzymes are closely held by scaffoldins or other reagents (Conrado et al., 2008; Dueber et al., 2009; Wilner et al., 2009; You et al., 2012). Lastly, improving the catalytic efficiency of enzymes would result in a potential increase of about a 10-fold in the reaction rate. Most enzymes have far lower catalytic efficiencies than those of catalytically perfect enzymes (Wolfenden and Snider, 2001) or the same catalytic enzymes with the highest efficiencies reported in the BRENDA database. The overall potential increases in hydrogen production rate can be estimated to be from 640,000 to 32,000,000-fold from the original starting point (Woodward et al., 2000). In previous years, we accomplished about a 750-fold enhancement in carbohydrate-to-hydrogen reaction rates and believe that further reaction enhancement to ~20 g H_2/L/h will be achievable soon. This is because microbial fermentation has already been reported to produce hydrogen with the rate of 23.6 H_2/L/h (Yoshida et al., 2005) and enzymatic reactions usually have faster reaction rates than microbial fermentation (Cooney et al., 2008; Zhang, 2009b; Zhu et al., 2014).

9.4 Future carbohydrate-to-hydrogen systems

The enzymatic carbohydrate-to-hydrogen production opens up several potential applications for the hydrogen economy, where hydrogen is an alternative transportation fuel or a short-term electricity storage carrier. The potential applications include from the most feasible one, i.e., local hydrogen-generating stations (Figure 9.5a), to the most ambitious application, i.e., SFCVs (Zhang, 2009a, 2010b) (Figure 9.5b).

9.4.1 Local hydrogen-generating stations

The near-term application is hydrogen production from local carbohydrate resources in small-size hydrogen generation reactors supplying hydrogen to local users, such as fuel cell vehicles (Figure 9.5a). In contrast with the hydrogen distribution from a centralized hydrogen generation facility to local hydrogen stations, the proposed hydrogen generation and distribution systems are based on local carbohydrates, which would require low capital investment and have a high safety level. Local hydrogen-generating stations system will be supplied with carbohydrates easily delivered from local areas. Each local hydrogen-generating station can have a relatively large size high-pressure hydrogen storage tank under the ground, which is refilled with hydrogen produced from underground bioreactors, like anaerobic digesters (Zhang et al., 2012). Such small-size decentralized hydrogen generation systems could promote hydrogen fuel cell vehicles without costly hydrogen distribution infrastructure.

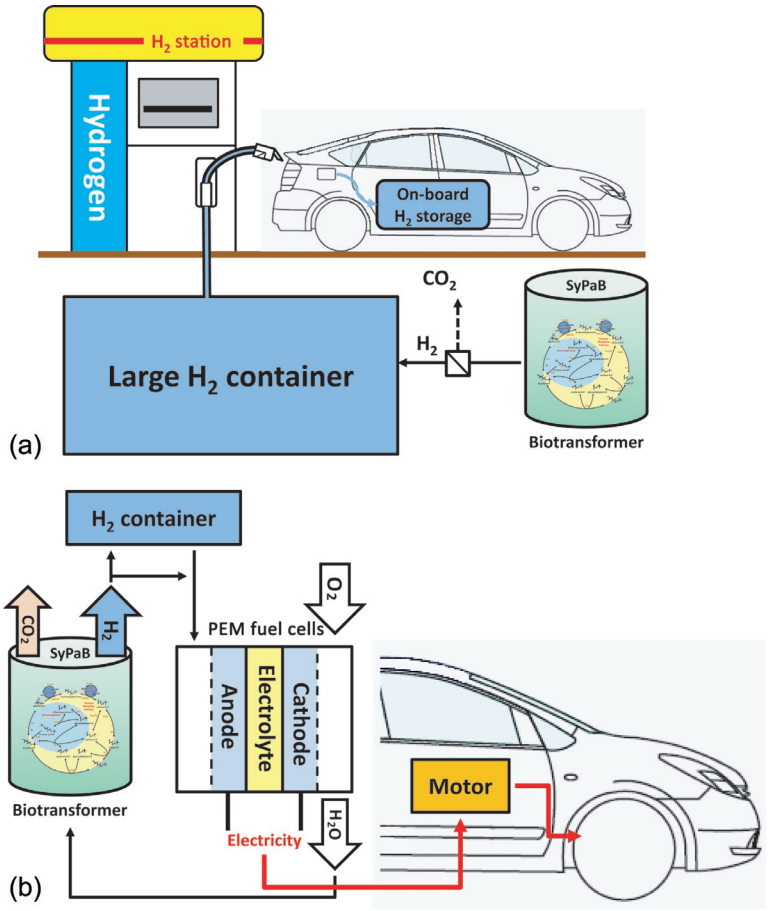

Figure 9.5 Future applications for hydrogen economy. Local hydrogen-generating station (a) and conceptual power train system for SFCVs using onboard sugar-to-hydrogen bioreactor (b).

These sugar-to-hydrogen stations may combine with fuel cells for providing low-cost electricity generation to individual houses or remote areas. The whole systems could have very high energy efficiency because the exothermic reaction of generating electricity from fuels cells is coupled with the endothermic reaction of producing hydrogen from carbohydrates via SyPaB, and the residual heat can cogenerate hot water (Zhang, 2009a).

9.4.2 *Sugar fuel cell vehicles*

The most ambitious application is SFCVs (Figure 9.5b). These hypothetical SFCVs can be powered by electricity generated from PEM fuel cells, which, in turn, can be supplied hydrogen generated from an onboard SyPaB bioreactor that can produce

on-demand hydrogen fast enough. SFCVs can be refilled with solid carbohydrate fuels or a carbohydrate/water slurry within minutes. Highly pure hydrogen generated via SyPaB can greatly decrease the complexity of the whole power train system and increase the operation performance (Zhang, 2009a). Water and heat generated from PEM fuel cells can be recycled to the bioreactor and maintain the reaction temperature of the bioreactor, respectively. Rechargeable batteries will be needed to assist the power train system during the vehicle starting-up or accelerating like typical FCVs (Zhang, 2009a). One important factor of onboard hydrogen storage/production systems is the spatial restriction. At the current hydrogen production rate of 150 mmol/L/h, it requires a reaction volume of 3.33 m^3 to produce 1 kg of hydrogen per hour. This reaction volume seems technically unrealistic, but it is expected to be able to increase the reaction rate by up to another 60-fold (see Section 9.3.3). The advantages of SFCVs are the highest biomass-to-wheel energy conversion efficiency (overall, 55%; carbohydrate-to-hydrogen, 122%; hydrogen-PEM fuel cells, 50%; electricity-motor, 90%), no special infrastructure needed, a high safety level, and the use of carbon-neutral carbohydrates (Zhang, 2009a).

9.5 Conclusions

Hydrogen is the most promising alternative transportation fuel, i.e., high-energy efficiency via PEM fuel cells (Zhang and Mielenz, 2011) and high-energy storage density relative to rechargeable batteries. Among a number of carbohydrate-to-hydrogen technologies (Figure 9.3), SyPaB is the most promising method to producing low-cost hydrogen due to the highest yield, highly pure hydrogen production, modest reaction conditions, competitive reaction rates, and low-cost bioreactor. As compared to other hydrogen carriers, renewable carbohydrates are attractive because of their abundance, renewability, low cost, and even distribution, and they have best hydrogen storage density and safety. However, a few obstacles need to be solved for commercial hydrogen production via SyPaB, such as low-cost and highly stable enzyme production and the use of biomimetic cofactors. It would be a dream that these remaining technical obstacles are solved by a single laboratorial effort. More realistically, it requires international collaborative research to solve those challenges and bring about a new carbon-neutral hydrogen economy.

9.6 Sources of future information and advice

The seminal concept of hydrogen produced from carbohydrates and water via a synthetic enzymatic pathway was first proposed in 2007 (Zhang et al., 2007). Later, high-yield hydrogen was produced from cellulosic materials (Ye et al., 2009) and xylose (Martin del Campo et al., 2013). The recent review covers the advances in SyPaB and remaining technical challenges (Rollin et al., 2013).

The concept of carbohydrates as a hydrogen carrier was proposed first in 2007 (Zhang et al., 2007), and its potential impact on the sustainable revolution was analyzed in 2013 (Zhang, 2013).

References

Ahmed, S., Krumpelt, M., 2001. Hydrogen from hydrocarbon fuels for fuel cells. Int. J. Hydrog. Energy 26 (4), 291–301.

Ansell, R.J., Lowe, C.R., 1999. Artificial redox coenzymes: biomimetic analogues of NAD^+. Appl. Microbiol. Biotechnol. 51 (6), 703–710.

Argun, H., Kargi, F., Kapdan, I.K., 2009. Hydrogen production by combined dark and light fermentation of ground wheat solution. Int. J. Hydrog. Energy 34 (10), 4305–4311.

Babu, S.P., 2002. Biomass Gasification for Hydrogen Production–Process Description and Research Needs. IEA, Gas Technology Institute, IL, USA, Revised draft (29.10.02.).

Bastian, S., Liu, X., Meyerowitz, J.T., Snow, C.D., Chen, M.M., Arnold, F.H., 2011. Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. Metab. Eng. 13 (3), 345–352.

Berner, R.A., 2003. The long-term carbon cycle, fossil fuels and atmospheric composition. Nature 426 (6964), 323–326.

Bogdanović, B., Felderhoff, M., Kaskel, S., Pommerin, A., Schlichte, K., Schüth, F., 2003. Improved hydrogen storage properties of Ti-doped sodium alanate using titanium nanoparticles as doping agents. Adv. Mater. 15 (12), 1012–1015.

Campbell, E., Wheeldon, I.R., Banta, S., 2010. Broadening the cofactor specificity of a thermostable alcohol dehydrogenase using rational protein design introduces novel kinetic transient behavior. Biotechnol. Bioeng. 107 (5), 763–774.

Campbell, E., Meredith, M., Minteer, S.D., Banta, S., 2012. Enzymatic biofuel cells utilizing a biomimetic cofactor. Chem. Commun. 48 (13), 1898–1900.

Carey, F.A., 2003. Organic Chemistry. McGraw-Hill, Boston, xxxii, 1191, 131 pp.

Carrette, L., Friedrich, K., Stimming, U., 2001. Fuel cells—fundamentals and applications. Fuel Cells 1 (1), 5–39.

Chao, B., Klebanoff, L., 2012. Hydrogen storage in interstitial metal hydrides. Hydrogen Storage Technol. Mater. Appl. 109.

Chheda, J.N., Huber, G.W., Dumesic, J.A., 2007. Liquid-phase catalytic processing of biomass-derived oxygenated hydrocarbons to fuels and chemicals. Angew. Chem. Int. Ed. 46 (38), 7164–7183.

Conrado, R.J., Varner, J.D., DeLisa, M.P., 2008. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. Curr. Opin. Biotechnol. 19 (5), 492–499.

Cooney, M.J., Svoboda, V., Lau, C., Martin, G., Minteer, S.D., 2008. Enzyme catalysed biofuel cells. Energy Environ. Sci. 1 (3), 320–337.

Corella, J., Aznar, M.-P., Gil, J., Caballero, M.A., 1999. Biomass gasification in fluidized bed: where to locate the dolomite to improve gasification? Energy Fuel 13 (6), 1122–1127.

Crabtree, G.W., Dresselhaus, M.S., Buchanan, M.V., 2004. The hydrogen economy. Phys. Today 57, 39.

Cummer, K.R., Brown, R.C., 2002. Ancillary equipment for biomass gasification. Biomass Bioenergy 23 (2), 113–128.

Deluga, G., Salge, J., Schmidt, L., Verykios, X., 2004. Renewable hydrogen from ethanol by autothermal reforming. Science 303 (5660), 993–997.

Demirbaş, A., 2002. Gaseous products from biomass by pyrolysis and gasification: effects of catalyst on hydrogen yield. Energy Convers. Manag. 43 (7), 897–909.

Döhr, O., Paine, M.J., Friedberg, T., Roberts, G.C., Wolf, C.R., 2001. Engineering of a functional human NADH-dependent cytochrome P450 system. Proc. Natl. Acad. Sci. U. S. A. 98 (1), 81–86.

Dresselhaus, M.S., 2004. Basic research needs for the hydrogen economy. Abstr. Pap. Am. Chem. Soc. 227, U1084–U1085.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.L.J., Keasling, J.D., 2009. Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. 27 (8), 753–759.

Eijsink, V.G.H., Bjork, A., Gaseidnes, S., Sirevag, R., Synstad, B., van den Burg, B., Vriend, G., 2004. Rational engineering of enzyme stability. J. Biotechnol. 113 (1–3), 105–120.

Eijsink, V.G.H., Gaseidnes, S., Borchert, T.V., van den Burg, B., 2005. Directed evolution of enzyme stability. Biomol. Eng. 22 (1–3), 21–30.

Falkowski, P., Scholes, R., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., 2000. The global carbon cycle: a test of our knowledge of earth as a system. Science 290 (5490), 291–296.

Gao, S., You, C., Renneckar, S., Bao, J., Zhang, Y.-H.P., 2014. New insights into enzymatic hydrolysis of heterogeneous cellulose by using carbohydrate-binding module 3 containing GFP and carbohydrate-binding module 17 containing CFP. Biotechnol. Biofuels 7 (1), 24.

Hallenbeck, P.C., Benemann, J.R., 2002. Biological hydrogen production; fundamentals and limiting processes. Int. J. Hydrog. Energy 27 (11), 1185–1193.

Haryanto, A., Fernando, S., Murali, N., Adhikari, S., 2005. Current status of hydrogen production techniques by steam reforming of ethanol: a review. Energy Fuel 19 (5), 2098–2106.

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

Huang, W.-D., Zhang, Y.-H.P., 2011a. Analysis of biofuels production from sugar based on three criteria: thermodynamics, bioenergetics, and product separation. Energy Environ. Sci. 4 (3), 784–792.

Huang, W.-D., Zhang, Y.-H.P., 2011b. Energy efficiency analysis: biomass-to-wheel efficiency related with biofuels production, fuel distribution, and powertrain systems. Plos One 6 (7), e22113.

Huang, Q., Wu, J.W., Xu, H.J., 2013. Biomimetic hydrogenation: a reusable NADH coenzyme model for hydrogenation of alpha, beta-epoxy ketones and 1,2-diketones. Tetrahedron Lett. 54 (29), 3877–3881.

Hull, J.F., Himeda, Y., Wang, W.H., Hashiguchi, B., Periana, R., Szalda, D.J., Muckerman, J.T., Fujita, E., 2012. Reversible hydrogen storage using CO_2 and a proton-switchable iridium catalyst in aqueous media under mild temperatures and pressures. Nat. Chem. 4 (5), 383–388.

Imamura, H., Masanari, K., Kusuhara, M., Katsumoto, H., Sumi, T., Sakata, Y., 2005. High hydrogen storage capacity of nanosized magnesium synthesized by high energy ball-milling. J. Alloys Compd. 386 (1), 211–216.

Ji, P., Feng, W., Chen, B., 2009. Production of ultrapure hydrogen from biomass gasification with air. Chem. Eng. Sci. 64 (3), 582–592.

Keaton, R.J., Blacquiere, J.M., Baker, R.T., 2007. Base metal catalyzed dehydrogenation of ammonia-borane for chemical hydrogen storage. J. Am. Chem. Soc. 129 (7), 1844–1845.

Kirk, R.E., Othmer, D.F., 2000. Kirk-Othmer Encyclopedia of Chemical Technology. Wiley, New York.

Kleerebezem, R., van Loosdrecht, M., 2007. Mixed culture biotechnology for bioenergy production. Curr. Opin. Biotechnol. 18 (3), 207–212.

Klerke, A., Christensen, C.H., Nørskov, J.K., Vegge, T., 2008. Ammonia for hydrogen storage: challenges and opportunities. J. Mater. Chem. 18 (20), 2304–2310.

Lee, S.M., Lee, Y.H., 2000. Hydrogen storage in single-walled carbon nanotubes. Appl. Phys. Lett. 76 (20), 2877–2879.
Lin, Y., Tanaka, S., 2006. Ethanol fermentation from biomass resources: current state and prospects. Appl. Microbiol. Biotechnol. 69 (6), 627–642.
Lin, S., Harada, M., Suzuki, Y., Hatano, H., 2005. Process analysis for hydrogen production by reaction integrated novel gasification (HyPr-RING). Energy Convers. Manag. 46 (6), 869–880.
Liu, W.F., Wang, P., 2007. Cofactor regeneration for sustainable enzymatic biosynthesis. Biotechnol. Adv. 25 (4), 369–384.
Lo, H.C., Fish, R.H., 2002. Biomimetic NAD^+ models for tandem cofactor regeneration, horse liver alcohol dehydrogenase recognition of 1,4-NADH derivatives, and chiral synthesis. Angew. Chem. Int. Ed. 41 (3), 478–481.
Logan, B.E., Regan, J.M., 2006. Microbial fuel cells-challenges and applications. Environ. Sci. Technol. 40 (17), 5172–5180.
Lutz, J., Hollmann, F., Ho, T.V., Schnyder, A., Fish, R.H., Schmid, A., 2004. Bioorganometallic chemistry: biocatalytic oxidation reactions with biomimetic NAD^+/NADH co-factors and $[Cp^*Rh(bpy)H]^+$ for selective organic synthesis. J. Organomet. Chem. 689 (25), 4783–4790.
Ma, K.S., Zhou, Z.H., Adams, M.W.W., 1994. Hydrogen-production from pyruvate by enzymes purified from the hyperthermophilic archaeon, *Pyrococcus furiosus*: a key role for NADPH. Fems Microbiol. Lett. 122 (3), 245–250.
Martin del Campo, J.S., Rollin, J., Myung, S., Chun, Y., Chandrayan, S., Patino, R., Adams, M.W.W., Zhang, Y.-H.P., 2013. High-yield production of dihydrogen from xylose by using a synthetic enzyme cascade in a cell-free system. Angew. Chem. Int. Ed. 52 (17), 4587–4590.
Mathews, J., Wang, G., 2009. Metabolic pathway engineering for enhanced biohydrogen production. Int. J. Hydrog. Energy 34 (17), 7404–7416.
Mattos, L.V., Jacobs, G., Davis, B.H., Noronha, F.b.B., 2012. Production of hydrogen from ethanol: review of reaction mechanism and catalyst deactivation. Chem. Rev. 112 (7), 4094–4123.
McNaught, A.D., 1996. International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology—Joint Commission on Biochemical Nomenclature—Nomenclature of carbohydrates—Recommendations 1996. Pure Appl. Chem. 68 (10), 1919–2008.
Milne, T.A., Abatzoglou, N., Evans, R.J., 1998. Biomass Gasifier "tars": Their Nature, Formation, and Conversion. National Renewable Energy Laboratory, Golden, CO.
Miyake, J., Mao, X., Kawamura, S., 1984. Photoproduction of hydrogen from glucose by a co-culture of a photosynthetic bacterium and Clostridium butyricum. J. Ferment. Technol. 62 (6), 531–535.
Myung, S., Wang, Y.R., Zhang, Y.-H.P., 2010. Fructose-1,6-bisphosphatase from a hyperthermophilic bacterium *Thermotoga maritima*: characterization, metabolite stability, and its implications. Process Biochem. 45 (12), 1882–1887.
Myung, S., Zhang, X.Z., Zhang, Y.-H.P., 2011. Ultra-stable phosphoglucose isomerase through immobilization of cellulose-binding module-tagged thermophilic enzyme on low-cost high-capacity cellulosic adsorbent. Biotechnol. Prog. 27 (4), 969–975.
Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: a comprehensive review. Renew. Sustain. Energy Rev. 14 (2), 578–597.
Narvaez, I., Corella, J., Orío, A., 1997. Fresh tar (from a biomass gasifier) elimination over a commercial steam-reforming catalyst. Kinetics and effect of different variables of operation. Ind. Eng. Chem. Res. 36 (2), 317–327.

Navarro, R., Pena, M., Fierro, J., 2007. Hydrogen production reactions from carbon feedstocks: fossil fuels and biomass. Chem. Rev. 107 (10), 3952–3991.
Ni, M., Leung, D.Y.C., Leung, M.K.H., Sumathy, K., 2006. An overview of hydrogen production from biomass. Fuel Process. Technol. 87 (5), 461–472.
Nielsen, M., Alberico, E., Baumann, W., Drexler, H.-J., Junge, H., Gladiali, S., Beller, M., 2013. Low-temperature aqueous-phase methanol dehydrogenation to hydrogen and carbon dioxide. Nature 495 (7439), 85–89.
Olah, G.A., 2005. Beyond oil and gas: the methanol economy. Angew. Chem. Int. Ed. 44 (18), 2636–2639.
Osman, M.H., Shah, A.A., Walsh, F.C., 2011. Recent progress and continuing challenges in bio-fuel cells. Part I: enzymatic cells. Biosens. Bioelectron. 26 (7), 3087–3102.
Panella, B., Hirscher, M., Roth, S., 2005. Hydrogen adsorption in different carbon nanostructures. Carbon 43 (10), 2209–2214.
Paul, C.E., Gargiulo, S., Opperman, D.J., Lavandera, I., Gotor-Fernandez, V., Gotor, V., Taglieber, A., Arends, I.W.C.E., Hollmann, F., 2013. Mimicking nature: synthetic nicotinamide cofactors for C=C bioreduction using enoate reductases. Org. Lett. 15 (1), 180–183.
Pérez-Hernández, R., Mendoza-Anaya, D., Martínez, A.G., Gómez-Cortés, A., 2012. Catalytic steam reforming of methanol to produce hydrogen on supported metal catalysts. Hydrogen Energy: Challenges Perspect, 149–256.
Probstein, R.F., Hicks, R.E., 2006. Synthetic Fuels. Courier Dover Publications, New York.
Reutemann, W., Kieczka, H., 1996. Formic acid. Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH, Weinheim, Germany.
Reyes, B.A., Pendergast, J.S., Yamazaki, S., 2008. Mammalian peripheral circadian oscillators are temperature compensated. J. Biol. Rhythms 23 (1), 95.
Rezaiyan, J., Cheremisinoff, N.P., 2005. Gasification Technologies: A Primer for Engineers and Scientists. CRC Press, Boca Raton, FL.
Rollin, J.A., Tam, T.K., Zhang, Y.-H.P., 2013. New biotechnology paradigm: cell-free biosystems for biomanufacturing. Green Chem. 15 (7), 1708–1719.
Sakintuna, B., Lamari-Darkrim, F., Hirscher, M., 2007. Metal hydride materials for solid hydrogen storage: a review. Int. J. Hydrog. Energy 32 (9), 1121–1140.
Schlapbach, L., Zuttel, A., 2001. Hydrogen-storage materials for mobile applications. Nature 414 (6861), 353–358.
Schuchmann, K., Muller, V., 2013. Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. Science 342 (6164), 1382–1385.
Sheldon, R.A., Sorgedrager, M., Janssen, M.H.A., 2007. Use of cross-linked enzyme aggregates (CLEAs) for performing biotransformations. Chim. Oggi Chem. Today 25 (1), 48–52.
Simell, P.A., Hakala, N.A., Haario, H.E., Krause, A.O.I., 1997. Catalytic decomposition of gasification gas tar with benzene as the model compound. Ind. Eng. Chem. Res. 36 (1), 42–51.
Simell, P.A., Hirvensalo, E.K., Smolander, V.T., Krause, A.O.I., 1999. Steam reforming of gasification gas tar over dolomite with benzene as a model compound. Ind. Eng. Chem. Res. 38 (4), 1250–1257.
Srivastava, D.K., Bernhard, S.A., 1986. Metabolite transfer via enzyme-enzyme complexes. Science 234 (4780), 1081–1086.
Stevens, D.J., 2001. Hot gas conditioning: recent progress with larger-scale biomass gasification systems. NREL Subcontractor Report (NREL/SR-510-29952).
Sutton, D., Kelleher, B., Ross, J.R., 2001. Review of literature on catalysts for biomass gasification. Fuel Process. Technol. 73 (3), 155–173.

Thauer, R.K., Jungermann, K., Decker, K., 1977. Energy conservation in chemotrophic anaerobic bacteria. Bacteriol. Rev. 41 (1), 100.

Thomas, G., Parks, G., 2006. Potential roles of ammonia in a hydrogen economy: a study of issues related to the use ammonia for on-board vehicular hydrogen storage. US Department of Energy, 23.

Vajo, J.J., Skeith, S.L., Mertens, F., 2005. Reversible storage of hydrogen in destabilized LiBH4. J. Phys. Chem. B 109 (9), 3719–3722.

Van Vucht, J., Kuijpers, F., Bruning, H., 1970. Reversible room-temperature absorption of large quantities of hydrogen by intermetallic compounds. Philips Res. Rep. 25, 133–140.

Wang, Y.R., Zhang, Y.-H.P., 2009. Overexpression and simple purification of the *Thermotoga maritima* 6-phosphogluconate dehydrogenase in *Escherichia coli* and its application for NADPH regeneration. Microb. Cell Fact. 8, 30.

Wang, Y.R., Zhang, Y.-H.P., 2010. A highly active phosphoglucomutase from *Clostridium thermocellum*: cloning, purification, characterization and enhanced thermostability. J. Appl. Microbiol. 108 (1), 39–46.

Wang, Y.R., Huang, W.D., Sathitsuksanoh, N., Zhu, Z.G., Zhang, Y.-H.P., 2011. Biohydrogenation from biomass sugar mediated by in vitro synthetic enzymatic pathways. Chem. Biol. 18 (3), 372–380.

Weckbecker, A., Groger, H., Hummel, W., 2010. Regeneration of nicotinamide coenzymes: principles and applications for the synthesis of chiral compounds. Biosyst. Eng. I: Creating Superior Biocatal. 120, 195–242.

Wilner, O.I., Weizmann, Y., Gill, R., Lioubashevski, O., Freeman, R., Willner, I., 2009. Enzyme cascades activated on topologically programmed DNA scaffolds. Nat. Nanotechnol. 4 (4), 249–254.

Wolfenden, R., Snider, M.J., 2001. The depth of chemical time and the power of enzymes as catalysts. Acc. Chem. Res. 34 (12), 938–945.

Woodward, J., Orr, M., Cordray, K., Greenbaum, E., 2000. Enzymatic production of biohydrogen. Nature 405, 1014–1015.

Ye, X.H., Wang, Y.R., Hopkins, R.C., Adams, M.W.W., Evans, B.R., Mielenz, J.R., Zhang, Y.-H.P., 2009. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. Chemsuschem 2 (2), 149–152.

Ye, X.H., Rollin, J., Zhang, Y.-H.P., 2010. Thermophilic alpha-glucan phosphorylase from *Clostridium thermocellum*: cloning, characterization and enhanced thermostability. J. Mol. Catal. B Enzymatic 65 (1–4), 110–116.

Ye, X., Zhang, C., Zhang, Y.-H.P., 2012. Engineering a large protein by combined rational and random approaches: stabilizing the *Clostridium thermocellum* cellobiose phosphorylase. Mol. BioSyst. 8, 1815–1823.

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 (11), 6762–6768.

You, C., Myung, S., Zhang, Y.-H.P., 2012. Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. Angew. Chem. Int. Ed. 51 (35), 8787–8790.

Yu, X., Pickup, P.G., 2008. Recent advances in direct formic acid fuel cells (DFAFC). J. Power Sources 182 (1), 124–132.

Zanzi, R., Sjöström, K., Björnbom, E., 2002. Rapid pyrolysis of agricultural residues at high temperature. Biomass Bioenergy 23 (5), 357–366.

Zhang, Y.-H.P., 2009a. A sweet out-of-the-box solution to the hydrogen economy: is the sugar-powered car science fiction? Energy Environ. Sci. 2 (3), 272–282.

Zhang, Y.-H.P., 2009b. Using extremophile enzymes to generate hydrogen for electricity. Microbe 4, 560–565.
Zhang, Y.-H.P., 2010a. Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. Biotechnol. Bioeng. 105 (4), 663–677.
Zhang, Y.-H.P., 2010b. Renewable carbohydrates are a potential high-density hydrogen carrier. Int. J. Hydrog. Energy 35 (19), 10334–10342.
Zhang, Y.-H.P., 2011a. Hydrogen production from carbohydrates: a mini-review, sustainable production of fuels, chemicals, and fibers from forest biomass. ACS Symp. Series 1067, 203–216.
Zhang, Y.-H.P., 2011b. Simpler is better: high-yield and potential low-cost biofuels production through cell-free synthetic pathway biotransformation (SyPaB). ACS Catal. 1 (9), 998–1009.
Zhang, Y.-H.P., 2011c. Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol. Adv. 29 (6), 715–725.
Zhang, Y.-H.P., 2013. Next generation biorefineries will solve the food, biofuels, and environmental trilemma in the energy-food-water nexus. Energy Sci. Eng. 1, 27–41.
Zhang, Y.-H.P., Mielenz, J.R., 2011. Renewable hydrogen carrier-carbohydrate: constructing the carbon-neutral carbohydrate economy. Energies 4 (2), 254–275.
Zhang, Y.-H.P., Evans, B.R., Mielenz, J.R., Hopkins, R.C., Adams, M.W.W., 2007. High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. Plos One 2 (5), e456.
Zhang, Y.-H.P., Sun, J.B., Zhong, J.J., 2010. Biofuel production by in vitro synthetic enzymatic pathway biotransformation. Curr. Opin. Biotechnol. 21 (5), 663–669.
Zhang, Y.-H.P., Xu, J.H., Zhong, J.J., 2012. A new high-energy density hydrogen carrier—carbohydrate—might be better than methanol. Int. J. Energy Res. 37 (7), 769–779.
Zhu, Z.G., Tam, T.K., Sun, F.F., You, C., Zhang, Y.-H.P., 2014. A high-energy-density sugar biobattery based on a synthetic enzymatic pathway. Nat. Commun. 5, 3026.
Züttel, A., 2003. Materials for hydrogen storage. Mater. Today 6 (9), 24–33.

Chapter 3. Complete enzymatic phosphorylation of starch with application to green hydrogen gas production at theoretical yield

Kim, J.-E., Kim, E.J., Chen, H., Huang, R., Adams, M.W.W and Zhang, Y.H.P. (2017). Complete enzymatic phosphorylation of starch for green hydrogen gas production at theoretical yield by in vitro metabolic engineering. (*Submitted*)

Complete enzymatic phosphorylation of starch with application to green hydrogen gas production at theoretical yield

Run title: New solar fuel – starch

Jae-Eung Kim,[1] Eui-Jin Kim,[1] Hui Chen,[1] Chang-Hao Wu,[2] Michael W.W. Adams,[2] Y.-H. Percival Zhang[1,3 *]

1 Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall, Blacksburg, Virginia 24061, USA

2 Department of Biochemistry and Molecular Biology University of Georgia, Athens, Georgia, 30602, USA

3 Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Avenue, Tianjin Airport Economic Area, Tianjin 300308, China

*Corresponding authors: YPZ, Email: ypzhang@vt.edu; zhang_yh@tib.cas.cn

Tel: (540) 231-7414 [O]; Fax: (540) 231-3199

Abstract

Starch is a natural energy storage compound and is hypothesized to be a high-energy density chemical compound or solar fuel. In contrast to industrial hydrolysis of starch to glucose, an alternative ATP-free phosphorylation of starch was designed to generate cost-effective glucose 6-phosphate by using five thermophilic enzymes (i.e., isoamylase, alpha-glucan phosphorylase, 4-α-glucanotransferase, phosphoglucomutase, and polyphosphate glucokinase). This enzymatic phosphorolysis is energetically advantageous because the energy of α-1,4-glycosidic bonds among anhydroglucose units is conserved in the form of phosphorylated glucose. Furthermore, we demonstrated an *in vitro* 17-thermophilic enzyme pathway that can convert all glucose units of starch, regardless of the branched and linear contents, to hydrogen at a theoretic yield (i.e., 12 H_2 per glucose), three times of the theoretical yield from dark microbial fermentation. The use of a biomimetic electron transport chain enabled to achieve a maximum volumetric productivity of 90.2 mmol of H_2/L/h at 20 g/L starch. The complete conversion of starch to hydrogen with this *in vitro* synthetic (enzymatic) biosystem suggests that starch as the best solar fuel becomes a high-density hydrogen storage compound with a gravimetric density of up to 14.8% H_2-based mass and an electricity density of more than 3,000 Wh/kg of starch.

Keywords: *in vitro* synthetic biology, *in vitro* metabolic engineering, starch phosphorylation, hydrogen production, hydrogen storage, thermophilic enzymes

Introduction

Starch is one of the most abundant and inexpensive renewable resources produced by plant photosynthesis. Starch harvested from seeds and tubers of cultivated crops is far more evenly distributed than fossil fuels. Besides most starch-containing cereals consumed as food and feed, approximately 70 million tonnes of starch from harvested grains are used for enzymatic hydrolysis and industrial fermentation for manufacturing numerous products, such as fructose syrup, ethanol, amino acids, and organic acids.[1] Starch is composed of two distinct glucan polymers: amylose, a linear chain of glucose units linked through α-1,4-glycosidic bond; and amylopectin, a highly branched glucose polymer including α-1,6-glycosidic linkage at branch points (**Fig. 1A**). Industrial enzymatic hydrolysis of starch to glucose is a two-step high temperature process with a yield of approximately 98% of the theoretical yield of glucose. During the first step of starch gelatinization and liquefaction α-1,4-glycosidic bonds are partially hydrolyzed by α-amylase at around 90-105 °C, converting the high-viscosity starch slurry to maltodextrin. In the second saccharification step, the remaining α-1,4-glycosidic bonds of maltodextrin are hydrolyzed by glucoamylase and β-amylase, and the branched points at α-1,6-glycosidic bonds are de-branched by pullulanase at ~55-65 °C.[2, 3] During this enzymatic hydrolysis, the glycosidic bond energy among anhydroglucose units is dissipated as waste heat energy.

Distributed production of cost-competitive green hydrogen gas from renewable energy sources is vital to the hydrogen economy. Hydrogen is a promising future energy carrier to replace fossil fuel-derived liquid fuels that are primarily used for transportation, and it offers high energy conversion efficiencies via fuel cells, produces nearly zero pollutants for end users, and has high

mass-specific energy density compared to rechargeable batteries. More than 50 million metric tons of hydrogen are annually produced all around the world through reforming and gasification processes of fossil fuels, and most hydrogen is used for ammonia synthesis and petroleum refining processes.[4] Nearly 96% of hydrogen are produced from fossil fuels.[5] Thermochemical processes of hydrogen production also require high capital investment for large-size central plants and cannot be scaled down economically for distributed operations.[6] Water splitting by solar energy to produce hydrogen is a promising clean and sustainable process. However, direct solar-powered water splitting suffers from very slow hydrogen generation rates due to low insolation flux (e.g., ~200 W/m^2) and photocatalysts with a limited range of light absorption and inefficient charge separation.[7] In addition, intermittent sunlight at a low energy concentration causes a major challenge to the scale-up of direct solar-powered hydrogen production. Indirect water splitting powered with solar electricity has rapid volumetric productivity of hydrogen (i.e., 40 mole of H_2/L/h), but its practical use is limited due to high costs of electricity.[8] For hydrogen to be a competitive energy carrier used for transportation, new technology must be developed to meet increasing needs of distributed hydrogen production in terms of both volumetric production rate and energy efficiency, while keeping the production costs low and reducing greenhouse gas emissions.

In vitro synthetic biology is an emergent biomanufacturing platform that does not have cellular constraints, such as cell membranes, complicated cellular regulation, and complex bioenergetics.[9-13] Some limitations of the whole-cell fermentation, e.g., a loss of energy and carbon source due to cellular replication and side-product formation, can be relieved by reconstituting purified enzymes into *in vitro* synthetic (enzymatic) biosystems. In addition, the *in*

vitro synthetic biosystem features several biomanufacturing advantages over whole-cell factories. These include fast volumetric productivities or biomanufacturing rates,[14] high product yields and energy efficiencies,[15] easy product separation,[14] tolerance to toxic substrates or products,[16] broad reaction conditions,[17] implementation of non-natural reaction,[18] and great engineering flexibility in terms of using enzymes as highly exchangeable biobricks.[19]

We have demonstrated *in vitro* hydrogen production from glycogen (animal starch) as a proof-of-concept study.[20] This system used mostly mesophilic enzymes, but exceeded the theoretical limit of microbial hydrogen production yield, that is, four H_2 per glucose. However, the maximum hydrogen production rate was only 0.4 mmol/L/h with 43% of the theoretical yield of 12 H_2 per glucose. These results were due to the highly-branched structure of glycogen, where glycogen phosphorylase can release approximately 30% of glucose units in glycogen.[21] We subsequently designed and demonstrated several *in vitro* enzymatic pathways for hydrogen production from various carbohydrates, including cellulosic materials,[22] xylose,[23] sucrose,[24] biomass mono-sugars,[25] and xylooligosaccharides.[26] However, all of these pathways suffer from the consumption of a significant amount of polyphosphate for the activation of monosaccharide sugars into sugar phosphates, resulting in potentially high operation costs from phosphate recycling and polyphosphate regeneration.

In the present study, an *in vitro* ATP-free enzymatic pathway for the complete phosphorylation of starch to glucose 6-phosphate (G6P) was designed with five thermophilic enzymes including isoamylase (IA), α-glucan phosphorylase (αGP), 4-α-glucanotransferase (4GT), polyphosphate glucokinase (PPGK), and phosphoglucomutase (PGM) (**Fig. 1B**). In combined with the complete

phosphorylation pathway of starch, an *in vitro* synthetic pathway comprised of numerous thermophilic enzymes, $NADP^+$, and benzyl viologen as an abiotic electron mediator instead of ferredoxin, an oxygen-sensitive protein, was designed to demonstrate high-speed hydrogen production from starch and water at a theoretical yield for the first time.

Results

In vitro Synthetic Pathway Design

An *in vitro* synthetic enzymatic pathway was designed for ATP-free phosphorylation of starch to G6P (**Fig. 1B**). In contrast to intracellular G6P generation via glycolysis which requires ATP, the *in vitro* enzymatic pathway comprised of five enzymes IA, αGP, 4GT, PGM and PPGK can accomplish ATP-free phosphorylation for releasing G6P from starch with phosphate ions. The α-1,6-glycosidic linkages of starch were debranched by a hyperthermophilic enzyme IA, yielding linear amylodextrin.[27] IA-treated starch (called amylodextrin later) and phosphate catalyzed by αGP were generated glucose 1-phosphate (G1P) efficiently. The chain length of amylodextrin continuously decreases for releasing G1P until amylodextrin becomes maltose or maltotriose, which was not utilized by αGP.[28] More G1P can be generated from maltose and maltotriose with a help of 4GT, which re-arranges one anhydroglucose unit from maltose to either another maltose or maltotriose to yield a short-chain amylodextrin with one glucose molecule released (**Fig. 1B**). Newly synthesized short-chain amylodextrin can be utilized by αGP to produce more G1P. Here the use of thermophilic 4GT is better than the use of maltose phosphorylase because there are no hyperthermophilic maltose phosphorylases reported. All G1P via starch phosphorolysis are converted to G6P catalyzed by PGM. One glucose residual per amylodextrin is converted to G6P catalyzed by PPGK at a cost of polyphosphate (i.e., one phosphate group

released from polyphosphate is used to activate glucose). In contrast to the classical glucose-producing hydrolysis of starch, this *in vitro* enzymatic starch phosphorylation pathway utilizes phosphate and α-1,4-glycosidic bond energy from starch to activate glucose units, yielding G6P. This enzymatic pathway enables cost-effective production of G6P by avoiding the use of costly ATP and utilizing all glucose units of starch for the first time.

In combination with the phosphorylation pathway, the *in vitro* enzymatic pathway comprised of 17 purified enzymes was designed to achieve the complete utilization of starch for hydrogen production (**Fig. 2**). This reconstituted ATP-free and cofactor-balanced enzymatic pathway can be grouped into four modules: (1) ATP-free phosphorylation of starch generating G6P (**Fig. 1B**); (2) NADPH generation via the oxidative pentose phosphate pathway (PPP); (3) hydrogen generation catalyzed by soluble [NiFe] hydrogenase from a hyperthermophilic archaeon *Pyrococcus furiosus* (SHI) from NADPH via a biomimetic electron transport chain (ETC) comprised of NADPH rubredoxin oxidoreductase (NROR) and benzyl viologen (BV) as an abiotic electron mediator;[29] and (4) G6P regeneration via the non-oxidative PPP and partial gluconeogenesis pathway. Phosphate generated in the fourth module is re-cycled by αGP for starch phosphorylation in the first module. The overall stoichiometric reaction can be written approximately as $C_6H_{10}O_5 + 7H_2O = 12H_2 + 6CO_2$.

Hydrogen Production at Theoretical Yield

All recombinant thermophilic enzymes except SHI were overexpressed in *E. coli* and purified to homogeneity for the enzyme cocktail reconstitution (**Table S1**). SHI was over-expressed in the native host, *P. furiosus*, whose optimal growth temperature is more than 100 ^{o}C.[25] The hydrogen

production experiments were conducted on maltodextrin, a partially hydrolyzed starch with a dextrose equivalent (DE) of 4.0-7.0 (**Fig. 3**). When maltodextrin was used as a substrate without the biomimetic ETC, the maximum volumetric hydrogen productivity was 5.1 mmol H_2/L/h at hour 4. This maximum hydrogen volumetric productivity was approximately 13-fold higher than that on glycogen.[20] Such an improvement on volumetric productivity was mainly due to an increased reaction temperature from 37 to 50 ^{o}C with the use of thermophilic enzymes instead of mesophilic enzymes, and the optimized enzyme ratio with help of kinetic modeling.[25] The integrated yield was 62% of the theoretical yield (i.e., 12 moles hydrogen per mole of glucose) after 30 h of the reaction.

The introduction of the biomimetic ETC, which transferred electrons from NADPH to oxidized BV catalyzed by NROR and from reduced BV to H_2 catalyzed by SHI, led to another 6.4-fold enhancement in the maximum volumetric productivity of 32.7 mmol of H_2/L/h after 2 h (**Fig. 3**). The hydrogen peak appearing earlier in this case showed an efficient removal of NADPH to H_2 generation than that without the biomimetic ETC. As a result, the integrated hydrogen yield was improved slightly to 65% after 8 h of the reaction.

The number-average degree of polymerization (DP) of maltodextrin (DE 4.0-7.0) was 23.8 determined by the phenol-sulfuric acid method and the modified bicinchoninic acid method.[30] After debranching with excess IA, the DP value was decreased to 19.3, suggesting that 23% of maltodextrin contains α-1,6-glycosidic branching points. These branching points of maltodextrin prevented αGP from progressively phosphorylating several glucose units next to the branching points. After IA-treatment yielding linear amylodextrin (**Fig. 1B**), the hydrogen yield was

increased to 77% and a slight improvement in the maximum volumetric productivity to 36.0 mmol/L/h was observed (**Fig. 3**). The addition of 4GT after 12 hours of the reaction enabled the generation of more G1P from maltotriose and maltose, resulting in an increase of the hydrogen yield from 77% to 95%. This enhanced yield was in good agreement with that value obtained with amylodextrin with the DP of 19.3, which has a potential of G1P generation of 94.8% (i.e., $\frac{DP-1}{DP} \times 100\%$. To utilize residual glucose molecules released by 4GT, PPGK along with a small amount of polyphosphate was added after 24 h. After 30 h, the integrated hydrogen yield was increased to 99% (**Fig. 3**). There was no detectable glucose, G1P, and G6P at the end of the reaction. For amylodextrin with the DP of 19.3, G1P via *in vitro* starch phosphorylation pathway accounts for 95% of the overall hydrogen yield while glucose via polyphosphate-based PPGK accounts for approximately 5% (i.e., $\frac{1}{DP} \times 100\%$).

The introduction of the biomimetic ETC was one of the most important steps to greatly increase volumetric productivity.[29] BV is also a perfect indicator of the *in vitro* hydrogen production and redox status of the reactants (**Fig. S1**). When oxidized BV (colorless) was initially added, the reaction solution was almost colorless. When the oxygen inside the reactor was completely evacuated by the nitrogen carrier gas, NADPH was generated from the PPP and colorless oxidized BV was converted to blue reduced BV. The development of the blue color was accompanied by the initiation of hydrogen production (**Fig. S1**). The reaction solution became deeper blue as more NADPH was generated that in turn generates more reduced BV. Reducing power decreased as the substrate was consumed, resulting in decreased volumetric productivity accompanied with a decrease in the blue color. When the substrate was depleted, no more hydrogen was produced leaving almost all of the BV in the colorless oxidized form (**Fig. S1**).

High-Speed Hydrogen Production

For a potential hydrogen generation in distributed (less costly) stationary bioreactors, the starch concentration in terms of glucose equivalent was increased 5-fold in a step-wise fashion until the maximum hydrogen volumetric productivity was obtained. An additional enzyme, 6-phosphogluconolactonase (6PGL) catalyzing the reaction between 6-phosphogluconolactone and 6-phosphogluconate,[31] was added for a high-speed hydrogen production from starch. The addition of 6PGL enabled the maximum hydrogen productivity to increase from 32.7 mmol/L/h to 42.6 mmol/L/h by 30% on 5 mM starch (in terms of glucose equivalents) (**Fig. 4**). When the starch concentration was increased to 25 mM and 125 mM (i.e., 20.3 g/L starch), the maximum hydrogen productivities increased to 69.2 mmol/L/h and 90.2 mmol/L/h, respectively. The productivity did not further increase upon increasing the starch concentration to 625 mM. The highest volumetric productivity of 90.2 mmol of H_2/L/h (i.e., 1.93 L of H_2/L/h) shows a technical feasibility for industrial stationary hydrogen generation systems, such as anaerobic digesters or beer fermenters. Also, the rapid formation of hydrogen gas bubbles was observed on the surface of the aqueous solution (**Fig. S2**).

Discussion

In contrast to industrial hydrolysis of starch to glucose with α-amylase, pullulanase, and glucoamylase,[3] the cost-effective starch phosphorolysis pathway was designed to achieve the complete utilization of starch by an enzyme cocktail of IA, αGP, 4GT, PGM, and PPGK (**Fig. 1B**). The phosphorolytic cleavage of starch catalyzed by αGP is energetically advantageous because the energy of α-1,4-glycosidic bonds among anhydroglucose units is conserved in the

form of phosphorylated glucose. When coupled with other enzymatic pathways that consume G6P, phosphate ions can be reutilized by αGP. Phosphate recycling in one pot helps keep nearly constant pH and phosphate concentration. In this study, we achieved a starch utilization efficiency of 95% without the use of polyphosphate (**Fig. 3**). Although a small amount of polyphosphate is required for the complete utilization of starch (i.e., one phosphate of polyphosphate per linear amylodextrin with the DP of ~20), it is anticipated that the use of polyphosphate could be decreased to lower levels when long DP starch materials, such as low DE maltodextrin, IA-treated intact natural starch, or long DP amylodextrin are used (**Fig. 1B**). This ATP-free generation of G6P from starch could open the door to the cost-effective production of various products, such as hydrogen (as described herein), fructose as a third generation sweetener,[32] tagatose as a fourth generation healthy sweetener,[33] bioelectricity,[34] and so on.

It was hypothesized that starch was a promising hydrogen storage carrier in 2007,[20] but we now validate this hypothesis by this new *in vitro* synthetic pathway (**Fig. 2**). This study using starch has many improvements over the prior work on glycogen.[20] These include: (1) the use of less costly starch instead of costly glycogen (animal starch with more branching points); (2) the complete utilization of the glucose units of starch (**Fig. 1B**); (3) a more than 200-fold enhancement in volumetric productivity of hydrogen; (4) the replacement of mesophilic enzymes with a set of thermophilic enzymes; and (5) increased reaction temperature from 37 to 50 °C, thereby lowering the chance of microbial contamination. According to the stoichiometric reaction $C_6H_{10}O_5 + 7H_2O = 12H_2 + 6CO_2$, dry starch powder has a mass H_2 storage density of 14.8 H_2-mass% if water is recycled, while the starch-water slurry has a density of 8.33 H_2-mass%

if water is not recycled. Both dehydrated starch and the starch slurry have high volumetric hydrogen storage densities of over 100 kg H_2/m^3, much higher than the Department of Energy's hydrogen storage goals. The theoretical yield achieved for the hydrogen production from starch with the *in vitro* synthetic biosystem could be a promising solution not only for cost-competitive green hydrogen generation at decentralized stationary bioreactors, such as biogas production in anaerobic digesters, but also for an off-board hydrogen storage solution.[35] Furthermore, starch becomes an electricity storage carrier with a density of more than 3000 Wh/kg of starch[35, 36] because of this ultra-low cost complete phosphorylation of starch (**Fig. 1B**).

Over the past decade we have validated the technical feasibility of high-yield hydrogen productions from glycogen,[20] cellodextrins,[22] xylose,[23] sucrose,[24] a mixture of biomass monosaccharides,[25] and xylo-oligosaccharides.[26] Hydrogen production rates and yields from different low-cost carbohydrates are summarized in **Table S2**. *In vitro* synthetic pathways surpass dark microbial fermentation because the latter suffers from the low theoretical yield, known as the Thauer limit, four moles of hydrogen per mole of glucose (33% energy efficiency compared to this *in vitro* enzymatic pathway).[8, 37] Therefore, neither natural nor metabolically engineered hydrogen-producing microorganism can offer cost-competitive hydrogen production because carbohydrates account for a major fraction of biocommodity prices.[9, 38] Upon the complete utilization of starch with its current price (e.g., approximately 0.30 US dollars per kg and lower in the future), it may be the most promising substrate for the hydrogen economy. In addition, the volumetric productivity of hydrogen with an *in vitro* enzymatic pathway has been increased from 0.4 to 90.2 mmol/L/h by more than 200-fold through the replacement of mesophilic enzymes with thermophilic enzymes (**Table S3**), optimization of enzyme ratios,[25]

and introduction of a biomimetic electron transport chain consisting of NROR and BV as an abiotic electron mediator instead of an oxygen-sensitive ferredoxin.[29] The highest productivity achieved here (i.e., 90.2 mmoles of H_2/L/h) is comparable to those of industrial biogas and hydrogen production.[39] It is expected that enzymatic hydrogen production rates can be further increased by an order of magnitude or more through increasing the reaction temperature to 80 °C (or even higher), increasing enzyme loading, using small-size biomimetic enzymes for better mass transfer,[40] creating substrate channeling in metabolons,[41] and using more highly active enzymes and integrated multi-enzyme modules.[9, 42]

Over the past decade the use of starch for bioenergy production has often been subject to the "food versus fuel" debate. However, recent statistics that the total global volume of biofuels continuously increases while crops prices decrease from 2011 to present showed a contradiction to the traditional belief.[43] A modeling approach to understand the complexity of the system dynamics of land, food and bioenergy also showed that bioenergy represents a relevant solution to a large number of possible scenarios, such as decreasing greenhouse gas emissions and increasing both global energy and food securities.[44] In comparison to ethanol production from starch, the production of hydrogen can offer more than three times higher biomass-to-wheel efficiency.[45] More importantly, two new methods could make sufficient artificial starch suitable for hydrogen production and storage via enzymatic biotransformation of nonfood biomass[18] and artificial photosynthesis[35, 36]. Therefore, the paradigm shift from "food versus fuel" to "food and fuel" may be promoted for the development of a sustainable low-carbon bioeconomy.[44]

This study suggests that starch could be one of the best solar fuels as compared to numerous potential solar fuels, such as methanol, hydrogen, CO, butanol [46-48] by considering its high-hydrogen storage density and high electricity-storage carrier.[35, 36] In addition to high energy storage density, starch has an unique feature: adjustable water solubility depending on its chain lengths, control of its branching points, and concentration. Long chain amylodextrin which has very low water solubility can be easily separated from aqueous solution to address the problems of separation of aqueous solar fuels (e.g., butanol, methanol) from the aqueous solutions and harvesting of gaseous solar fuels (e.g., H_2, methane, CO).

Remaining challenges for the commercialization of *in vitro* enzymatic hydrogen production include replacing the expensive and labile coenzyme -- $NADP^+$, and further improving enzyme thermal stability, such as that of PPGK.[49] The cost and stability of natural coenzymes could be addressed by replacing them with more stable and less costly (biomimetic) coenzymes[9] while more thermostable enzymes might be discovered and/or obtained by enzyme engineering and immobilization. It is anticipated that the commercial success of the *in vitro* synthetic biology platform will take place within next few years through more collaboration with industrial partners in this emerging area.

Materials and Methods

Characterization and Preparation of Starch

Partially hydrolyzed starch, maltodextrin (DE 4.0–7.0), was used as an intact substrate. DE was defined as the number of reducing end equivalents per unit dry weight and was calculated using the following formula: $DE = 180/(162 \times DP + 18) \times 100$.[3] Number average DP can be calculated

as the ratio of glucosyl monomer concentration divided by the reducing-end concentration. The glucosyl monomer concentration and the reducing-end concentration were determined by the phenol-sulfuric acid method and the modified bicinchoninic acid method, respectively.[30] Starch or maltodextrin was debranched by IA in 5 mM acetate buffer (pH 5.5) containing 0.5 mM $MgCl_2$ at a mass ratio of starch/IA = 1000:1 at 80 °C for 24 h.

Preparation of Recombinant Enzymes

All thermophilic enzymes except SHI were produced by heterologous expression in *E. coli* BL21(DE3) containing the respective expression plasmids in LB media. H_2ase was produced by homologous overexpression in the native host *P. furiosus.*[25] Details on the enzymes and purification methods are given in **Table S1** and ***supplementary materials and methods***.

Enzyme Cocktail for *in vitro* Hydrogen Production

The enzyme loadings for hydrogen production from starch are listed in **Table S2**. All enzymes were stored in 50% (wt/wt) glycerol at -20 °C, except for PPGK which was freshly prepared and immobilized in RAC. Appropriate volumes of each enzyme, except IA, 4GT, PPGK and NROR, were combined and diluted to 0.1% glycerol with the addition of 20 mM HEPES buffer (pH 7.5), and then re-concentrated with 10,000 MWCO Amicon centrifugal filters from Millipore. The final reaction mixture with concentrated enzymes was adjusted to 100 mM HEPES buffer (pH 7.5) containing 5 mM intact or IA-treated maltodextrin, 5 mM sodium phosphate, 5 mM $NADP^+$, 10 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 0.5 mM thiamine pyrophosphate. For high-rate hydrogen production, the appropriate volume of NROR and 2 mM benzyl viologen were added as supplements. For the complete starch utilization to hydrogen production, 4GT and PPGK were

added sequentially and 5 mM polyphosphate was added with PPGK. For protection against microbial growth, 25 μg/mL of kanamycin and 0.01% (w/v) sodium azide were added to all reactions. Maltodextrin concentration was increased stepwise when the maximum hydrogen rate was reached (i.e., 25, 125 and 625 mM) and the sodium phosphate concentration was also increased (i.e., 25 and 50 mM) according to maltodextrin concentration. The reactor was sealed and the reaction solution was agitated with a magnetic stir bar. Temperature, carrier gas flow rate and hydrogen signal were monitored continuously throughout all experiments. Hydrogen production was measured in a continuous flow system purged with 50 mL/min ultrapure nitrogen (Airgas, Christiansburg, VA) as described previously.[25]

Acknowledgements

This study is mainly supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy, Fuel Cell Technologies Office under Award Number DE-EE0006968. Funding to YPZ for this work was partially supported by the Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture.

Contributions

YPZ conceived of the project, oversaw, and coordinated research; JEK and YPZ designed experiments; JEK performed experiments; EJK, HC, CHW, and MWA contributed experimental materials; JEK, EJK, and YPZ analyzed data; JEK made figures; and YPZ, JEK and MWA wrote the paper.

References

Foley, J.A. et al. Solutions for a cultivated planet. *Nature* **478**, 337-342 (2011).

Chaplin, M.F. & Bucke, C. Enzyme technology. (Cambridge University Press, Cambridge, England; 1990).

van der Maarel, M.J.E.C., van der Veen, B., Uitdehaag, J.C.M., Leemhuis, H. & Dijkhuizen, L. Properties and applications of starch-converting enzymes of the alpha-amylase family. *J. Biotechnol.* **94**, 137-155 (2002).

Navarro, R., Pena, M. & Fierro, J. Hydrogen production reactions from carbon feedstocks: fossil fuels and biomass. *Chem. Rev.* **107**, 3952-3991 (2007).

Ngoh, S.K. & Njomo, D. An overview of hydrogen gas production from solar energy. *Renew. Sustain. Energy Rev.* **16**, 6782-6792 (2012).

Clomburg, J.M., Crumbley, A.M. & Gonzalez, R. Industrial biomanufacturing: The future of chemical production. *Science* **355** (2017).

Maeda, K. & Domen, K. Photocatalytic water splitting: Recent progress and future challenges. *J. Phys. Chem. Lett.* **1**, 2655-2661 (2010).

Levin, D.B. & Chahine, R. Challenges for renewable hydrogen production from biomass. *Int. J. Hydrogen Energy* **35**, 4962-4969 (2010).

Rollin, J.A., Tam, T.K. & Zhang, Y.-H.P. New biotechnology paradigm: cell-free biosystems for biomanufacturing. *Green Chem* **15**, 1708-1719 (2013).

Hold, C., Billerbeck, S. & Panke, S. Forward design of a complex enzyme cascade reaction. *Nat. Commun.* **7**, 12971 (2016).

Orelle, C. et al. Protein synthesis by ribosomes with tethered subunits. *Nature* **524**, 119-124 (2015).

Pardee, K. et al. Portable, on-demand biomolecular manufacturing. *Cell* **167**, 248-259.e212 (2016).

Opgenorth, P.H., Korman, T.P. & Bowie, J.U. A synthetic biochemistry module for production of bio-based chemicals from glucose. *Nat. Chem. Biol.* **12**, 393–395 (2016).

Hodgman, C.E. & Jewett, M.C. Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng.* **14**, 261-269 (2012).

Opgenorth, P.H., Korman, T.P. & Bowie, J.U. A synthetic biochemistry molecular purge valve module that maintains redox balance. *Nat. Commun.* **5**, 4113 (2014).

Guterl, J.K. et al. Cell-free metabolic engineering: Production of chemicals by minimized reaction cascades. *ChemSusChem* **5**, 2165-2172 (2012).

Panke, S., Held, M. & Wubbolts, M. Trends and innovations in industrial biocatalysis for the production of fine chemicals. *Curr. Opin. Biotechnol.* **15**, 272-279 (2004).

You, C. et al. Enzymatic transformation of nonfood biomass to starch. *Proc. Natl. Acad. Sci. USA* **110**, 7182-7187 (2013).

Zhang, Y.-H.P., Sun, J.B. & Zhong, J.J. Biofuel production by *in vitro* synthetic enzymatic pathway biotransformation. *Curr. Opin. Biotechnol.* **21**, 663-669 (2010).

Zhang, Y.-H.P., Evans, B.R., Mielenz, J.R., Hopkins, R.C. & Adams, M.W.W. High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. *PLoS One* **2**, e456 (2007).

Lew, C.R., Guin, S. & Theodorescu, D. Targeting glycogen metabolism in bladder cancer. *Nat. Rev. Urol.* **12**, 383-391 (2015).

Ye, X.H. et al. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. *ChemSusChem* **2**, 149-152 (2009).
Martin del Campo, J.S. et al. High-yield production of dihydrogen from xylose by using a synthetic enzyme cascade in a cell-free system. *Angew. Chem. Int. Ed.* **52**, 4587-4590 (2013).
Myung, S. et al. *In vitro* metabolic engineering of hydrogen production at theoretical yield from sucrose. *Metab. Eng.* **24**, 70-77 (2014).
Rollin, J.A. et al. High-yield hydrogen production from biomass by *in vitro* metabolic engineering: Mixed sugars coutilization and kinetic modeling. *Proc. Natl. Acad. Sci. USA* **112**, 4964-4969 (2015).
Moustafa, H.M.A. et al. Water splitting for high-yield hydrogen production energized by biomass xylooligosaccharides catalyzed by an enzyme cocktail. *ChemCatChem* **8**, 2898-2902 (2016).
Cheng, K., Zhang, F., Sun, F.F., Chen, H.G. & Zhang, Y.-H.P. Doubling power output of starch biobattery treated by the most thermostable isoamylase from an archaeon *Sulfolobus tokodaii*. *Sci Rep* **5**, 13184 (2015).
Zhou, W., You, C., Ma, H.W., Ma, Y.H. & Zhang, Y.-H.P. One-pot biosynthesis of high-concentration alpha-glucose 1-phosphate from starch by sequential addition of three hyperthermophilic enzymes. *J. Agri. Food Chem.* **64**, 1777-1783 (2016).
Kim, E.J., Wu, C.H., Adams, M.W.W. & Zhang, Y.-H.P. Exceptionally high rates of biological hydrogen production by biomimetic *in vitro* synthetic enzymatic pathways. *Chemistry* **22**, 16047-16051 (2016).
Zhang, Y.-H.P. & Lynd, L.R. Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis. *Biomacromolecules* **6**, 1510-1515 (2005).
Zhu, Z.G. & Zhang, Y.-H.P. *In vitro* metabolic engineering of bioelectricity generation for the complete oxidation of glucose. *Metab. Eng.* **39**, 110-116 (2017).
Moradian, A. & Benner, S.A. A biomimetic biotechnological process for converting starch to fructose - Thermodynamic and evolutionary considerations in applied enzymology. *J. Am. Chem. Soc.* **114**, 6980-6987 (1992).
Oh, D.-k., Hong, S.-H. & Lee, S.-H. (2014).
Zhu, Z.G., Tam, T.K., Sun, F.F., You, C. & Zhang, Y.-H.P. A high-energy-density sugar biobattery based on a synthetic enzymatic pathway. *Nat. Commun.* **5**, 3026 (2014).
Zhang, Y.-H.P. & Huang, W.D. Constructing the electricity-carbohydrate-hydrogen cycle for a sustainability revolution. *Trends Biotechnol.* **30**, 301-306 (2012).
Zhang, Y.-H.P. Next generation biorefineries will solve the food, biofuels, and environmental trilemma in the energy-food-water nexus. *Energy Sci. Eng.* **1**, 27-41 (2013).
Thauer, R.K., Jungermann, K. & Decker, K. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* **41**, 100-180 (1977).
Maeda, T., Sanchez-Torres, V. & Wood, T.K. Hydrogen production by recombinant *Escherichia coli* strains. *Microb. Biotechnol.* **5**, 214-225 (2012).
Kotay, S.M. & Das, D. Biohydrogen as a renewable energy resource - Prospects and potentials. *Int. J. Hydrogen Energy* **33**, 258-263 (2008).
Paul, C.E., Arends, I.W.C.E. & Hollmann, F. Is simpler better? Synthetic nicotinamide cofactor analogues for redox chemistry. *ACS Catal.* **4**, 788-797 (2014).

Wheeldon, I. et al. Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8**, 299-309 (2016).
Brasen, C., Esser, D., Rauch, B. & Siebers, B. Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.* **78**, 89-175 (2014).
Rosillo-Calle, F. A review of biomass energy-shortcomings and concerns. *J. Chem. Technol. Biotechnol.* **91**, 1933-1945 (2016).
Strapasson, A. Modelling the limits of bioenergy. *Systems Res. Behav. Sci.* **33**, 686-690 (2016).
Huang, W.-D. & Zhang, Y.-H.P. Energy efficiency analysis: Biomass-to-wheel efficiency related with biofuels production, fuel distribution, and powertrain systems. *PLoS One* **6**, e22113 (2011).
Woolerton, T.W., Sheard, S., Chaudhary, Y.S. & Armstrong, F.A. Enzymes and bio-inspired electrocatalysts in solar fuel devices. *Energy Environ. Sci.* **5**, 7470-7490 (2012).
Centi, G. & Perathoner, S. Towards solar fuels from water and CO_2. *ChemSusChem* **3**, 195-208 (2010).
Artero, V. & Fontecave, M. Solar fuels generation and molecular systems: is it homogeneous or heterogeneous catalysis? *Chem. Soc. Rev.* **42**, 2338-2356 (2013).
Liao, H.H., Myung, S.W. & Zhang, Y.-H.P. One-step purification and immobilization of thermophilic polyphosphate glucokinase from *Thermobifida fusca* YX: glucose-6-phosphate generation without ATP. *Appl. Microbiol. Biotechnol.* **93**, 1109-1117 (2012).

Figure Legends

Figure 1. Structural diagram for plant starch (A). Each glucose unit is connected by the α-1,4-glycosidic bond and the branches are connected by α-1,6-glycosidic bond. Enzymatic pathway depicting the complete phosphorylation of starch to glucose 6-phosphate without the use of ATP.

Figure 2. Enzymatic pathway for *in vitro* hydrogen production through complete utilization of starch. Details and full names of enzymes are included in **Table S2**.

Figure 3. Hydrogen production profile and yield from starch with an in vitro enzymatic pathway without a ETC (red lines), with the pathway containing a biomimetic ETC (blue lines), and from IA-treated starch (amylodextrin) with the ETC-containing pathway supplemented with 4GT and PPGK (black lines). The solid lines represent the volumetric productivities of hydrogen and the dashed lines represent the integrated hydrogen yields.

Figure 4. The time profile of the maximum volumetric productivity of hydrogen in step-wise addition of more starch at 50^{o}C. Starting with 5 mM starch, the maximum hydrogen productivities were measured with starch concentrations increased by 5-fold step-wise until no further improvement on volumetric productivity at 625 mM starch (data not shown) was observed.

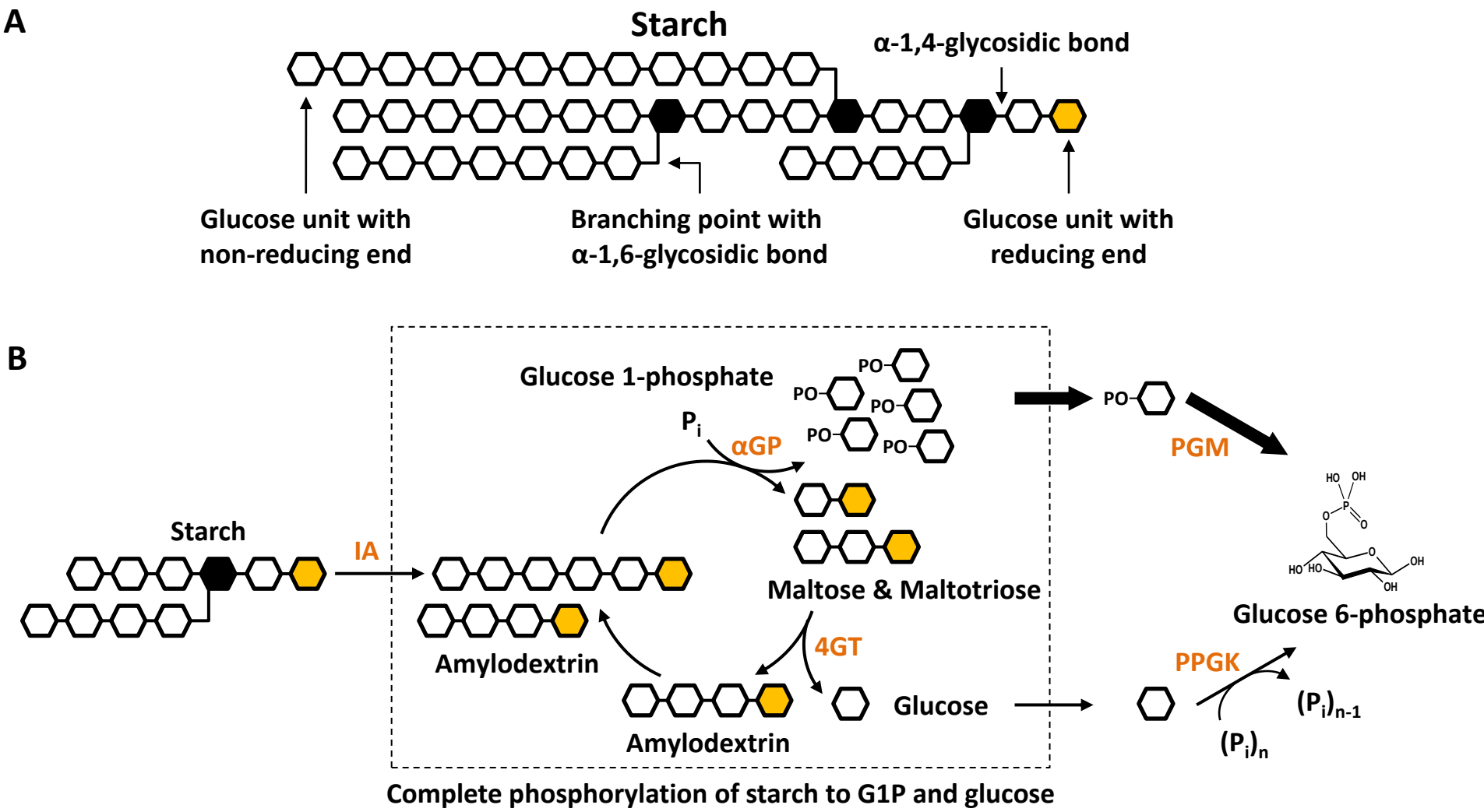
A
Starch
α-1,4-glycosidic bond
Glucose unit with non-reducing end
Branching point with α-1,6-glycosidic bond
Glucose unit with reducing end
B
Glucose 1-phosphate
PO
Pi
αGP
Starch
IA
Maltose & Maltotriose
Amylodextrin
4GT
Glucose
Amylodextrin
PGM
Glucose 6-phosphate
PPGK
(Pi)n-1
(Pi)n
Complete phosphorylation of starch to G1P and glucose

Figure 1.

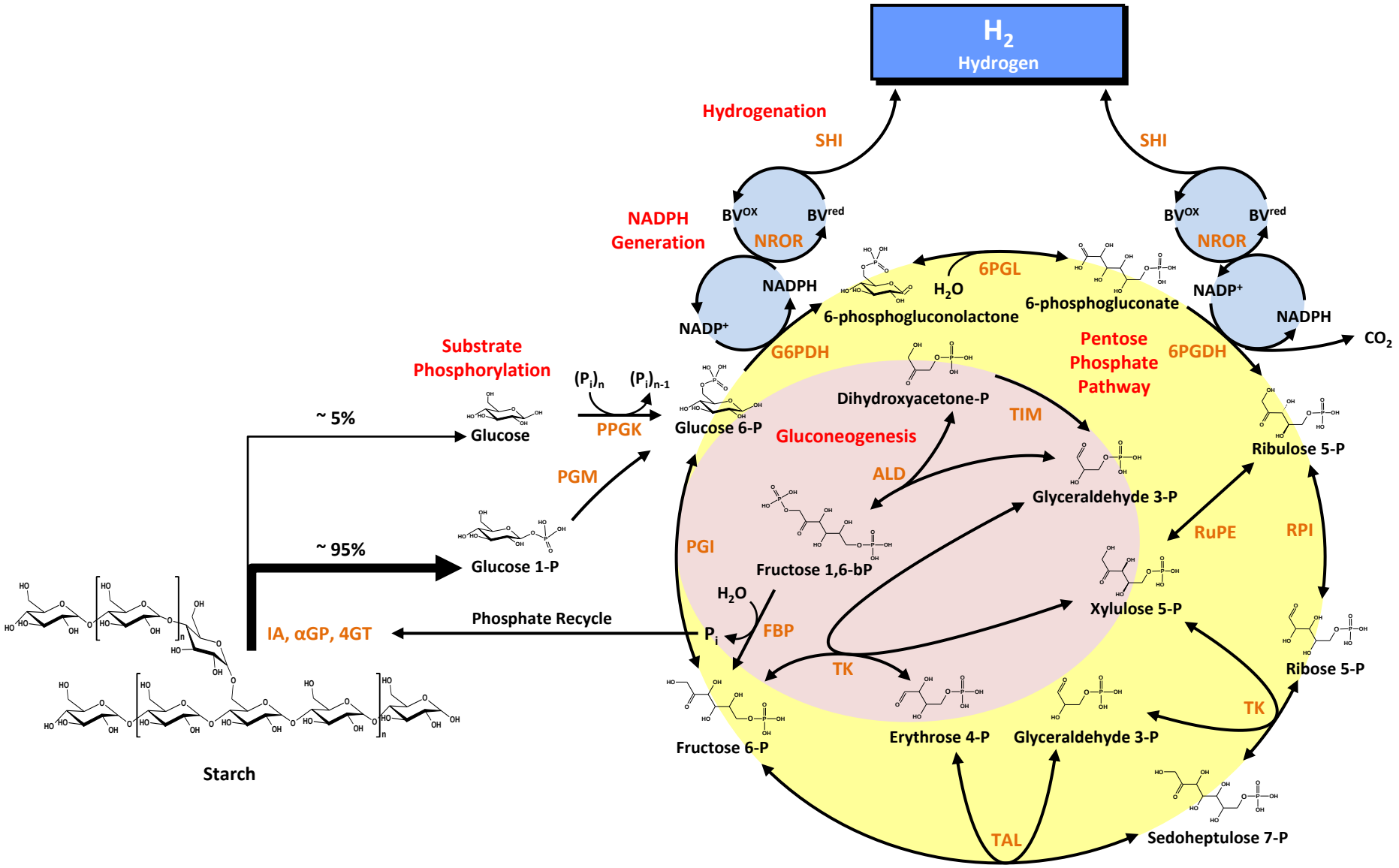
H2
Hydrogen
Hydrogenation
SHI
SHI
NADPH
Generation
BVOX
BVred
NROR
BVOX
BVred
NROR
NADPH
NADP+
NADP+
NADPH
6PGL
H2O
6-phosphogluconolactone
6-phosphogluconate
CO2
Substrate
Phosphorylation
(Pi)n
(Pi)n-1
PPGK
G6PDH
Pentose
Phosphate
Pathway
6PGDH
~ 5%
Glucose
Glucose 6-P
Dihydroxyacetone-P
TIM
Gluconeogenesis
Ribulose 5-P
PGM
ALD
Glyceraldehyde 3-P
~ 95%
Glucose 1-P
PGI
RuPE
RPI
Fructose 1,6-bP
H2O
Xylulose 5-P
Phosphate Recycle
IA, αGP, 4GT
Pi
FBP
Ribose 5-P
TK
TK
Fructose 6-P
Erythrose 4-P
Glyceraldehyde 3-P
TAL
Sedoheptulose 7-P
Starch

Figure 2.

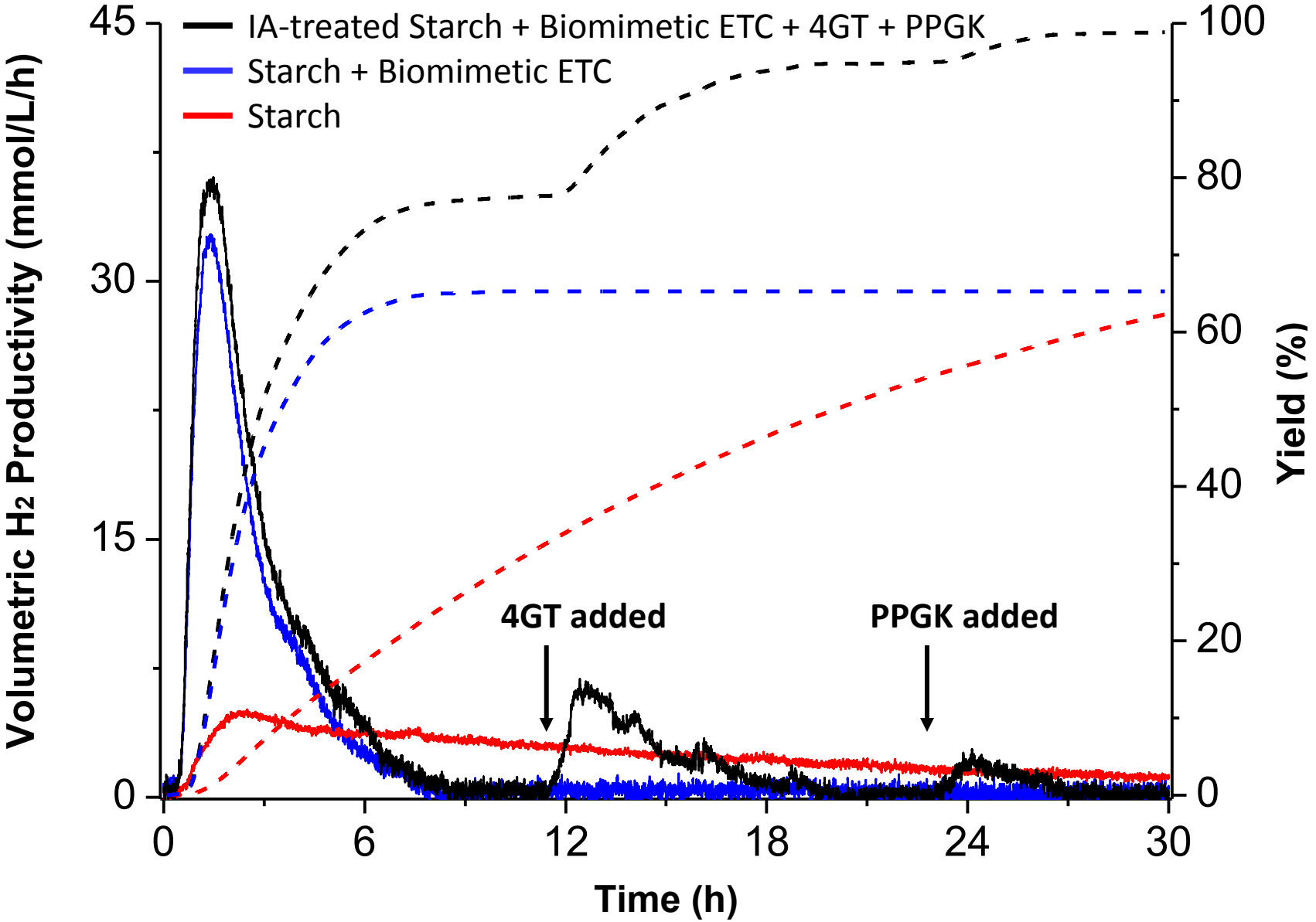
IA-treated Starch + Biomimetic ETC + 4GT + PPGK
Starch + Biomimetic ETC
Starch
Volumetric H2 Productivity (mmol/L/h)
Yield (%)
Time (h)
4GT added
PPGK added
0
15
30
45
0
20
40
60
80
100
0
6
12
18
24
30

Figure 3.

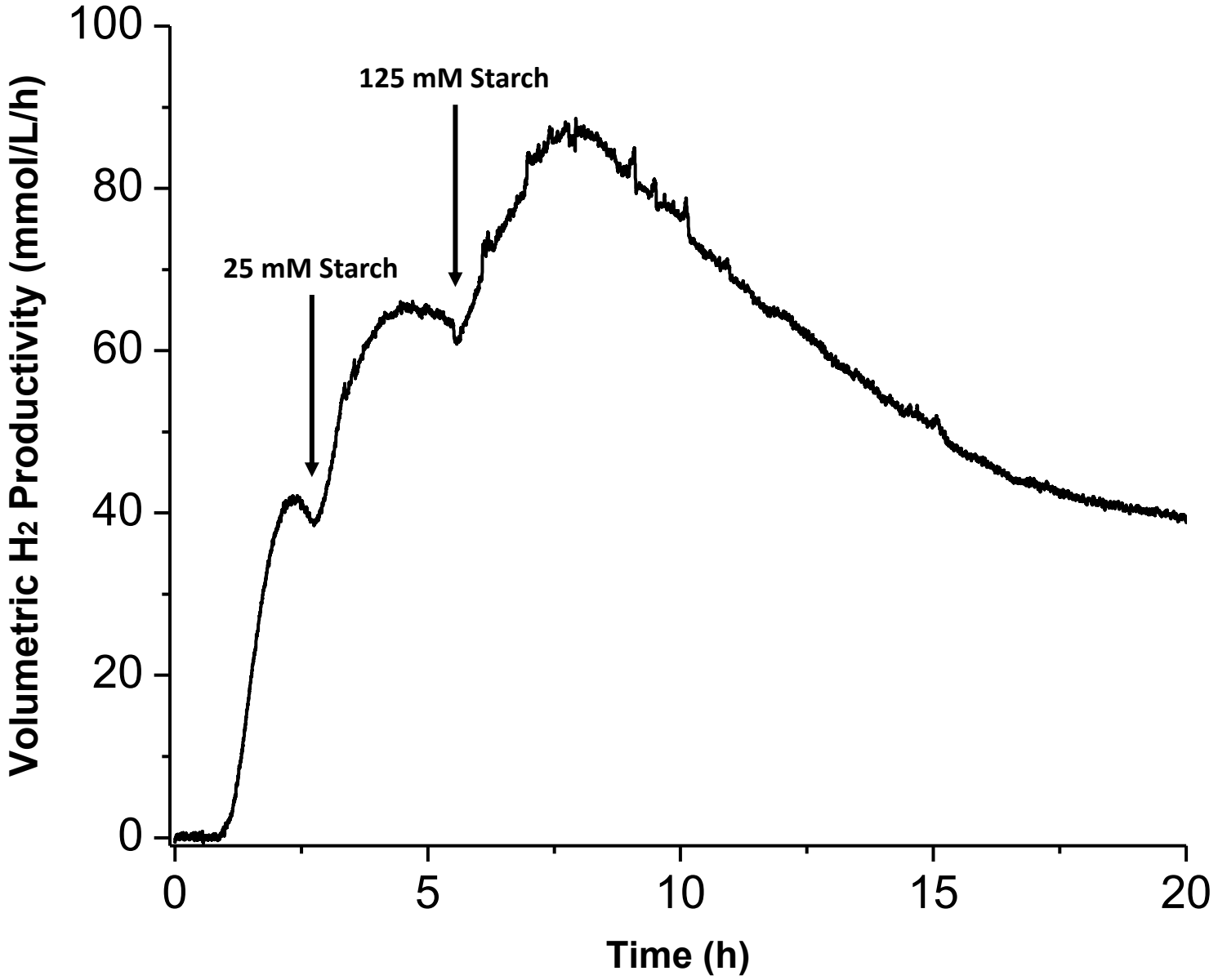
100
80
60
40
20
0
Volumetric H2 Productivity (mmol/L/h)
125 mM Starch
25 mM Starch
0
5
10
15
20
Time (h)

Figure 4.

Chapter 4. Biosynthesis of D-xylulose 5-Phosphate from D-xylose and polyphosphate through a minimized two-enzyme cascade

Kim, J.-E. and Zhang, Y.H.P. (2016). Biosynthesis of D-xylulose 5-phosphate from D-xylose and polyphosphate through a minimized two-enzyme cascade. Biotechnology and bioengineering, 113(2), pp.275-282. Reproduced by permission of John Wiley & Sons.

ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Biosynthesis of D-Xylulose 5-Phosphate From D-Xylose and Polyphosphate Through a Minimized Two-Enzyme Cascade

Jae-Eung Kim,[1] Y.-H. Percival Zhang[1,2,3,4]

[1]Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall, Blacksburg 24061, Virginia; telephone: +540-231-7414; fax: +540-231-3199; e-mail: ypzhang@vt.edu

[2]Cell Free Bioinnovations Inc., Blacksburg, Virginia

[3]Institute for Critical Technology and Applied Science (ICTAS), Virginia Tech, Blacksburg, Virginia

[4]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China

ABSTRACT: Sugar phosphates cannot be produced easily by microbial fermentation because negatively-charged compounds cannot be secreted across intact cell membrane. D-xylulose 5-phosphate (Xu5P), a very expensive sugar phosphate, was synthesized from D-xylose and polyphosphate catalyzed by enzyme cascades in one pot. The synthetic enzymatic pathway comprised of xylose isomerase and xylulokinase was designed to produce Xu5P, along with a third enzyme, polyphosphate kinase, responsible for in site ATP regeneration. Due to the promiscuous activity of the ATP-based xylulokinase from a hyperthermophilic bacterium *Thermotoga maritima* on polyphosphate, the number of enzymes in the pathway was minimized to two without polyphosphate kinase. The reactions catalyzed by the two-enzyme and three-enzyme pathways were compared for Xu5P production, and the reaction conditions were optimized by examining effects of reaction temperature, enzyme ratio and substrate concentration. The optimized two-enzyme system produced 32 mM Xu5P from 50 mM xylose and polyphosphate after 36 h at 45°C. Biosynthesis of less costly Xu5P from D-xylose and polyphosphate could be highly feasible via this minimized two-enzyme pathway.

Biotechnol. Bioeng. 2016;113: 275–282.

KEYWORDS: sugar phosphate; D-xylulose 5-phosphate; synthetic enzymatic pathway; in vitro synthetic biosystem; xylose; ATP regeneration

Correspondence to: Y.-H. Percival Zhang

Contract grant sponsor: NSF SBIR II
Contract grant number: IIP-1353266
Contract grant sponsor: Virginia Agricultural Experiment Station
Contract grant sponsor: National Institute of Food and Agriculture, U.S. Department of Agriculture
Received 20 April 2015; Revision received 22 July 2015; Accepted 27 July 2015
Accepted manuscript online 4 August 2015;
Article first published online 3 September 2015 in Wiley Online Library (http://onlinelibrary.wiley.com/doi/10.1002/bit.25718/abstract).
DOI 10.1002/bit.25718

Introduction

Because many organic syntheses involving sugars are difficult to carry out, new biological routes are constantly being sought (Bhaskar et al., 2012). For example, chemical synthesis of D-xylulose 5-phosphate (Xu5P) involves complicated steps of protection and de-protection of functional groups, resulting in low yields associated with high production costs (Shaeri et al., 2008). As compared to chemical catalysis, biotransformation features better chemical selectivity, more modest and environmentally friendly reaction conditions (Bornscheuer et al., 2012; Kwon et al., 2012; Zhang, 2015). Although whole-cell microorganisms have been used to produce a large number of products, such as beer, wine, lactic acid, butanol, and so on (de Carvalho, 2011), they have some inherent weaknesses, such as relatively low product yields, complicated cellular regulations, difficulties in the production of some specific products, for example, sugar phosphates. In contrast, enzyme-based biocatalysis is becoming an important alternative (Bornscheuer et al., 2012; Zhang, 2015). Recently, in vitro synthetic biosystems comprised of synthetic or artificial enzymatic pathways emerge as a promising biomanufacturing platform for producing desired products better than microorganisms in terms of product yield, reaction rate, product separation cost, tolerance of toxic product or substrate and, easy process control (Kwon et al., 2012; Tessaro et al., 2015; Zhang, 2015). Via this novel platform, theoretical maximum yields of hydrogen production have been achieved from sugars and water, which are 2–3 times higher than theoretical yields from microbial fermentations (Martin del Campo et al., 2013b; Myung et al., 2014; Rollin et al., 2015); a nearly theoretical yield of 1,3-propandiol production has been produced from glycerol (Rieckenberg et al., 2014); non-food cellulose has been biotransformed to starch (You et al., 2013); high-yield butanol has been produced from glucose (Guterl et al., 2012; Krutsakorn et al., 2013); 2-deoxyribose 5-phosphate has been produced from fructose and polyphosphate (Honda et al., 2010); dihydroxyacetone

phosphate has been synthesized via a partial glycolysis pathway (Bujara et al., 2011; Meyer et al., 2007).

Xu5P is a key sugar phosphate intermediate of the pentose phosphate pathway. It is used as an essential substrate for the enzyme assay of D-ribulose 5-phosphate 3-epimerase and transketolase (Lee et al., 2008; Wood, 1973). Also, Xu5P is an important metabolite related to the study of metabolic diseases (Iizuka and Horikawa, 2008; Shaeri et al., 2008). For example, Xu5P has been reported as an activator for protein phosphatase 2A, which dephosphorylates a carbohydrate response element binding protein and activates its transactivity. The activated carbohydrate response element binding protein controls approximately 50% of hepatic lipogenesis by regulating glycolytic and lipogenic gene expression (Iizuka and Horikawa, 2008). Despite of its importance in those highlighted fields, the commercial availability of Xu5P is limited to few vendors with the list selling price of more than $2,500 for 10 mg.

Several studies of enzymatic biosynthesis of Xu5P have been carried out based on the same enzymatic pathway containing transketolase to synthesize Xu5P from two initial substrates, hydroxypyruvate with either fructose 1,6-bisphosphate, glyceraldehyde 3-phosphate or dihydroxyacetone phosphate (Shaeri et al., 2008; Wood, 1973; Zimmermann et al., 1999) (Fig. 1A). Two more enzymes, aldolase and/or triosephosphate isomerase, may be required depending on given substrates. Although this enzymatic pathway has a flexibility in selecting initial substrates and the last irreversible Xu5P-synthesizing reaction by releasing CO_2, it suffers from costly substrates and labile intermediates—glyceraldehyde 3-phosphate and dihydroxyacetone phosphate, resulting in a narrow biocatalytic window (Shaeri et al., 2008).

A partial pathway of the natural xylose-utilization pathway may be used to produce Xu5P from D-xylose by using two cascade enzymes—xylose isomerase (XI) and xylulokinase (XK) (Fig. 1B). However, its feasibility for Xu5P synthesis was not reported, to our limited knowledge. Because ATP is a costly substrate, the ATP regeneration system is of importance for reducing ATP use, driving the reaction to completion, simplifying product isolation, and preventing the accumulation of inhibitory cofactor by-products (Zhao and van der Donk, 2003). Several methods have been developed to regenerate ATP from ADP using whole-cell based regeneration methods (Endo and Koizumi, 2001) or enzymatic methods, such as pyruvate kinase (Hirschbein et al., 1982), acetate kinase (Shih and Whitesides, 1977), or polyphosphate kinase (PPK) (Shiba et al., 2000). ATP can also be regenerated from AMP with polyphosphate using polyphosphate AMP phosphotransferase

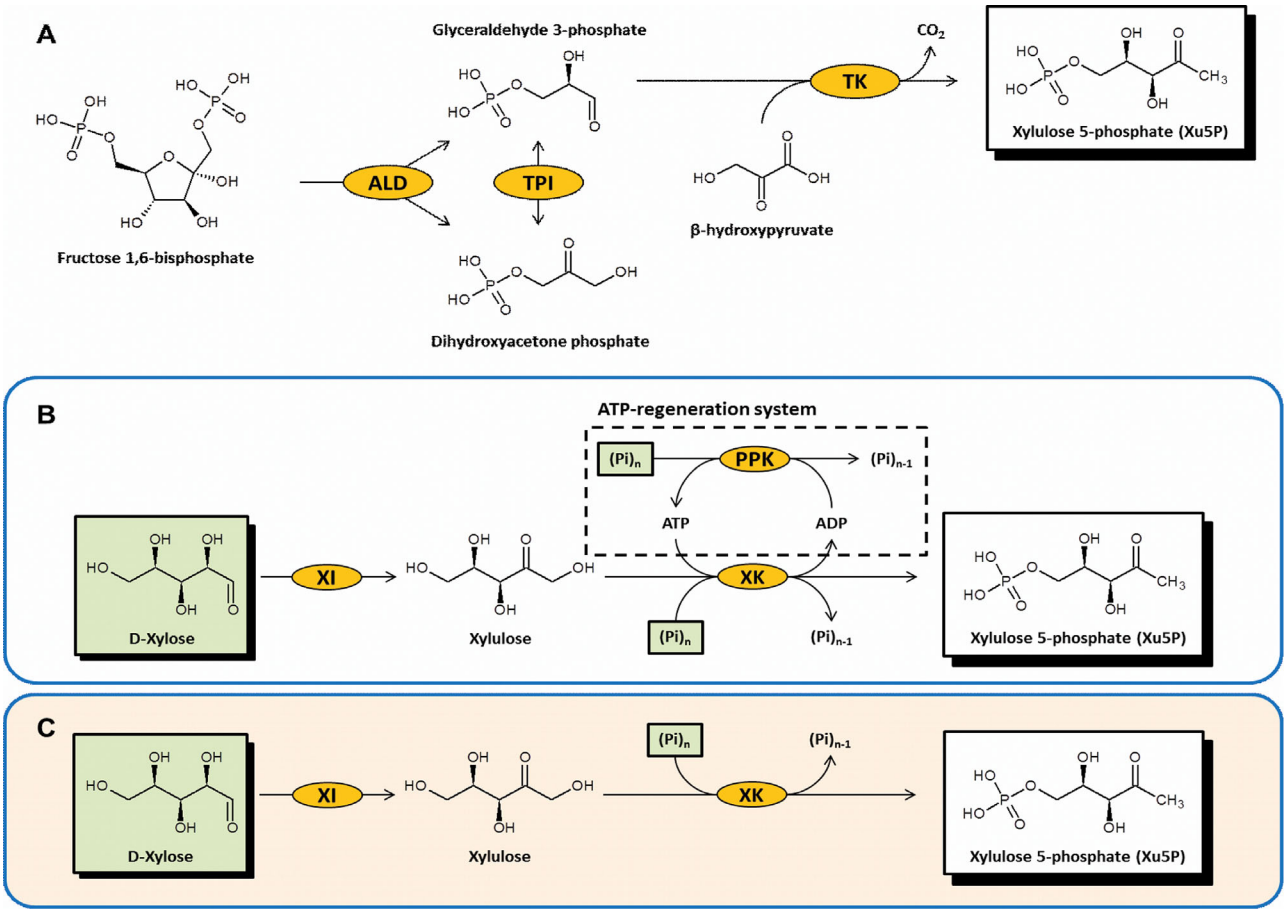

Figure 1. Scheme of enzymatic pathways for the production of xylulose 5-phosphate by the enzymatic pathway for Xu5P synthesis from fructose 1,6-bisphosphate or glyceraldehyde 3-phosphate or dihydroxyacetone phosphate with hydroxypyruvate (**A**), by the three-enzyme pathway containing XI/XK/PPK with ATP-regenerating system (**B**) and by the minimized two-enzyme pathway containing XI/XK (**C**).

conjugated with either acetylate kinase (Resnick and Zehnder, 2000) or PPK (Kameda et al., 2001). Among these methods, the ATP regeneration system based on PPK is the most preferred method because it is simple and requires readily available low-cost polyphosphate only (Honda et al., 2010; Iwamoto et al., 2007; Sato et al., 2007). An in-depth comparison of various ATP regeneration methodologies has been reviewed recently (Andexer and Richter, 2015).

In this study, we investigated the biosynthesis of Xu5P from xylose and polyphosphate mediated by two de novo synthetic enzymatic pathways (Fig. 1B and C). The three-enzyme pathway contained XI, XK and PPK with in situ ATP regeneration (Fig. 1B). Also, the minimized two-enzyme pathway comprised of XI and XK without ATP regeneration (Fig. 1C) was investigated because a recently-discovered XK from *Thermotoga maritima* exhibited a significant promiscuous activity on polyphosphate (Martin del Campo et al., 2013a).

Materials and Methods

Chemicals and Strains

All chemicals were reagent or higher grade, and purchased from Sigma–Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), unless otherwise noted. Polyphosphate purchased from Sigma–Aldrich was sodium hexametaphosphate. Two other polyphosphates were purchased from eBay (sodium hexametaphosphate grahams salt, Chemistry Connection) and from Tianjin Bodi Chemical Engineering Co. (Tianjin, China), respectively. Immobilized xylose isomerase from *Streptomyces murinus* was purchased from Sigma–Aldrich. The high-fidelity Phusion DNA polymerase and the protein marker were purchased from New England Biolabs (Ipswich, MA). Primers were purchased from IDT (Coralvile, IA). Genomic DNA samples of *T. maritima* MSB8 (ATCC43589) and *Thermus thermophiles* HB27 were purchased from the American Type Culture Collection (Manassas, VA). *E. coli* TOP10 was used for DNA manipulation and plasmid amplification. *E. coli* BL21 (DE3) (Invitrogen, Carlsbad, CA) and pET20b (+) were used for protein expression.

Preparation of Enzymes

Recombinant *T. maritima* xylulokinase (XK) expressed in *E. coli* was prepared and its specific activity was measured as described previously (Martin del Campo et al., 2013a). The coding DNA sequence (CDS) TTC0637 encoded for PPK was amplified by PCR with a pair of primers IF (5′-**TAACT TTAAG AAGGA GATAT ACATA TG**CAC CTCCT TCCCG AAGCA AGCTG GCTC-3′) and IR (5′-TCAGT GGTGG TGGTG GTGGT GCTCG AG**TAG CTCCA GGCGC TGGGC GTCGT GGCG**-3′) based on genomic DNA of *T. thermophiles* HB27. A linear vector backbone was amplified based on pET20b (+) with a pair of primers VF (5′-CGCCA CGACG CCCAG CGCCT GGAGC TA**CTC GAGCA CCACC ACCAC CACCA CTGA**-3′) and VR (5′-**GAGCC AGCTT GCTTC GGGAA GGAGG TG**CAT ATGTA TATCT CCTTC TTAAA GTTA-3′). VF and VR contain the last 27 bp of the 3′ terminus of the insertion sequence (underlined) and the first 27 bp of the 5′ terminus of the vector sequence (highlighted). The two PCR products were assembled by prolonged overlap extension PCR (POE-PCR) (You et al., 2012). POE-PCR conditions were as followings: initial denaturation (30 s at 98°C), 30 cycles of denaturation (20 s at 98°C), annealing (10 s at 60°C), and elongation (165 s at 72°C), and a final extension step (10 min at 72°C). The POE-PCR product was transferred to *E. coli* TOP10, yielding plasmid pET20b-ttcPPK.

To enhance PPK expression in *E. coli*, the DNA sequence of PPK was codon-optimized by GenScript (Piscataway, NJ), yielding plasmid pET20b-ttcPPK-co. The recombinant plasmid pET20b-ttcPPK-co was transferred into *E. coli* BL21 (DE3) to express the recombinant PPK in the LB medium. The *E. coli* culture was grown at 37°C in 250 mL Erlenmeyer flasks containing 50 mL of the LB medium plus 100 μg/mL of ampicillin. The expression of PPK was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when the culture A_{600} reached about 0.6, then cultivated at 18°C for 16 h. The cells were harvested by centrifugation at 4°C and washed with 50 mM of Tris–HCl buffer (pH 8.5) containing 0.5 M NaCl. The cell pellets were suspended in the same buffer and lysed by ultrasonication in an ice bath (Fisher Scientific sonic Dismembrator Model 500; 3 s pulse on and 8 s off, total 180 s at 30% amplitude). After heat treatment at 70°C for 15 min to deactivate *E. coli* endogenous enzymes followed by centrifugation, the insoluble membrane-bound PPK in the pellets was collected. The specific activity of PPK on hexametaphosphate was measured as described elsewhere (Iwamoto et al., 2007; Restiawaty et al., 2011). Purities of XK and PPK were analyzed by SDS-PAGE.

HPLC Analysis

Concentrations of D-xylose, D-xylulose, and D-xylose 5-phosphate were measured by the SHIMADZU high performance liquid chromatography (HPLC) equipped with a refractive index detector. Samples were separated on a Bio-Rad Aminex HPLC organic acid column (HPX-87H 300 × 7.8 mm^2) at 60°C with a mobile phase of 5 mM sulfuric acid solution at a rate of 0.6 mL/min. Under these conditions, the retention times of xylose, xylulose and Xu5P were 9.73, 10.16, and 6.52 min, respectively.

Enzymatic Synthesis of Xu5P

One-pot enzymatic reactions were conducted in 3 mL of the 30 mM sodium citrate buffer (pH 6.2) containing 10 mM $MgCl_2$, 20 mM polyphosphate, 20 mM xylose, and 0.05% sodium azide. Polyphosphate was replaced with 2 mM or 20 mM ATP as the control reactions. The two enzymes: XI and XK were added into the reaction solution with the final concentration of 1.0 U/mL for each enzyme. PPK (0.02 U/mL) was added for ATP regeneration. Reaction profiles were measured by withdrawing samples at different times. H_2SO_4-acidified samples were frozen overnight and thawed prior to centrifugation to remove any undissolved small particles. The supernatants were used for the quantification analysis by HPLC.

Effects of Reaction Temperature

The reactions containing 20 mM xylose with 20 mM polyphosphate catalyzed by the two enzymes (XI and XK) and three enzymes

(XI, XK and PPK) were carried out in three different temperatures: 30°C, 45°C, and 60°C. The three enzyme system was supplemented with 2 mM ATP.

Effect of Enzyme Ratio

The reactions catalyzed by XI and XK were carried out at a fixed total enzyme loading of 5 U/mL at 45°C. The XI and XK loadings were 1, 1.7, 2.5, 3.4, and 4 U/mL as well as 4, 3.4, 2.5, 1.7, and 1 U/mL, respectively. The reaction solution contained 20 mM xylose and 20 mM polyphosphate. The synthesis of Xu5P was analyzed at 18 and 36 h.

Effect of Substrate Concentration on Synthesis of Xu5P

The reactions catalyzed by XI and XK at a unit ratio of 2 were carried out in a various xylose concentration from 10, 20, 50 to 100 mM at 45°C. In these reactions, the concentration ratio of polyphosphate to xylose was 1:1, and the total enzyme loading to xylose was kept at a ratio of 2.5 U/10 mM xylose. Samples were taken at 36 h to analyze Xu5P production by HPLC. The Xu5P conversion was calculated with the amount of Xu5P produced divided by the initial xylose concentration.

Results

Overexpression and Purification of Recombinant Enzymes

Two plasmids, pET20b-TmXylK and pET20b-ttcPPK encoding XK and PPK, respectively, were constructed and verified by DNA sequencing. XK and PPK were expressed in *E. coli* BL21 (DE3) growing in the LB medium. The recombinant protein synthesis was induced by IPTG addition at 18°C after the absorbance of the culture at 600 nm reached about 0.6. According to SDS-PAGE gel, a majority of XK was expressed in the soluble form (Fig. 2, Lane S under XK). The XK with a C-terminal His-tag was purified to homogeneity by using a nickel-resin (Fig. 2, Lane P under XK), exhibiting a single band in a molecular weight of approximately 55 kDa, which is in agreement with the calculated molecular mass of 57.5 kDa from its deduced amino acid sequence. The specific activity of XK on ATP was 20 U/mg at 30°C, 32 U/mg at 45°C, and 49 U/mg at 60°C. The wild-type *ppk* gene from *T. thermophiles* was not overexpressed well even after optimizing a series of expression conditions, such as IPTG concentrations, cell culture temperatures, and different *E. coli* hosts (e.g., *E. coli* BL21 Rosetta) (data not shown). To enhance PPK expression in *E. coli*, a codon-optimized *ppk* gene was synthesized without rare codons. The recombinant codon-optimized PPK were expressed (Fig. 2**,** Lane T under PPK), but its majority was present in the insoluble fraction (Fig. 2, Lane I under PPK). Different from most soluble recombinant proteins, PPK is an outer membrane-associated protein (Restiawaty et al., 2011) so that the recombinant PPK was present in the pellets of the cell extract. Because this enzyme is hyperthermostable, the heat treatment at 75°C for 15 min deactivated *E. coli* cellular proteins but not this enzyme. The active PPK in the pellets (Fig. 2, Lane H under PPK) was used to regenerate ATP. The specific activity of PPK on hexametaphosphate was 0.0006 U/mg at 30°C, 0.020 U/mg at 45°C, and 0.039 U/mg at 60°C.

HPLC Analysis

To quantitatively determine substrates (e.g., xylose, ATP, phosphate, polyphosphate), intermediates (e.g., xylulose, ADP), and product (i.e., xylulose 5-phosphate), the several mobile phases for HPLC equipped with the Amex HPX-87H column were tested to minimize the interference from mobile phases and buffer salts. Among several possible mobile phases, such as phosphoric acid, sulfuric acid and trifluoroacetic acid, 5 mM sulfuric acid was chosen for HPLC analysis. Several buffer solutions were also tested in order to avoid the interference to the product measurement. Sodium citrate buffer had a retention time of 8.01 min without a significant overlap with any of chemicals listed above. No considerable reduction of the apparent enzyme activity was found in 30 mM sodium citrate buffer (pH 6.2) as

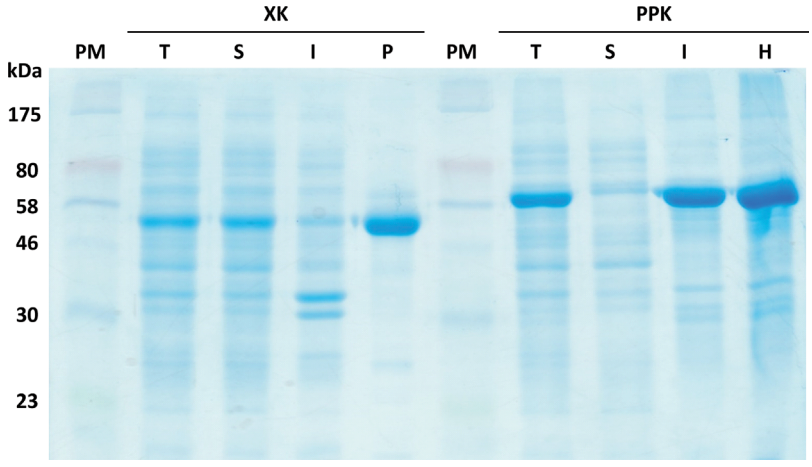

Figure 2. SDS-PAGE gel analysis of recombinant enzyme expression and purification of XK and PPK in *E. coli*. **PM**, protein marker; **T**, total cell lysate; **S**, soluble fraction of cell lysate; **I**, insoluble fraction of cell lysate; **P**, purified protein by Ni-charged resin; **H**, Heat treated protein.

compared to HEPES and Tris buffers (data not shown). Thus, sodium citrate buffer was chosen as an enzymatic buffer (Fig. 3). Mg^{2+}, a metal required for XK activity, has a retention time of 6.22 min.

D-xylose in the citrate buffer containing Mg^{2+} catalyzed by XI resulted in two peaks: xylose at a retention time of 9.73 min and xylulose at a retention time of 10.16 min (Fig. 3A). When borax was added, it switched the equilibrium toward xylulose, resulting in a much large peak for xylulose (Fig. 3B). ATP peak has a retention time of 6.78 min (Fig. 3C), very close to the Xu5P peak at a retention time of 6.52 min (Fig. 3D). When the reaction was catalyzed by XI/XK supplemented with ATP, the Xu5P peak was not observed while there was an ATP peak at the beginning of the reaction (Fig. 3C). At the end of the reaction, the Xu5P peak appeared and the ATP peak disappeared, suggesting the formation of Xu5P (Fig. 3D).

Enzymatic Synthesis of Xu5P

The Xu5P profiles catalyzed by the two-enzyme (XI and XK) and three-enzyme (XI, XK and PPK) systems are shown in Figure 4. In the two-enzyme system supplemented with 2 and 20 mM ATP, approximately 2 and 8 mM Xu5P were produced after the first hour, respectively. For the reaction containing 2 mM ATP, Xu5P production stopped due to the ATP depletion. For the reaction with 20 mM ATP, 10.8 mM Xu5P was produced after 2 h. Later, Xu5P generation rates significantly decreased possibly due to the product inhibition. The final Xu5P level from the reaction with 20 mM ATP was about 17 mM after 48 h. The two-enzyme reaction supplemented with 2 or 20 mM ATP achieved the final conversion of 10% (100% based on ATP) or 85% (85% based on ATP), respectively.

To decrease ATP consumption, another enzyme PPK (0.02 U/mL) was added, resulting in the three-enzyme system (XI, XK, and PPK) with a decreased ATP concentration of 2 mM. Twenty millimolar polyphosphate was added for ATP regeneration. The initial production rate was as fast as the two-enzyme reaction with 2 mM ATP for the first hour. After the complete consumption of initial ATP, more Xu5P was generated gradually due to ATP regeneration. The final Xu5P concentration was 14.5 mM Xu5P after 48 h, resulting in the xylose conversion of 73%.

Because XK from *T. maritima* has a significant promiscuous activity on polyphosphate (Martin del Campo et al., 2013a), the three-enzyme system with ATP regeneration could be minimized to the two-system without ATP. This two-enzyme (XI/XK) reaction was supplemented with 20 mM polyphosphate without PPK and ATP. Since XK retains lower specific activity on polyphosphate than the specific activity on ATP (Martin del Campo et al., 2013a), the minimized two-enzyme reaction based on polyphosphate showed slower production rates than the three-enzyme system and the two-enzyme system supplemented with 20 mM ATP. The minimized two-enzyme mixture gradually produced 11.6 mM Xu5P for 48 h. The comparison of the three-enzyme reaction with ATP regeneration to the minimized two-enzyme reaction indicated that the addition of PPK for ATP regeneration resulted in only a marginal increase in the production rate and about 14% improvement in the product conversion.

Effect of Reaction Temperature on the Synthesis of Xu5P

Xu5P synthesis catalyzed by the two-enzyme (XI/XK) system without ATP regeneration and the three-enzyme (XI/XK/PPK) system with ATP regeneration were investigated at three different reaction temperatures of 30, 45, and 60 °C (Fig. 5). The Xu5P production rates increased as increasing the reaction temperature from 30 to 45°C, but product concentration was very low at 60°C for both cases possibly due to the degradation of Xu5P (Fig. 5A and B). At 30 and 45°C, the three-enzyme systems exhibited only slightly faster production rate than the two-enzyme systems, indicating a marginal benefit from ATP regeneration by PPK. From 30 to 45°C, the increment in the production rate was higher in the three-enzyme system than in the two-enzyme system. This increment could be explained by more increased activities of PPK at higher

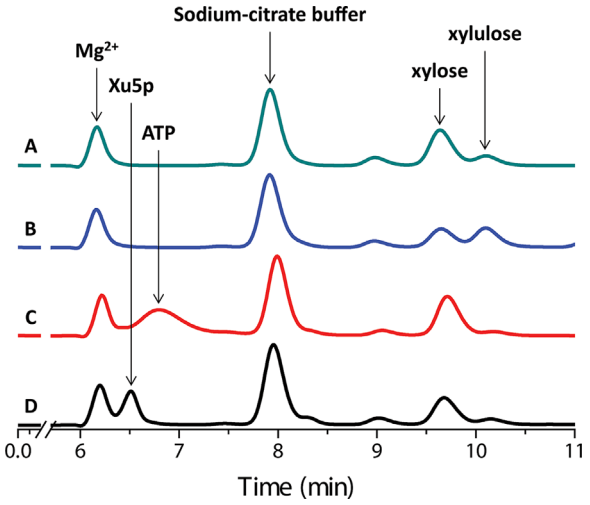

Figure 3. HPLC chromatography of the Bio-Rad Aminex HPX-87H column with a refractive index detector. The retention times of xylose, xylulose, ATP and Xu5P were 9.73, 10.16, 6.80, and 6.52 min, respectively (**A**,**C**, and **D**). The addition of borax shifted the enzymatic reaction of XI in favor to xylulose formation (**B**).

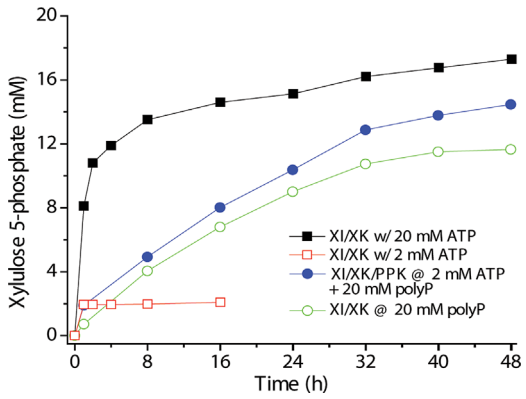

Figure 4. Xylulose 5-phosphate production profiles at 30°C by using XI and XK with 20 mM ATP (■), XI and XK with 2 mM ATP (□), XI, XK and PPK with 2 mM ATP and 20 mM polyphosphate (●), XI and XK with 20 mM polyphosphate (○). 1.0 U/mL for each XI and XK was used. 0.02 U/mL of PPK was used for ATP regeneration.

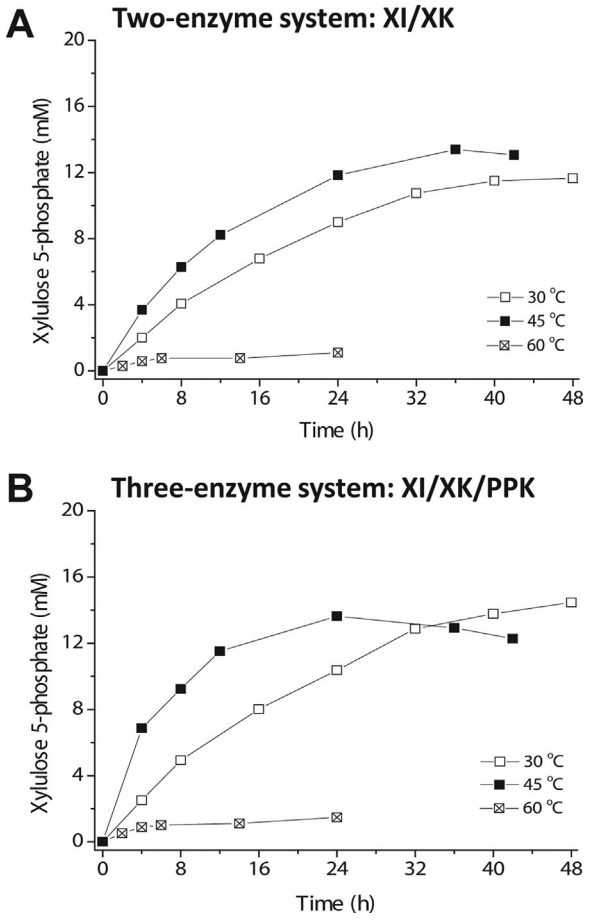

Figure 5. Effect of reaction temperature and ATP-recycling system by PPK on Xu5P production rates and conversions. Enzymatic reactions were supplemented with 20 mM xylose and 20 mM polyphosphate as substrates. 2 mM ATP was supplied to the three-enzyme system with ATP-regeneration.

temperature than other enzymes. The highest conversions of the two-enzyme system at 30 and 45°C were about 60% and 70%, respectively. From the three-enzyme system, about 72% and 68% of the maximum conversion were achieved from 30 and 45°C, respectively. Due to a margin enhancement of product conversion in the three-enzyme system, the minimized two-enzyme system without ATP at 45°C was used for further investigation of effects of enzyme loading and substrate loading.

Effect of Enzyme Ratio on the Synthesis of Xu5P

The minimized XI/XK system was further investigated to find an optimal enzyme ratio at 45°C (Fig. 6). The total enzyme loading was 5 U/mL, and the ratio between XI and XK was changed from 0.25 to 4. The conversions gradually increased from about 33% to 68% as the XI/XK ratio increased from 0.25 to 2. The highest conversions of Xu5P production were obtained at the XI/XK ratio of 2 regardless of reaction time (18 or 36 h), producing 9.8 mM and 13.5 mM Xu5P, respectively. At the ratio of 4, the conversions decreased possible due to the limited amount of XI.

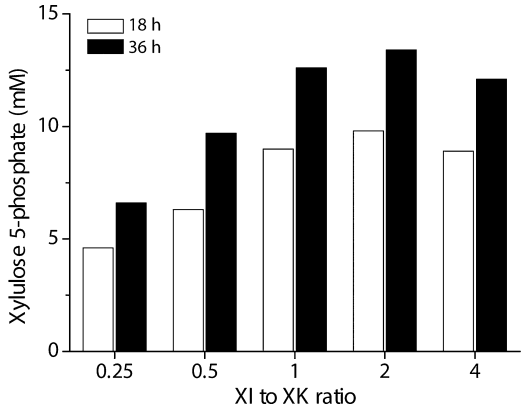

Figure 6. Effect of enzyme loading ratio of XI to XK. The two-enzyme system (XI and XK) supplemented with 20 mM xylose and 20 mM polyphosphate was carries out with the total enzyme loading of 5 U/mL. The ratio between XI and XK was changed from 0.25 to 4.

Effect of Substrate Concentration on the Synthesis of Xu5P

Enzymatic synthesis of Xu5P catalyzed by the minimized XI/XK cocktail was conducted with varying substrate concentrations from 10 to 100 mM xylose with polyphosphate corresponding to xylose concentrations at a fixed substrate/enzyme ratio at 45°C (Fig. 7). The Xu5P titer increased from approximately 9 mM to 32 mM with the increasing substrate concentration from 10 to 50 mM. However, the product conversions decreased from 90% to 26% when the substrate concentration increased from 10 to 100 mM. The higher substrate concentration did not result in higher product titer possibly due to high ionic strength and/or substrate/product inhibition.

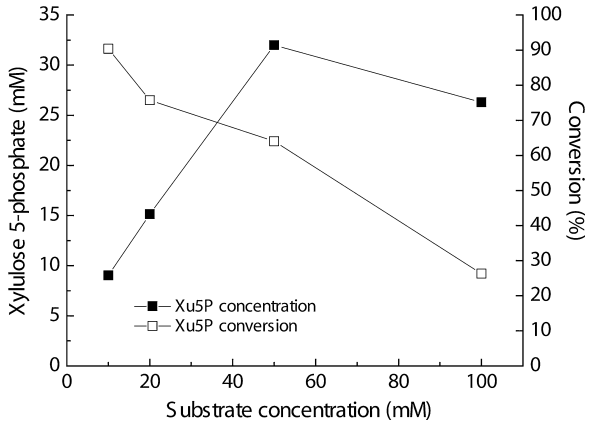

Figure 7. Effect of substrate loading on Xu5P production. The reactions were supplemented with varying concentrations of xylose and polyphosphate from 10 to 100 mM, and the total enzyme (XI and XK) was added according to the substrate concentration to maintain the ratio of enzyme to substrate constant.

Discussion

The minimized two-enzyme pathway containing XI and XK was designed and validated to produce Xu5P from xylose and polyphosphate without ATP regeneration (Fig. 1C) because the hyperthermophilic bacterium *T. maritima* XK can phosphorylate xylulose with polyphosphate (Martin del Campo et al., 2013a). Although the first step catalyzed by XI was reversible, the second step catalyzed by XK had a negative Gibbs energy of -12.6 ± 4.0 kJ/mol (http://equilibrator.weizmann.ac.il/), resulting in relatively high product conversions by pulling xylulose toward Xu5P synthesis. This analysis explained why the addition of borax into the whole cascade did not increase the product conversion (data not shown). This minimized two-enzyme pathway was simpler in terms of enzyme and substrate than the three-enzyme pathway featuring in situ ATP regeneration, but had slight decreases in product titer and volumetric productivity (Fig. 5).

Compared to the previous traditional three-enzyme pathway comprised of ALD, TIP and TK (Shaeri et al., 2008), this minimized two-enzyme pathway had numerous advantages. First, much cheaper substrates (xylose and polyphosphate) were used instead of costly fructose 1,6-biphosphate and hydroxypyruvate. Xylose is the most abundant pentose in the nature, especially in plant cell walls. Polyphosphate, an inorganic polymer containing tens to hundreds of phosphate residues linked by high-energy phosphoanhydride bonds polyphosphate, is used as a direct phosphorus donor for biochemical reactions (Kulaev and Kulakovskaya, 2000). Polyphosphate is a much less costly phosphorus donor than ATP because of its simple preparation process. Second, the two-enzyme pathway had broader reaction windows in terms of pH, temperature and substrate concentration because the intermediate (xylulose) is far more stable than DHAP and G3P, which are infamous for their rapid degradation at high temperature (Du et al., 2011; Kouril et al., 2013; Zhang, 2011). Third, the two-enzyme pathway has simpler process control because the traditional three-enzyme system generated CO_2 from the aqueous solution, requiring an extra effort to control the decreasing pH of the reaction buffer (Shaeri et al., 2008) or use high-salt buffers. Fourth, 32 mM Xu5P was produced from 50 mM xylose (Fig. 7), higher than the Xu5P levels catalyzed by the ALD, TIP and TK mixture (Shaeri et al., 2008), suggesting potentially low Xu5P separation cost. Fifth, both XI and XK enzymes had much better stability at working temperatures (e.g., 30–60°C) than the enzymes used for the traditional enzymatic pathway (Shaeri et al., 2008), because they originated from thermophilic microorganisms.

Although polyphosphate is a low-cost phosphate group donor because it is produced via chemical or biological approaches and sold even as a fertilizer, it was pretty challenging to purchase high-quality polyphosphate. Sigma–Aldrich discontinued some types of polyphosphate and sold hexametaphosphate only now. Two other polyphosphates, sodium hexametaphosphate grahams salts purchased from eBay (Chemistry Connection) and the chemical reagent polyphosphate from Tianjin Bodi Chemical Engineering Co., were tested to compare with Sigma–Aldrich's one. Polyphosphate from Chemistry Connection had similar water solubility to Sigma's, and XK showed similar activities on both substrates. However, polyphosphate from Tianjin Bodi Chemical Engineering Co. had a very low water solubility, indicating that it had a very high degree of polymerization. XK also showed very low activities on this substrate (data not shown).

Most sugar phosphates (e.g., Xu5P, glucose 1-phosphate, dihydroxyacetone phosphate) are monosaccharides having substituted phosphate groups. Some of them are important metabolites in glycolysis and pentose phosphate pathway for transferring or storing energy and being the backbone for DNA and RNA synthesis. Due to the phosphate group, negatively charged and hydrophilic sugar phosphates cannot pass through intact cell membrane so that they cannot be fermented efficiently by regular microorganisms. Enzymatic biocatalysis including this study provides great opportunities to synthesize sugar phosphates without concerns of cell membrane, product separation and possible contamination from other metabolites or microbial pathogens (Bujara et al., 2011; Fessner, 2014; Honda et al., 2010; Meyer et al., 2007).

Conclusions

D-Xylulose 5-phosphate was biosynthesized from D-xylose and polyphosphate catalyzed by a minimized two-enzyme pathway comprised of xylose isomerase and polyphosphate-utilizing xylulokinase. Approximately 32 mM Xu5P was produced from 50 mM xylose, achieving 64% conversion. This simple cascade reaction without involvement of labile metabolites could provide a cost-effective way to product Xu5P under the broad reaction conditions.

This work could not be finished without the support of Virginia Tech. In addition, funding to YPZ for this work was provided in part, by NSF SBIR II award (IIP-1353266), DOE STTR I award and by the Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture.

References

Andexer JN, Richter M. 2015. Emerging enzymes for ATP regeneration in biocatalytic processes. ChemBioChem 16(3):380–386.

Bhaskar U, Sterner E, Hickey A, Onishi A, Zhang F, Dordick J, Linhardt R. 2012. Engineering of routes to heparin and related polysaccharides. Appl Microbiol Biotechnol 93(1):1–16.

Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. 2012. Engineering the third wave of biocatalysis. Nature 485(7397):185–194.

Bujara M, Schümperli M, Pellaux R, Heinemann M, Panke S. 2011. Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. Nat Chem Biol 7(5):271–277.

de Carvalho CCCR. 2011. Enzymatic and whole cell catalysis: Finding new strategies for old processes. Biotechnol Adv 29(1):75–83.

Du J, Say RF, Lü W, Fuchs G, Einsle O. 2011. Active-site remodelling in the bifunctional fructose-1, 6-bisphosphate aldolase/phosphatase. Nature 478(7370):534–537.

Endo T, Koizumi S. 2001. Microbial conversion with cofactor regeneration using genetically engineered bacteria. Adv Synth Catal 343(6-7):521–526.

Fessner WD. 2014. Systems Biocatalysis: Development and engineering of cell-free "Artificial Metabolisms" for preparative multi-enzymatic synthesis. New Biotechnol DOI: 10.1016/j.nbt.2014.11.007.

Guterl JK, Garbe D, Carsten J, Steffler F, Sommer B, Reiße S, Philipp A, Haack M, Rühmann B, Kettling U, et al. 2012. Cell-free metabolic engineering—Production of chemicals via minimized reaction cascades. ChemSusChem 5(11):2165–2172.

Hirschbein BL, Mazenod FP, Whitesides GM. 1982. Synthesis of phosphoenolypyruvate and its use in ATP cofactor regeneration. J Org Chem 47(19):3765–3766.

Honda K, Maya S, Omasa T, Hirota R, Kuroda A, Ohtake H. 2010. Production of 2-deoxyribose 5-phosphate from fructose to demonstrate a potential of artificial bio-synthetic pathway using thermophilic enzymes. J Biotechnol 148(4): 204–207.

Iizuka K, Horikawa Y. 2008. ChREBP: A glucose-activated transcription factor involved in the development of metabolic syndrome. Endocr J 55(4):617–624.

Iwamoto S, Motomura K, Shinoda Y, Urata M, Kato J, Takiguchi N, Ohtake H, Hirota R, Kuroda A. 2007. Use of an Escherichia coli recombinant producing thermostable polyphosphate kinase as an ATP regenerator to produce fructose 1,6-diphosphate. Appl Environ Microbiol 73(17):5676–5678.

Kameda A, Shiba T, Kawazoe Y, Satoh Y, Ihara Y, Munekata M, Ishige K, Noguchi T. 2001. A novel ATP regeneration system using polyphosphate-AMP phosphotransferase and polyphosphate kinase. J Biosci Bioeng 91(6):557–563.

Kouril T, Esser D, Kort J, Westerhoff HV, Siebers B, Snoep JL. 2013. Intermediate instability at high temperature leads to low pathway efficiency for an in vitro reconstituted system of gluconeogenesis in *Sulfolobus solfataricus*. FEBS J 280(18):4666–4680.

Krutsakorn B, Honda K, Ye X, Imagawa T, Bei X, Okano K, Ohtake H. 2013. In vitro production of n-butanol from glucose. Metab Eng 20:84–91.

Kulaev I, Kulakovskaya T. 2000. Polyphosphate and phosphate pump. Annu Rev Microbiol 54(1):709–734.

Kwon SJ, Mora-Pale M, Lee MY, Dordick JS. 2012. Expanding nature's small molecule diversity via in vitro biosynthetic pathway engineering. Curr Opin Chem Biol 16(1):186–195.

Lee JY, Cheong DE, Kim GJ. 2008. A novel assay system for the measurement of transketolase activity using xylulokinase from *Saccharomyces cerevisiae*. Biotechnol Lett 30(5):899–904.

Martin del Campo JS, Chun Y, Kim JE, Patiño R, Zhang YHP. 2013a. Discovery and characterization of a novel ATP/polyphosphate xylulokinase from a hyperthermophilic bacterium *Thermotoga maritima*. J Ind Microbiol Biotechnol 40(7):661–669.

Martin del Campo JS, Rollin J, Myung S, You C, Chandrayan S, Patino R, Adams MWW, Zhang YHP. 2013b. High-yield production of dihydrogen from xylose by using a synthetic enzyme cascade in a cell-free system. Angew Chem Int Ed 52(17):4587–4590.

Meyer A, Pellaux R, Panke S. 2007. Bioengineering novel in vitro metabolic pathways using synthetic biology. Curr Opin Microbiol 10(3):246–253.

Myung S, Rollin J, You C, Sun F, Chandrayan S, Adams MWW, Zhang Y-HP. 2014. In vitro metabolic engineering of hydrogen production at theoretical yield from sucrose. Metab Eng 24:70–77.

Resnick SM, Zehnder AJB. 2000. In vitro ATP regeneration from polyphosphate and AMP by polyphosphate: AMP phosphotransferase and adenylate kinase from *Acinetobacter johnsonii* 210A. Appl Environ Microbiol 66(5):2045–2051.

Restiawaty E, Iwasa Y, Maya S, Honda K, Omasa T, Hirota R, Kuroda A, Ohtake H. 2011. Feasibility of thermophilic adenosine triphosphate-regeneration system using Thermus thermophilus polyphosphate kinase. Proc Biochem 46(9): 1747–1752.

Rieckenberg F, Ardao I, Rujananon R, Zeng AP. 2014. Cell-free synthesis of 1,3-propanediol from glycerol with a high yield. Eng Life Sci 14(4):380–386.

Rollin JA, del Campo JM, Myung S, Sun F, You C, Bakovic A, Castro R, Chandrayan SK, Wu C-H, Adams MW. 2015. High-yield hydrogen production from biomass by in vitro metabolic engineering: Mixed sugars coutilization and kinetic modeling. Proc Natl Acad Sci USA 112(16):4964–4969.

Sato M, Masuda Y, Kirimura K, Kino K. 2007. Thermostable ATP regeneration system using polyphosphate kinase from *Thermosynechococcus elongatus* BP-1 for D-amino acid dipeptide synthesis. J Biosci Bioeng 103(2):179–184.

Shaeri J, Wright I, Rathbone EB, Wohlgemuth R, Woodley JM. 2008. Characterization of Enzymatic D-Xylulose 5-Phosphate Synthesis. Biotechnol Bioeng 101(4): 761–767.

Shiba T, Tsutsumi K, Ishige K, Noguchi T. 2000. Inorganic polyphosphate and polyphosphate kinase: their novel biological functions and applications. Biochemistry (Moscow) 65(3):315–323.

Shih YS, Whitesides GM. 1977. Large-scale ATP-requiring enzymic phosphorylation of creatine can be driven by enzymic ATP regeneration. J Org Chem 42(25): 4165–4166.

Tessaro D, Pollegioni L, Piubelli L, D'Arrigo P, Servi S. 2015. Systems Biocatalysis: An Artificial Metabolism for Interconversion of Functional Groups. ACS Catal 5(3):1604–1608.

Wood T. 1973. The Preparation Of Xylulose 5-Phosphate Using Transketolase. Prep Biochem 3(6):509–515.

You C, Chen H, Myung S, Sathitsuksanoh N, Ma H, Zhang XZ, Li J, Zhang YHP. 2013. Enzymatic transformation of nonfood biomass to starch. Proc Natl Acad Sci USA 110(18):7182–7187.

You C, Zhang XZ, Zhang YHP. 2012. Simple Cloning via direct transformation of PCR product (DNA multimer) to Escherichia coli and Bacillus subtilis. Appl Environ Microbiol 78(5):1593–1595.

Zhang YHP. 2011. Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol Adv 29(6):715–725.

Zhang YHP. 2015. Production of biofuels and biochemicals by in vitro synthetic biosystems: Opportunities and challenges. Biotechnol Adv DOI: 10.1016/j.biotechadv.2014.10.009.

Zhao HM, van der Donk WA. 2003. Regeneration of cofactors for use in biocatallysis. Curr Opin Biotechnol 14(6):583–589.

Zimmermann FT, Schneider A, Schorken U, Sprenger GA, Fessner WD. 1999. Efficient multi-enzymatic synthesis of D-xylulose 5-phosphate. Tetrahedron: Asymmetry 10(9):1643–1646

Chapter 5. Facile construction of random gene mutagenesis library for directed evolution without the use of restriction enzyme in *Escherichia coli*

Kim, J.-E., Huang, R., Chen, H., You, C. and Zhang, Y.H.P. (2016). Facile construction of random gene mutagenesis library for directed evolution without the use of restriction enzyme in *Escherichia coli*. Biotechnology Journal, 11(9), pp.1142-1150. Reproduced by permission of John Wiley & Sons.

Biotechnology Journal
www.biotechnology-journal.com

Biotechnol. J. 2016, 11

DOI 10.1002/biot.201600121

Biotech Method

Facile construction of random gene mutagenesis library for directed evolution without the use of restriction enzymes in *Escherichia coli*

Jae-Eung Kim[1], Rui Huang[1], Hui Chen[1], Chun You[2] and Y.-H. Percival Zhang[1,2]

[1] Biological Systems Engineering Department, Virginia Tech, Blacksburg, VA, USA
[2] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin Airport Economic Area, Tianjin, China

A foolproof protocol was developed for the construction of mutant DNA library for directed protein evolution. First, a library of linear mutant genes was generated by error-prone PCR or molecular shuffling, and a linear vector backbone was prepared by high-fidelity PCR. Second, the amplified insert and vector fragments were assembled by overlap-extension PCR with a pair of 5'-phosphorylated primers. Third, full-length linear plasmids with phosphorylated 5'-ends were self-ligated with T4 ligase, yielding circular plasmids encoding mutant variants suitable for high-efficiency transformation. Self-made competent *Escherichia coli* BL21(DE3) showed a transformation efficiency of 2.4×10^5 cfu/μg of the self-ligated circular plasmid. Using this method, three mutants of mCherry fluorescent protein were found to alter their colors and fluorescent intensities under visible and UV lights, respectively. Also, one mutant of 6-phosphorogluconate dehydrogenase from the thermophilic bacterium *Moorella thermoacetica* was found to show a 3.5-fold improved catalytic efficiency (k_{cat}/K_m) on NAD^+ as compared to the wild-type. This protocol is DNA-sequence independent, and does not require restriction enzymes, special *E. coli* hosts, or labor-intensive optimization. In addition, this protocol can be used for subcloning the relatively long DNA sequences into any position of plasmids.

Received	03 MAR 2016
Revised	21 JUN 2016
Accepted	27 JUN 2016
Accepted article online	01 JUL 2016

Keywords: Cherry fluorescent protein · Directed evolution · *Moorella thermoacetica* · Mutant library construction · 6-phosphorogluconate dehydrogenase

1 Introduction

Protein engineering becomes a fundamental tool for the development of new biocatalysts possessing necessary properties for biocatalysis, such as thermostability, high catalytic activity, tolerance toward organic solvents and a wide range of pH, high substrate selectivity, and so on. Rational design based on site-directed mutagenesis and directed evolution based on random mutagenesis are two complementary strategies of protein engineering. Rational design selects a few amino acids to change based on detailed knowledge of protein structure, computational model, and catalytic mechanism of the enzyme. The selected amino acids are subjected to change through site-directed mutagenesis. Conversely, directed evolution adopts a random mutagenic process of error-prone PCR (epPCR) and/or molecular shuffling to create a library of genetic mutants [1–3]. High-throughput screening or selection systems are subsequently required to identify mutants with desired phenotypes. It is very important to construct large-size and high-quality mutant libraries for directed evolution [4–6]. Error-prone PCR is the simplest and easiest way to generate a library of DNA variants, and this method has been extensively studied to construct mutant DNA libraries [7, 8].

Correspondence: Prof. Y.-H. Percival Zhang , Biological Systems Engineering Department, Virginia Polytechnic Institute and State University, 210-A Seitz Hall, Blacksburg, VA 24060, USA
Email: ypzhang@vt.edu

Abbreviations: epPCR, error-prone PCR; **OE-PCR**, overlap extension PCR; **POE-PCR**, prolonged overlap extension PCR; **PDLGO**, protein domain library generation by overlap extension

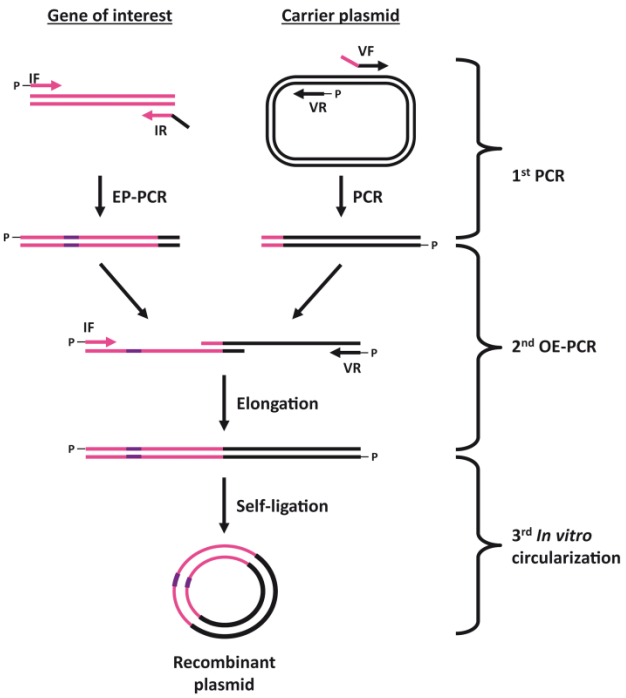

Figure 1. Scheme of the mutant library construction. First, the linear mutant DNA inserts were generated by error-prone PCR, and the linear vector backbone was generated by high-fidelity PCR. The insert and vector fragments were allowed to anneal at their complementary overlapping ends introduced by IR and VF. Second, the two independently PCR-amplified fragments were assembled together by overlap extension PCR (OE-PCR). Third, OE-PCR products with phosphorylated 5'-ends could be self-ligated by T4 ligase. Then, the ligated recombinant plasmids were used to transform competent *E. coli* cells.

The largest technical challenge for this construction is to create high quality mutant libraries in circular plasmid forms suitable for high-efficiency transformation [9, 10]. The typical cloning technique is based on the digestion with two restriction enzymes followed by ligation, where an insert gene amplified by PCR primers introducing restriction enzyme sites and a vector plasmid are digested by the same two restriction enzymes and they are joined together through ligation. This simple technique, however, has several disadvantages, such as limited choices of restricting enzymes, low efficiency of double restriction enzyme digestion, consequentially low recovery rate of DNA mutants, and careful ligation optimization of the insert to vector ratio, resulting in low transformation efficiencies. This method is time-consuming and requires high lab skills to prepare large-size and high-quality mutant libraries for directed evolution.

In order to overcome the deficiencies of double-restriction-enzyme digestion and ligation for mutant library construction, several advanced methods have been developed for directed evolution. A ligation-independent method, Megaprimer PCR of Whole Plasmid, was developed based on the QuikChange method [11]. This method requires a high concentration of methylated parental plasmid as a PCR template and a large amount of *Dpn*I to remove the wild-type DNA template from the final mutant library. Despite of efficient *Dpn*I digestion, remaining-intact wild-type DNA template can be carried over and decrease the quality of final mutant library [12]. The parental plasmid contamination is free in the Kunkel mutagenesis method, and this method has been adopted by Maynard and co-workers to develop a restriction enzyme-independent technique [9, 13]. However, this method involves a special host (dut^+ ung^+ *E. coli*) and its complicated procedure requires high laboratory skills as well. A restriction enzyme- and ligation-independent method was developed for *Bacillus subtilis* because this host prefers assimilating DNA multimers [14]. Later, this PCR-based method was modified to facilitate the generation of large-size circular plasmid libraries for directed evolution in *E. coli* [4]. This method, however, requires a large amount of a restriction enzyme to cut DNA multimers generated by prolonged overlap extension PCR. Various mutagenesis methods have been summarized and compared in a recent review [5].

Here we developed a foolproof protocol for the mutant library construction by a hybrid of error-prone PCR and overlap-extension PCR followed by self-ligation (Fig. 1). First, the insert fragment was prepared from a gene of interest with a 5'-phosphorylated forward primer (IF) and a reverse primer (IR) by epPCR. The vector fragment was prepared by regular high-fidelity PCR from a carrier plasmid with a forward primer (VF) and a 5'-phosphorylated reverse primer (VR). Primers IR and VF shared the complementary sequences added to their 5'-ends such that the right end of the amplified insert fragment is complementary to the left end of the amplified vector fragment. Second, the amplified insert and vector fragments were assembled together by overlap extension PCR (OE-PCR) with a pair of 5'-phosphorylated primers, IF and VR. Third, the OE-PCR product was self-ligated with T4 ligase and transformed into competent *E. coli* cells with high transformation efficiencies. Mutant libraries of *mCherry* gene were constructed by using this protocol for quick validation of this method because mutant mCherry proteins can be easily identified based on their color. This method was further applied to construct mutant libraries of 6-phosphoglucose dehydrogenase (*6pgdh*) gene from a thermophilic bacterium *Moorella thermoacetica*. 6PGDH has played an important role in high-yield hydrogen production [15] and bioelectricity generation [16]. The wild-type 6PGDH shows much higher cofactor preference for $NADP^+$ than NAD^+. It is important to change this enzyme cofactor from NADP to NAD because the latter is less costly and shows higher stability.

2 Materials and methods

2.1 Materials

All chemicals were reagent or higher grade, and purchased from Sigma–Aldrich (St. Louis, MO, USA) or Fisher Scientific (Pittsburgh, PA, USA), unless otherwise noted. All enzymes used for the experiments were purchased from New England Biolabs (Ipswich, MA, USA). The high-fidelity Q5 DNA polymerase was used for regular cloning and OE-PCR, and the NEB regular Taq polymerase was used for epPCR. Primers were purchased from IDT (Coralville, IA, USA). The PCR thermocycler was Eppendorf temperature gradient MasterCycler(Hauppauge, NY, USA).

2.2 Strains and culture conditions

The *E. coli* BL21(DE3) (Invitrogen, Carlsbad, CA, USA) was used for recombinant protein expression. The *E. coli* TOP10 (Invitrogen, Carlsbad, CA, USA) was used for DNA manipulation and plasmid amplification. All microorganisms were grown in Luria-Bertani (LB) medium. 1.5% w/v agar was used for LB-agar plates. The concentration of ampicillin was 100 μg/mL, and the concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG) was 10 μM for overexpression of recombinant proteins on agar plates or 100 μM in liquid cultures. Chemically competent cells of *E. coli* BL21(DE3) and TOP10 were prepared as described elsewhere [17].

2.3 Preparation of plasmid pET20b-*mCherry* and pET28a-P_{tac}-*6pgdh*

The plasmid encoding mCherry fluorescent protein under the control of T7 promoter was prepared from a previously constructed plasmid, pET20b-*mCherry-cbm* [4]. pET20b-*mCherry-cbm* was amplified with a pair of primers (FP: CTCGA GCACC ACCAT CACC; RP: CTTGT ACAGC TCGTC CATG) using the NEB Q5 polymerase to remove the carbohydrate binding module (CBM). The linearized PCR product was treated with the NEB T4 polynucleotide kinase, then ligated by the NEB T4 DNA ligase. The ligated plasmid was transformed into *E. coli* TOP10, yielding plasmid pET20b-*mCherry*.

The plasmid encoding 6-phosphogluconate dehydrogenase (6PGDH) under the control of two cascade promoters, T7 and P_{tac}, was constructed by inserting the *6pgdh* gene amplified from *M. thermoacetica* genomic DNA and P_{tac} into pET28a vector. The insert gene *6pgdh* was amplified with a pair of primers (6PG_F: CTTTA AGAAG GAGAT ATACA TATGC ATATC GGCTT AGTCG GTCTT GGTCG; 6PG_R: ATCTC AGTGG TGGTG GTGGT GGTGT TTCCG GGCCA CGTCG TGGCC). A vector backbone was amplified with a pair of primers (p28_F: CGACC AAGAC CGACT AAGCC GATAT GCATA TGTAT ATCTC CTTCT TAAAG; p28_R: GGCCA CGACG TGGCC CGGAA ACACC ACCAC CACCA CCACT GAGAT) from pET28a. The two PCR products were assembled by prolonged overlap extension PCR (POE-PCR) [18]. The POE-PCR product was transformed to *E. coli* TOP10, yielding pET28a-*6pgdh*. pET28a-*6pgdh* was PCR-amplified with a pair of 5′-end phosphorylated primers (T7_Tac_F: GGCTC GTATA ATGTG TGGAA TTGTG AGCGG ATAAC AATTC; T7_Tac_R: GATGA TTAAT TGTCA ACCTA TAGTG AGTCG TATTA ATTTC G), each of which contains half of P_{tac} promoter sequence. The PCR product was self-ligated by NEB Quick Ligation Kit, then transformed to *E. coli* TOP10, yielding pET28a-P_{tac}-*6pgdh*.

2.4 Preparation of mutant DNA library

Mutant libraries of *mCherry* gene were generated by epPCR from pET20b-*mCherry* as a template DNA with a pair of primer (Mmcherry-IF: ACATA TGGTG AGCAA GGGCG AG; Mmcherry-IR: GTGAT GGTGG TGCTC GAGCT TGTAC AGCTC GTCCA TGC). Mmcherry-IF was ordered as a 5′-end phosphorylated primer. The last 30 bp of the 5′ terminus of Mmcherry-IR (underlined) were designed to introduce complementary sequences to the vector fragment. The epPCR was carried out under three reaction conditions with different concentrations of 50, 100 and 200 μM of $MnCl_2$. The reaction solution with a total volume of 50 μl contained 3 ng/μl plasmid pET20b-*mCherry*, 0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP, 1 mM dTTP, 5 mM $MgCl_2$, various amounts of $MnCl_2$, 0.4 μM primers (each Mmcherry-IF and Mmcherry-IR), and 0.05 U/μl NEB Taq polymerase [14]. The epPCR conditions were initial denaturation (120 s at 94°C), 18 cycles of denaturation (30 s at 94°C), annealing (30 s at 56°C), and extension (42 s at 68°C), and a final extension step (5 min at 68°C). The linear vector fragment was amplified from pET20b-*mCherry* by using the NEB high-fidelity Q5 DNA polymerase with a pair of primers (Mmcherry-VF: GAGCT GTACA AGCTC GAGCA CCACC ATCAC CATCA CTGAG; Mmcherry-VR: ATATC TCCTT CTTAA AGTTA AAC). The first 30 bp of the 5′ terminus of Mmcherry-VF (underlined) were designed to introduce complementary sequences to the insert fragment. Mmcherry-VR was ordered as a 5′-end phosphorylated primer. The reaction solution with a total volume of 50 μl contained pET20b-*mCherry* (0.02 ng/μl), dNTPs (0.2 mM each), primers (0.5 μM each), and the Q5 DNA polymerase (0.02 U/μl). The PCR was conducted as follows: initial denaturation (60 s at 98°C), 18 cycles of denaturation (20 s at 98°C), annealing (20 s at 60°C), and extension (72 s at 72°C), and a final extension step (5 min at 72°C). The PCR products were recovered from 0.8% w/v agarose gel by using a Zymoclean Gel DNA Recovery Kit (Irvine, CA, USA).

Mutant libraries of *6pgdh* gene were generated by epPCR from pET28a-P_{tac}-*6pgdh* as a template DNA with a

pair of primers (M6pgdh-IF: CAATT CCCCT CTAGA AATAA TTTTG TTTAA CTTTA AGAAG GAGAT ATCAT ATG; M6pgdh-IR: ATAAT GGCCG CTGCC GGGCG GCCCG GTATG GAGAT AACCT CCGGG TACGG T) following the same procedures to construct mutant libraries of mCherry protein as described above. M6pgdh-IR (51 bp) was designed to introduce complementary sequences to the vector fragment. The epPCR for 6PGDH libraries was carried out under two reaction conditions with different concentrations of 100 and 200 μM of $MnCl_2$. All other epPCR conditions for 6PGDH libraries were the same as described above, except for the extension time (34 s at 68°C). The linear vector fragment was amplified from pET28a-P_{tac}-*6pgdh* by using the NEB high-fidelity Q5 DNA polymerase with a pair of primers (M6pgdh-VF: ACCGT ACCCG GAGGT TATCT CCATA CCGGG CCGCC CGGCA GCGGC CATTA T; M6pgdh-VR: TTATC CGCTC ACAAT TCCAC ACATT ATACG AGCCG ATG). M6pgdh-VF (51 bp) was designed to introduce complementary sequences to the insert fragment. The PCR conditions using high-fidelity Q5 DNA polymerase for 6PGDH libraries were the same as described above, except for the extension time (113 s at 72°C). The PCR products were recovered from 0.8% w/v agarose gel by using the Zymoclean Gel DNA Recovery Kit.

The gel-purified insert and vector fragments were spliced through OE-PCR. The reaction solution with a total volume of 50 μl contained dNTPs (0.2 mM each), 0.5 μM primers (each Mmcherry-IF and Mmcherry-VR for mCherry protein library or each M6pgdh-IF and M6pgdh-VR for 6PGDH library), the insert fragment (10 ng/μl), the vector fragment (equimolar with the insert fragment), and the Q5 DNA polymerase (0.02 U/μl). OE-PCR was conducted as followings: initial denaturation (60 s at 98°C), 15 cycles of denaturation (20 s at 98°C), annealing (20 s at 60°C), and extension (86 s at 72°C for mCherry protein or 124 s at 72°C for 6PGDH), and a final extension step (5 min at 72°C). The OE-PCR product was recovered from 0.8% w/v agarose gel by using the Zymoclean Gel DNA Recovery Kit. The purified linear plasmid DNAs were self-ligated by using the NEB Quick Ligation Kit at 25°C for 5 min. The ligation product was transformed into the competent *E. coli* BL21(DE3). The transformation process was conducted as followings. The competent *E. coli* BL21 in 1.5-mL regular centrifuge tube was kept on ice for 10 min after completely thawed. The ligation product (e.g. 1–5 ng) was added to 100 μl of the competent cells and then gently mixed. After stored on ice for another 30 min, the competent cell tubes were heat treated to 42°C for 90 s, and then immediately transferred to an ice bath to cool down for 5 min. Then, 900 μl of the LB liquid media was added to the tubes for recovering cultivation at 37°C for 45 min. After centrifugation at 3000 ×g for 3 min, the supernatant was discarded. The cell pellets were re-suspended with the remaining LB liquid media. Re-suspended cells were spread on the LB-agar plates containing 100 μg/mL ampicillin and 10 μM IPTG for mCherry protein or 50 μg/mL kanamycin for 6PGDH. The plates were incubated at 37°C overnight or until colonies appeared.

2.5 Protein overexpression and purification

The *E. coli* BL21(DE3) cells harboring the protein expression plasmids were grown at 37°C in 250 mL Erlenmeyer flasks containing 150 mL of LB medium with appropriate antibiotics. The protein expression was induced with 0.1 mM IPTG when the culture A_{600} reached about 0.8, and then cultivated at 18°C for 16 h. The *E. coli* cells were harvested by centrifugation and washed with 50 mM of HEPES buffer (pH 7.2) containing 0.5 M NaCl. The cell pellets were re-suspended in the same buffer and lysed by sonication in an ice bath (Fisher Scientific sonic dismembrator Model 500; 3 s pulse on and 6 s off, total 180 s at 50% amplitude). After sonication, the cell lysate was centrifuged to isolate the supernatant containing soluble target proteins from the cell lysate. His-tagged target proteins in the supernatant were purified using the ThermoFisher HisPur Ni-NTA Agarose Resin. The target protein was eluted with 100 mM HEPES buffer (pH 7.5) containing 500 mM imidazole after several washing steps with 100 mM HEPES buffer (pH 7.5) containing 50 mM imidazole.

2.6 Screening of 6PGDH mutants and activity assays

The double-layer screening system was used to select positive 6PGDH mutants with increased activity on NAD^+. The *E. coli* TOP10 cells transformed with mutant libraries of 6PGDH were grown on the kanamycin-containing LB-agar plates at 37°C overnight. The colonies on the plates were heat-treated at 60°C for 1 h to deactivate *E. coli* cellular enzymes, degrade metabolites and cofactors, and render the cell membrane semi-permeable. 10 mL of 0.5% agarose solution containing 50 mM Tris-HCl (pH 7.5), 50 μM tetranitroblue tetrazolium, 10 μM phenazine methosulfate, 2 mM 6-phosphogluconate, and 1 mM NAD^+ was applied to the heat-treated colonies on the agar plates. After incubation at room temperature for 3–4 h, colonies appearing with a halo of dark brown color around them were picked for colony PCR to amplify mutant *6pgdh* genes. Amplified genes were re-inserted back to pET28a-P_{tac} vector by using restriction enzyme digestion and ligation. After transformation to *E. coli* TOP10 for plasmid amplification, plasmids were extracted and sent for DNA sequencing.

The activity assay of 6PGDH was done as described elsewhere [16]. Briefly, activities of 6PGDH were measured in 100 mM HEPES buffer (pH 7.5) with 2 mM 6-phosphogluconate, 2 mM $NAD(P)^+$, 5 mM $MgCl_2$ and 0.5 mM $MnCl_2$, 1 μg/mL 6PGDH at 50°C for 5 min. The formation

of NAD(P)H was measured at 340 nm by Beckman Coulter UV/Vis spectrophotometer (Fullerton, CA, USA). The K_m values of 6PGDH on NAD(P)$^+$ were measured under the same conditions of the activity assay with NAD(P)$^+$ concentrations from 5 µM to 5 mM. Regression equations were derived by using GraphPad Prism 5 from Graphpad Software Inc.(La Jolla, CA, USA).

3 Results

3.1 Construction and transformation of pET20b-*mCherry* to *E. coli*

Wild-type mCherry protein was used for a proof-of-concept experiment for validating this method. Plasmid pET20b-*mCherry* constructed from pET20b-*mCherry-cbm* by removing CBM was used as a template to amplify the insert and vector fragment. These two independently PCR-amplified fragments were assembled together by OE-PCR and the full-length DNA fragments were self-ligated to yield circular plasmids (Fig. 1). For the high-transformation efficiency experiment, the PCR amplification of both insert and vector fragments was conducted by using a high-fidelity DNA polymerase. The insert and vector fragments had an overlapping region (30 bp) at the other side of each fragment introduced by a pair of primers (IR and VF) during OE-PCR (Fig. 2A). The insert (730 bp) and vector (3598 bp) fragments were generated as shown in lanes 1 and 2 of Fig. 2B, respectively. The amplified insert and vector fragments were assembled together to generate a full-length linear plasmid (4298 bp) by OE-PCR (Fig. 2B, lane 3). Primers, IF and VR, used for this method were ordered as 5'-end phosphorylated primers. Thus, the linear plasmids after OE-PCR resulted to have phosphorylated 5'-ends, facilitating self-ligation with T4 ligase.

After OE-PCR followed by in vitro circularization, the self-ligated pET20b-*mCherry* encoding wild-type mCherry protein under the control of a T7 promoter was transformed into competent *E. coli* BL21 cells. Transformed cells were grown on LB-agar plates containing 10 µM of IPTG at 37°C. Almost all colonies that appeared after overnight growth exhibited a cherry color, suggesting that the method can be used to construct a large number of colonies (Fig. 2C). No colonies appeared from the transformation of competent *E. coli* BL21 with only linear vector fragments. The transformation efficiency of fleshly-made competent BL21(DE3) with an intact plasmid was 4.0×10^6 cfu/µg DNA, while the transformation efficiency with ligated plasmids resulted in 2.4×10^5 cfu/µg DNA. Plasmids from three randomly selected colonies showed the wild-type mCherry sequence and no errors in overlap regions validated by DNA sequencing.

3.2 Three mutant mCherry fluorescent proteins exhibit different colors

In order to generate mutant libraries of mCherry fluorescent protein, the insert fragment amplified from pET20b-*mCherry* was conducted by epPCR with various concentrations of Mn^{2+} to adjust mutation rates. The vector fragment was amplified by a high-fidelity DNA polymerase. After OE-PCR followed by in vitro circularization, circular plasmids encoding variants of mutant *mCherry* gene were transformed into competent *E. coli* BL21(DE3) cells for library expression and mutant analysis. Three different concentrations of Mn^{2+} (50, 100 and 200 µM) used for epPCR to prepare three mutant libraries of mCherry resulted in different mutation rates by showing that high Mn^{2+} levels led to low percentages of cherry colonies on plates (Fig. 3A–C). Among several mutants exhibiting different colors, three easily-identified mutants

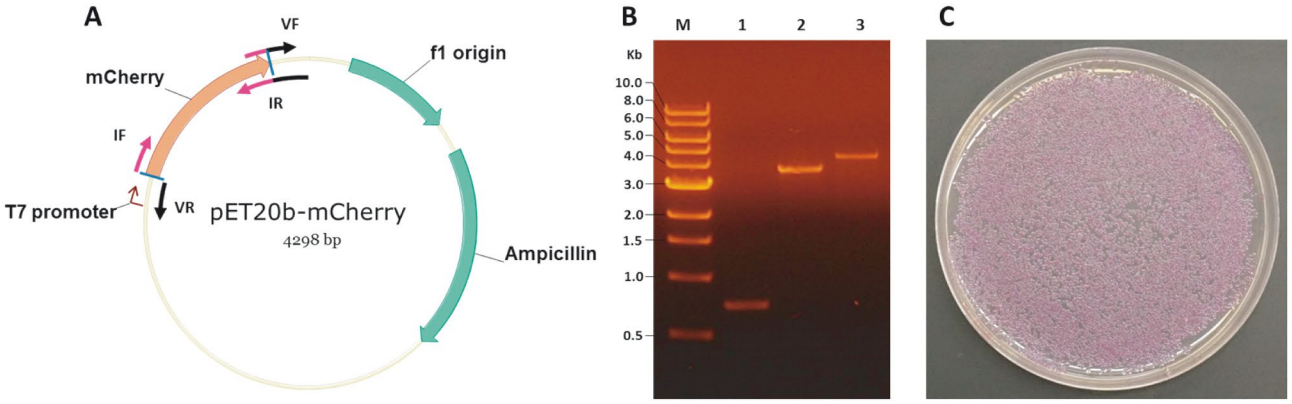

Figure 2. Construction of pET20b-*mCherry* plasmid. Plasmid map of pET20b-*mCherry* with primers (arrows) used for PCR amplification in the transcription directions (**A**). T7 promoter is denoted as an arrow symbol, and located upstream of mCherry DNA sequence. DNA analysis by 0.8% w/v agarose gel electrophoresis (**B**). Lane M: NEB 1 kb DNA ladder; lane 1: *mCherry* DNA amplified by regular PCR; lane 2: linear vector amplified by regular PCR; lane 3: full-length linear plasmid generated by OE-PCR. Self-made competent *E. coli* BL21(DE3) cells transformed with the self-ligated circular pET20b-*mCherry* generated via this method (**C**).

were selected for DNA sequencing. Two mutants, M1 and M2, were selected from the mutant library constructed with 100 μM Mn^{2+}, and one mutant, M3, was selected from the mutant library constructed with 200 μM Mn^{2+} (Fig. 3B and 3C). Each mutant had one active mutation, resulting in one amino acid substitution (M1: M71T; M2: M71L; M3: I202T). M1 had two more base changes, but they were silent mutations. The selected mutants and wild-type mCherry protein were overexpressed and purified for comparison under natural visible light and ultraviolet excitation at 302 nm (Fig. 3D and 3E). M1 and M2 exhibited pink and yellow colors, respectively, in visible light, and shared the same site of the amino acid for substitution. This result suggested that the amino acid site of 71 is essential to mCherry protein for its original color, and active mutations at this site can be effective to find new fluorescent proteins. Under UV excitation at 302 nm, M1 remained strongly fluorescent, while M2 and M3 lost their fluorescent signals greatly.

3.3 Improved catalytic efficiency on NAD^+ of 6PGDH

Plasmid pET28a-P_{tac}-*6pgdh* was used as a template to amplify the insert and vector fragment. These two independently PCR-amplified fragments were assembled by OE-PCR and self-ligated to yield circular plasmids. The insert and vector fragments had an overlapping region (51 bp) at the other side of each fragment introduced by a pair of primers (IR and VF) during OE-PCR (Fig. 4A). The insert (557 bp) and vector (5647 bp) fragments were generated as shown in lanes 1 and 2 of Fig. 4B, respectively. The amplified insert and vector fragments were assembled together to generate a full-length linear plasmid (6153 bp) by OE-PCR (Fig. 4B, lane 3). Amplification of the insert fragment from pET28a-P_{tac}-*6pgdh* was conducted by epPCR under 100 μM Mn^{2+} concentration.

Circular plasmids encoding variants of mutant *6pgdh* gene were transformed into competent *E. coli* TOP10 cells. Screening mutant 6PGDH with the desired pheno-

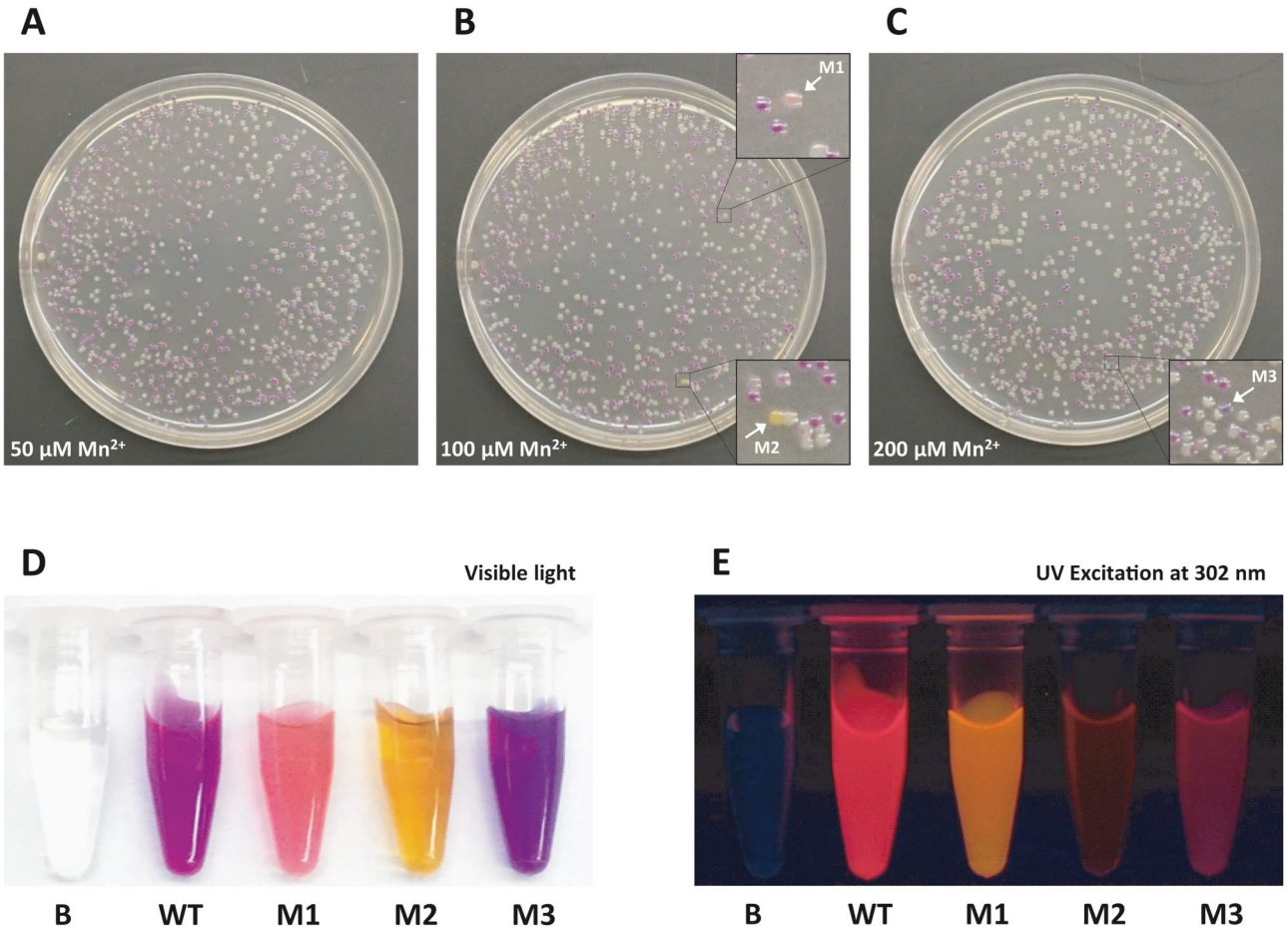

Figure 3. Screening of mCherry fluorescent mutants expressed in *E. coli* BL21(DE3) cells grown on LB-agar plates, and overexpression and purification of mCherry proteins. Three different Mn^{2+} concentrations, 50 μM Mn^{2+} (**A**), 100 μM Mn^{2+} (**B**), and 200 μM Mn^{2+} (**C**), were used to construct mutant libraries of *mCherry* gene. Three easily-identified mutant, M1–M3, were selected for overexpression and purification. Purified proteins in the same concentration were exposed to visible light (**D**) and UV excitation at 302 nm (**E**). B, blank water; WT, wild-type mCherry protein; M1–M3, mCherry mutants M1, M2, and M3.

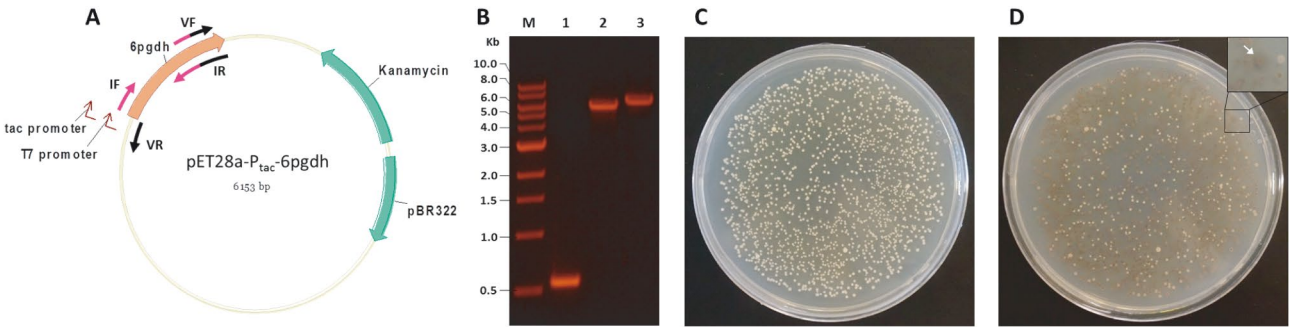

Figure 4. Construction of pET28a-P$_{tac}$-*6pgdh* plasmid and screening mutant library of *6pgdh* gene. Plasmid map of pET28a-P$_{tac}$-*6pgdh* with primers (arrows) used for PCR amplification in the transcription directions (**A**). T7 and P$_{tac}$ promoter are denoted as arrow symbols, and located upstream of mCherry DNA sequence. DNA analysis by 0.8% (w/v) agarose gel electrophoresis (**B**). Lane M: NEB 1 kb DNA ladder; lane 1: *6pgdh* DNA amplified by epPCR; lane 2: linear vector amplified by regular PCR; lane 3: full-length linear plasmid generated by OE-PCR. The mutant library of the *6pgdh* gene was screened by a double-layer screening system after transformation into *E. coli* TOP10. An initial screening plate at 0 h (**C**) and after 3–4 h (**D**). The most-easily identified mutant (indicated by an white arrow in (**D**) was selected for DNA recovery and further analysis.

type of enhanced activity on NAD^+ was performed by a double-layer screening system. Reduced NADH generated by 6PGDH mutants can reduce colorless oxidized TNBT into a reduced form with dark brown color. No color changed immediately after applying the second layer (Fig. 4C). It took approximately 3–4 h to develop dark brown color haloes around colonies (Fig. 4D). One mutant 6PGDH (indicated by an arrow in Fig. 4D) was identified. This mutant had one amino acid substitution (T32H). The apparent K_m and k_{cat} values of this mutant on NAD^+ were 1172 μM and 36.8 s^{-1}, respectively, resulting in 3.5-fold improvement of the catalytic efficiency (k_{cat}/K_m) on NAD^+ compared to the wild-type 6PGDH, whose apparent K_m and k_{cat} values on NAD^+ were 1397 μM and 12.6 s^{-1}, respectively.

4 Discussion

We developed a foolproof protocol to construct large-size DNA mutant libraries for directed evolution in *E. coli*. The insert fragment prepared by epPCR and the vector fragment prepared by regular PCR were assembled together by OE-PCR. The full-length DNA fragments were self-ligated to yield high-quality mutant libraries in circular plasmid forms suitable for high-efficiency transformation. From two cases demonstrated here, two different overlap regions (30 bp for mCherry protein and 51 bp for 6PGDH) were designed and tested, and their GC-content percentage were 57 and 65%, respectively. The PCR results of both cases were successful with annealing temperature at 60°C. The more significant quality than the length of overlap region and %GC was that of primers introducing complementary sequences during OE-PCR. It is highly recommended to order highly purified error-free primers for OE-PCR. OE-PCR has been widely adopted for many cloning, site-directed mutagenesis techniques, and DNA recombination methods [19–22]. In a hybrid with epPCR, the OE-PCR was adopted to develop a method, called "protein domain library generation by overlap extension" (PDLGO), generating random mutations in a confined region within a target gene sequence [23]. PDLGO still uses restriction sites to clone a chimerical sequence library into an expression vector, inheriting the disadvantages of the traditional cloning technique demanding restriction enzyme digestion and ligation.

Compared to previously developed cloning techniques for directed evolution, our method described here has several advantages. First, our method is a restriction enzyme-free random mutant library construction method without any restraints of DNA sequences for directed evolution in *E. coli*. The typical cloning technique requires two unique restriction enzyme sites for cloning, which consequently introduces extra amino acids to a target protein with possibility to influence the properties of the target protein. The method developed from the Kunkel mutagenesis is also independent of the use of restriction enzymes, but it requires a special host [9]. The method based on DNA multimers [4] can utilize any single-cutting site of the restriction enzyme available in the plasmid, but it requires a large amount of the restriction enzyme (e.g. >200 U/50 μl of DNA multimer product) to generate a linear mutant DNA library. Second, this new method is easy to perform and optimize overall procedures of mutant library construction. Compared to the method based on DNA multimers [4], which requires two-side overlap extension PCR, this new method involves one-side overlap extension PCR. The final products of OE-PCR are one-fragment DNA products with 5′-phosphorylated ends. Thus this new method does not require ligation optimization to achieve a high transformation efficiency, whereas the typical cloning technique involved time- and labor-intensive work to optimize the ratio of the insert to vector fragment for ligation [24]. Third, this new method can decrease the background of mutant library without undesirable DNA molecules, such as wild-type plasmid

and empty plasmid backbone. Only the full-length linear plasmids can be self-ligated to form circular plasmids, which can be transformed into competent *E. coli* cells for the target protein expression.

This new method consists of two sequential PCRs, epPCR and OE-PCR, followed by self-ligation step. Error-prone PCR has been extensively studied to accomplish high-quality non-biased random mutagenesis, and many different random mutagenesis methods were compared for their mutational spectra [8]. There are several factors that affect the quality of mutant libraries, including different DNA polymerases, compositions of dNTPs, concentrations of Mg^{2+} and Mn^{2+}, different target genes, concentrations of template DNA used for PCR, and PCR cycle numbers. Exploring various random mutagenesis conditions is often necessary to find positive mutants with desired phenotypes [25, 26]. For example, the mutant mCherry protein exhibiting yellow color was only appearing from the mutant library constructed with 100 μM Mn^{2+} (Fig. 3B). On the other hand, it has been reported that OE-PCR potentially introduces undesired random errors, because any disincorporation of nucleotides in the early rounds of amplification can be propagated into subsequent rounds of amplification by representing a significant portion of the final product [27, 28]. Therefore, this method is prone to undesired mutation possibly introduced during two sequential PCR steps, one of which is OE-PCR. In order to minimize such potential errors by reducing amplification ratio, high DNA template concentrations of insert and vector fragments along with low cycle numbers during OE-PCR are recommended. Nearly all colonies from the self-ligated pET20b-*mCherry* encoding wild-type mCherry protein exhibited a cherry color with less than 0.5% colonies in a light cherry color (Fig. 2C). These colonies in the light cherry color were possibly due to undesired errors introduced in the plasmid backbone over sequential PCR steps. The ratio of colonies in the light cherry color approximately accorded with an estimated percentage error of PCR products by the PCR Fidelity Calculator (https://www.thermofisher.com).

Among the three selected mutant mCherry proteins, the most color-changed mutants, M1 and M2, shared the same site of substituted amino acid. The substitutions, M71T from mutant M1 and M71L from M2, yielded pink and yellow colors, respectively, in visible light (Fig. 3D). This result suggested that Met71 is an important position for the diverse colors of many fluorescent proteins. Yellow-to-red fluorescent proteins have been studied to increase the speed of maturation and improve their quantum efficiency and photostability [29]. This study revealed that Gln66 of mRFP1, homologous to Met71 of mCherry, plays an important role as substitutions Q66T and Q66C resulted in blue-shifted mutants with respected to mRFP1. A mutant fluorescent protein, mOrange derived from DsRed by Shaner et al. [29], has the same amino acid substitution at Met71 site to Thr71 as M1 has, exhibiting similar colors under both visible light and UV excitation.

A mutant 6PGDH (T32H) showed the 3.5-fold improved catalytic efficiency (k_{cat}/K_m) on NAD^+, while the 1.1-fold lower catalytic efficiency on $NADP^+$ as compared to the wild-type 6PGDH. The single amino acid substitution, T32H, possibly decreased the binding affinity of 6PGDH with $NADP^+$ by breaking hydrogen bonds between the enzyme and the cofactor, resulting in the decreased catalytic efficiency on $NADP^+$. Similar threonine replacement for sheep liver 6PGDH also exhibited decreased catalytic efficiency on $NADP^+$ [30].

In conclusion, our method is foolproof because it is easy to perform and optimize overall procedure for the construction of high-quality mutant DNA library for directed evolution. Also, this method can be used as a subcloning technique to insert any DNA sequence to any position of the plasmid.

This project cannot be carried out without the support of the Biological System Engineering Department, Virginia Polytechnic Institute and State University, Virginia, USA. Funding to YPZ for this work was partially supported by DOE EERE award (DE-EE0006968)and the Virginia Agricultural Experiment Station and the Hatch Program of the National Institute of Food and Agriculture, US Department of Agriculture.

The authors declare no financial or commercial conflict of interest.

5 References

[1] Reetz, M. T., Wilensek, S., Zha, D., Jaeger, K. E., Directed evolution of an enantioselective enzyme through combinatorial multiple-cassette mutagenesis. *Angew. Chem. Int. Ed.* 2001, *40*, 3589–3591.

[2] Hossain, G. S., Shin, H.-d., Li, J., Wang, M. et al., Integrating error-prone PCR and DNA shuffling as an effective molecular evolution strategy for the production of α-ketoglutaric acid by L-amino acid deaminase. *RSC Adv.* 2016, *6*, 46149–46158.

[3] Cirino, P. C., Qian, S., Protein engineering as an enabling tool for synthetic biology, in: Zhao, H. (Ed.), *Synth. Biol.*, Elsevier 2013, pp. 23–42.

[4] You, C., Zhang, Y. H. P., Easy preparation of a large-size random gene mutagenesis library in *Escherichia coli*. *Anal. Biochem.* 2012, *428*, 7–12.

[5] Ruff, A. J., Dennig, A., Schwaneberg, U., To get what we aim for – progress in diversity generation methods. *FEBS J.* 2013, *280*, 2961–2978.

[6] Reetz, M. T., Carballeira, J. D., Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* 2007, *2*, 891–903.

[7] Cirino, P. C., Mayer, K. M., Umeno, D., Generating mutant libraries using error-prone PCR, in: Arnold, F. H., Georgiou, G., *Directed Evolution Library Creation*, Springer 2003, pp. 3–9.

[8] Wong, T. S., Roccatano, D., Zacharias, M., Schwaneberg, U., A statistical analysis of random mutagenesis methods used for directed protein evolution. *J. Mol. Biol.* 2006, *355*, 858–871.
[9] Pai, J. C., Entzminger, K. C., Maynard, J. A., Restriction enzyme-free construction of random gene mutagenesis libraries in *Escherichia coli*. *Anal. Biochem.* 2012, *421*, 640–648.
[10] Koyanagi, T., Yoshida, E., Minami, H., Katayama, T., Kumagai, H., A rapid, simple, and effective method of constructing a randomly mutagenized plasmid library free from ligation. *Biosci. Biotechnol. Biochem.* 2008, *72*, 1134–1137.
[11] Miyazaki, K., Takenouchi, M., Creating random mutagenesis libraries using megaprimer PCR of whole plasmid. *Biotechniques* 2002, *33*, 1033–1034.
[12] Tobias, A. V., Preparing libraries in *Escherichia coli*, in: Arnold, F. H., Georgiou, G., *Directed Evolution Library Creation*, Springer 2003, pp. 11–16.
[13] Kunkel, T. A., Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 1985, *82*, 488–492.
[14] Zhang, X. Z., Zhang, Y. H. P., Simple, fast and high-efficiency transformation system for directed evolution of cellulase in *Bacillus subtilis*. *Microb. Biotechnol.* 2011, *4*, 98–105.
[15] Rollin, J. A., del Campo, J. M., Myung, S., Sun, F. et al., High-yield hydrogen production from biomass by in vitro metabolic engineering: Mixed sugars coutilization and kinetic modeling. *Proc. Natl. Acad. Sci. USA* 2015, *112*, 4964–4969.
[16] Zhu, Z. G., Tam, T. K., Sun, F. F., You, C., Zhang, Y.-H. P., A high-energy-density sugar biobattery based on a synthetic enzymatic pathway. *Nat. Commun.* 2014, *5*, 3026.
[17] Sambrook, J., Russell, D. W., The Inoue method for preparation and transformation of competent *E. coli*: "Ultra competent" cells. *CSH Protoc.* 2006, *2006*, pdb.prot3944.
[18] You, C., Zhang, X.-Z., Zhang, Y.-H. P., Simple cloning via direct transformation of PCR product (DNA multimer) to *Escherichia coli* and *Bacillus subtilis*. *Appl. Environ. Microbiol.* 2012, *78*, 1593–1595.
[19] Bryksin, A. V., Matsumura, I., Overlap extension PCR cloning: A simple and reliable way to create recombinant plasmids. *Biotechniques* 2010, *48*, 463.
[20] Heckman, K. L., Pease, L. R., Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* 2007, *2*, 924–932.
[21] Kolkman, J. A., Stemmer, W. P., Directed evolution of proteins by exon shuffling. *Nat. Biotechnol.* 2001, *19*, 423–428.
[22] An, Y., Ji, J., Wu, W., Lv, A. et al., A rapid and efficient method for multiple-site mutagenesis with a modified overlap extension PCR. *Appl. Microbiol. Biotechnol.* 2005, *68*, 774–778.
[23] Gratz, A., Jose, J., Protein domain library generation by overlap extension (PDLGO): A tool for enzyme engineering. *Anal. Biochem.* 2008, *378*, 171–176.
[24] Arnold, F. H., Georgiou, G., Directed evolution library creation. *Methods Mol. Biol.* 2003, *231*.
[25] Drummond, D. A., Iverson, B. L., Georgiou, G., Arnold, F. H., Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J. Mol. Biol.* 2005, *350*, 806–816.
[26] Romero, P. A., Arnold, F. H., Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* 2009, *10*, 866–876.
[27] Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., Pease, L. R., Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 1989, *77*, 51–59.
[28] Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A. et al., A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 7877–7882.
[29] Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. et al., Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* 2004, *22*, 1567–1572.
[30] Li, L., Cook, P. F., The 2'-phosphate of NADP is responsible for proper orientation of the nicotinamide ring in the oxidative decarboxylation reaction catalyzed by sheep liver 6-phosphogluconate dehydrogenase. *J. Biol. Chem.* 2006, *281*, 36803–36810.

Chapter 6. Development of the first monosaccharide 4-epimerase for a zero-calorie sweetener, L-arabinose, production from D-xylose

In preparation

Abstract

More and more evidences suggest that over-consumption of sucrose and high fructose corn syrup is the ultimate cause of high blood pressure, high cholesterol, heart disease, diabetes, and liver disease while fat and sodium may be the lesser of the evils in out diet. The cost-competitive production of the fourth generation sweeteners featuring perfect sweet taste, low calorie, and great safety is highly desired. L-arabinose is a natural zero-calorie sweetener that can neutralize sucrose assimilation and promote growth of beneficial microorganisms in the small intestine as a prebiotic. Here we design an in vitro three-enzyme pathway that converts biomass sugar D-xylose to zero calorie sweetener L-arabinose. To bridge D-xylulose, made from D-xylose catalyzed by xylose isomerase, with L-ribulose, we mined promiscuous activities of D-xylulose 4-epimerase (Xu4E). Tagaturonate-fructuronate epimerase (UxaE) from *Thermotoga maritima* was found to have the highest Xu4E activity among 11 different epimerases we examined for the promiscuous activity on D-xylulose. Furthermore, we developed a sensitivity-adjustable genetic sensor-based high-throughput screening method, which can detect intracellular L-arabinose concentrations and converts into an observable phenotype accordingly. The activity of Xu4E was enhanced by more than 25-fold via three rounds of directed evolution. This enzyme is not only the first monosaccharide 4-epimerase, but also the first enzyme accomplishing D- to L-interconversion of monosaccharides. The production of high-value L-arabinose would bring great economic benefits to biorefineries, analogue to plastics from oil refineries.

Introduction

Sucrose has been one of the most important commodities for thousands of years due to following reasons: human instinctive appetite for sweet foods and drinks; the complementarity that sucrose brings to the other flavors in food; its preservation and fermentation properties; and the calories it provides. The global market for all sweeteners was more than $70 billion in 2012, and the sucrose industry accounted for over 150 million tonnes (Philippe et al., 2014). Now the integrated consortia of plantations and sweetener industry become one of the most influential components of the food industry. In need of an alternative less costly sweetener, high fructose corn syrup (HFCS) was developed from 1970s and up to 25 million tonnes of HFCS per year is consumed globally. Over past decades, a dramatic increase in the amount of added sucrose and HFCS is found to be correlated with increasing number of people in diabetes, obesity, cancer, and cardiovascular disease (Goran et al., 2013; Malik et al., 2006). Recent reports exposed that the sucrose industry sponsored the misleading studies favoring its interests, scapegoating such negative correlation due to more consumption of lipids (Kearns et al., 2016; Nestle, 2016). More and more evidences suggest that the epidemic obesity and metabolic syndromes are mainly driven by sucrose and fructose over-consumption with its unique metabolic fate in liver (Goran et al., 2013; Lanaspa et al., 2013; Lyssiotis and Cantley, 2013). The World Health Organization (WHO) and Food and Agricultural Organization (FAO) recommend low-sugar dietary guidelines for preventing diet-related chronic diseases. The health-related issues with sugar and caloric sweeteners initiated intensive efforts to find non-nutritive sweeteners, such as stevia, aspartame, and saccharin. Many of these new products, however, have concerns in unpleasant aftertaste, costs, and long-term safety (Burbank and Fraumeni, 1970; Kroger et al., 2006; Soffritti et al., 2006; Suez et al., 2014).

L-arabinose is an US Food and Drug Administration (FDA)-approved zero-calorie natural sweetener with 50% sweetness of sucrose and the similar taste (Hao et al., 2015). In addition to the non-caloric sweetener, L-arabinose is regarded as a sucrose neutralizer because it is an inhibitor of intestinal sucrase and suppresses sucrose uptake (Krog-Mikkelsen et al., 2011; Seri et al., 1996). Numerous studies indicate that the addition of 3-4% L-arabinose into sucrose can drastically decrease the uptake of sucrose and consequently decrease lipogenesis (Osaki et al., 2001). Another health benefit of L-arabinose consumption is that the unutilized sucrose with L-arabinose has potential prebiotic properties that promote the growth of beneficial gut microbes (Gibson and Roberfroid, 1995). Currently, L-arabinose is produced as a co-product of xylitol made from the acid hydrolysis of corn cobs. The low percentage contents for L-arabinose in biomass (e.g., typically 2-3%) and limited supplies of high L-arabinose-rich feedstock result in its high production costs ($~30/kg). Currently, a blend of sucrose and L-arabinose is commercially available in Japan, Korea and China, but its high price prevents its wide acceptance in the market.

D-xylose, the most abundant pentose in nature, is a major component of the hydrolysate of lignocellulose. Because it is not economically feasible to separate D-xylose from glucose and other small amounts of rare sugars, such as arabinose, galactose, and mannose, many engineered microorganisms have been developed to co-utilize C5 and C6 sugars for the production of biochemicals and biofuels, such as ethanol (Alper and Stephanopoulos, 2009; Ha et al., 2011; Zhang et al., 2015). High processing costs and low selling prices of cellulosic ethanol imposed great challenges on the economic viability of second generation biorefineries. It has been a long

sought to identify high-value-added biochemicals with a large market size that can be co-produced from biomass in biorefinery industry. Such co-products with a great market potential could generate more revenues, subsidizing biofuel production (Zhang, 2011).

Epimerases (EC 5.1.3.X) are among isomerase enzymes that catalyze a stereochemical inversion of configuration at a single stereocenter of substrates with several centers of asymmetry. Monosaccharide epimerases can catalyze the epimerization occurring at C1 position (EC 5.1.3.3: aldose 1-epimerase), C2 position (EC 5.1.3.8: N-acylglucosamine 2-epimerase), and C3 position (EC 5.1.3.31: D-tagatose 3-epimerase), but enzymes that catalyze 4-epimerization on monosaccharides are unknown in nature. Natural 4-epimerases require the presence of phosphate or uridine diphosphate (UDP) group for substrate binding: L-ribulose 5-phosphate 4-epimerase (EC 5.1.3.4); UDP-glucose 4-epimerase (EC 5.1.3.2); UDP-D-xylose 4-epimerase (EC 5.1.3.5). Creation and development of the first monosaccharide 4-epimerase are expected to bring a high potential for cost-competitive production of L-arabinose and other rare sugars (Beerens et al., 2017).

In this study we mined numerous 4-epimerase enzymes from four EC enzymes for discovering their promiscuous activities on D-xylulose as a substrate, then named it as D-xylulose 4-epimerase (Xu4E) (**Figure 1a**). Directed evolution was applied to improve the specific activity of Xu4E by developing a high-throughput screening (HTS) method based on a transcription-factor-driven biosensor that detects intracellular L-arabinose levels and converts into an easily discernable phenotype such as a fluorescent color in different intensities. A novel three-enzyme

pathway comprised of xylose isomerase (XI), Xu4E, and L-arabinose isomerase (AI) was constructed to enable the upgrade of D-xylose to L-arabinose (**Figure 1b**).

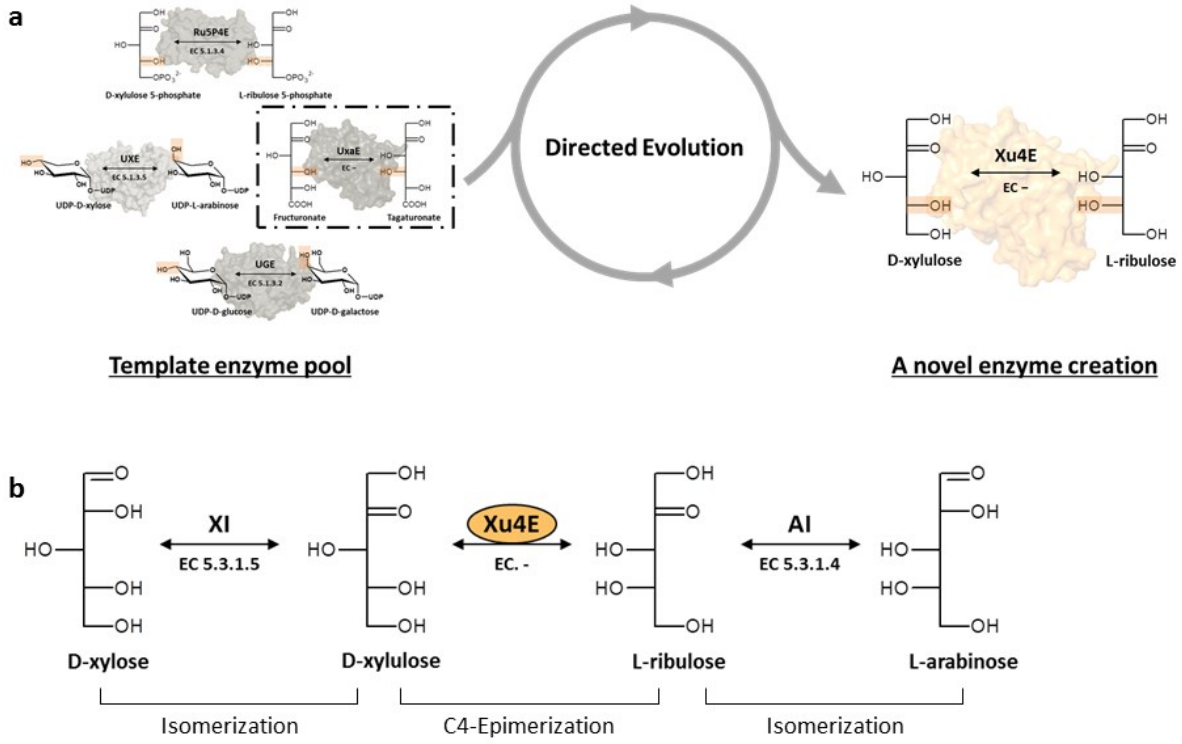

Figure 1. Directed evolution of UxaE to create a novel enzyme, xylulose 4-epimerase (Xu4E), catalyzing 4-epimerization reaction on D-xylulose as a substrate (a). Various epimerases were examined to find the most suitable candidate for directed evolution. A novel three-enzyme pathway comprised of xylose isomerase (XI), Xu4E, and L-arabinose isomerase (AI) was constructed to produce a zero calorie sweetener, L-arabinose, from D-xylose (b).

Materials and Methods

Chemicals and strains

All chemicals were reagent-grade or higher and purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), unless otherwise noted. All enzymes for molecular biology experiments were purchased from New England Biolabs (NEB, Ipswich, MA). Primers were purchased from IDT (Coralvile, IA). Genomic DNA samples of *Thermotoga maritima MSB8* (ATCC43589) and *Escherichia coli* K-12 MG1655 were purchased from the American Type

Culture Collection (Manassas, VA). *E. coli* TOP10 was used for DNA manipulation, plasmid amplification and recombinant protein expression. *E. coli* BL21 (DE3) (Invitrogen, Carlsbad, CA) was used for recombinant protein expression. The Luria-Bertani (LB) medium (Miller formula) with either 100 μg/L ampicillin or 50 μg/L kanamycin was used for *E. coli* cell culture and recombinant protein expression. Strains and plasmids used in this study are listed in Table 1. All oligonucleotides used in this study are listed in Table 2.

Table 1. Strains and plasmids used in this study[a]

Strains/ plasmids	Relevant characteristics	Reference
Strains		
TOP10	Wild type	
JK01	TOP10, *ΔxylB*::*araA* (Km^R)	This study
JK02	JK01, *Δkan*::*FRT-Cm-FRT* (Cm^R)	This study
JK03	JK02, *Δkan*::*FRT*	This study
Plasmids		
pET28a		
pET28-AK	pET28a carrying *araA* gene	This study
pET28a-*tm_uxaE*	pET28a carrying *uxaE* gene	This study
pET28a-P$_{tac}$-*tm_uxaE*	pET28a carrying uxaE gene under P$_{tac}$ promoter	This study
pXZ198_*WT_araC*-P$_{BAD}$-*mCherry*	pXZ198 carrying P$_{BAD}$ expression system and *mCherry* gene	This study
pET28a-P$_{tac}$-*tm_uxaE*-WT_*araC*-P$_{BAD}$-*mCherry* (pScV-*uxaE*)	pET28a carrying a fluorescence-based biosensor consisting of P$_{BAD}$ expression system with wild-type *araC* gene and *mCherry* gene	This study
pET28a-P$_{tac}$-*tm_uxaE*-Mut_*araC*-P$_{BAD}$-*mCherry* (pHScV-*uxaE*)	pET28a carrying a fluorescence-based biosensor consisting of P$_{BAD}$ expression system with truncated *araC* gene and *mCherry* gene	This study
pET20b_tmAI	pET20b carrying *araA* gene from *T. maritima*	This study

[a]Km, kanamycin; Cm, chloramphenicol.

Table 2. Oligonucleotides used in this study

Primers	Sequences (5' → 3')
F_TM0440	GAGATATACCCATATGGTCTTGAAAGTGTTCAAAGACC
epF_TM0440	CCATATGGTCTTGAA
R_TM0440	GGTGGTGGTGCTCGAGCCCCTCCAGCAGATCCACGTGCC
F_pET28a	GCTGGAGGGGCTCGAGCACCACCACCACCACCACTG
R_pET28a	CTTTCAAGACCATATGGGTATATCTCCTTCTTAAAG
epR_pET28a	GTATATCTCCTTCTTAAAGT
DualP_tac_F	TTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAA
DualP_tac_R	ACACATTATACGAGCCGATGATTAATTGTCAACCTATAGTGAGTCGTATT
Ara_F	AGGCTTTAATTTGCAAGCTCACGTGCATATGTTATGACAACTTGACGGCTACATCA
Ara_R	GTTATCCTCCTCGCCCTTGCTCACCATTTTTTATAACCTCCTTAGAGCTC
mCherry_F	GAGCTCTAAGGAGGTTATAAAAAATGGTGAGCAAGGGCGAGGAGGATAAC
mCherry_R	GTTTTATTTGATGCCTCTAGCCAGCTTACTTGTACAGCTCGTCCATGCCG
T1T2_F	CGGCATGGACGAGCTGTACAAGTAAGCTGGCTAGAGGCATCAAATAAAAC
T1T2_R	CACATTTCCCCGAAAAGTGCCACCTTCTAGAAGCTGTTTGTAGAAACGCAAAAAGG

Sc_F	ATGCCTGCAGGTCGACTCTAGAAGCTGTTTG
Sc_R	GCTTCCTTTCAAGCTTTTATGACAACTTGACGGC
HSc_R	GCTTCCTTTCAAGCTTTTAACAACCGGCACGGAAC
ScV_F	AGCTTCTAGAGTCGACCCGGGATCTCGACGCTCTCC
ScV_R	CGGTTGTTAAAAGCTTCCGGGATCTCGACCGATGC
AK_IF	GGATTAGCAGAGCGAGGTATGTAGGATGACGATTTTTGATAATTATGAAG
AK_IR	CCACTTCAAGAACTCTGTAGCACCGTTAGCGACGAAACCCGTAATACAC
AK_VF	GTGTATTACGGGTTTCGTCGCTAACGGTGCTACAGAGTTCTTGAAGTGG
AK_VR	CTTCATAATTATCAAAAATCGTCATCCTACATACCTCGCTCTGCTAATCC
KO_xylB_F	GGCCCGGTTATCGGTAGCGATACCGGGCATTTTTTTAAGGAACGATCGATATGACGATTTTTG ATAATTATGAAG
KO_xylB_R	AGATATATAGATGTGAATTATCCCCCACCCGGTCAGGCAGGGGATAACGTCAGGTGGCACTTT TCGGGGAAATG

Construction of screening plasmids with two different sensitivities

Plasmid pET28a-P_{tac}-*tm_uxaE*-*WT_araC*-P_{BAD}-*mCherry* (pScV-*uxaE*) was constructed as follows. The inserted TM0440 (*tm_uxaE*) gene was amplified from *T. maritima* genomic DNA by PCR with a pair of primers F_TM0440 and R_TM0440. The linearized vector backbone was amplified from pET28a by PCR with a pair of primers F_pET28a and R_pET28a. The two PCR products were digested with NdeI and XhoI, and ligated by NEB Quick Ligation Kit. The ligated plasmid was transformed into *E. coli* TOP10, yielding plasmid pET28a-*tm_uxaE*. pET28a-*tm_uxaE* was PCR-amplified with a pair of primers DualP_tac_F and DualP_tac_R, each of which contains homology P_{tac} promoter sequences. The PCR product was digested with DpnI, and directly transformed into *E. coli* TOP10, yielding plasmid pET28a-P_{tac}-*tm_uxaE*. Yeast transformation-associated recombination cloning was done as previously described (Shao et al., 2009). The *araBAD* promoter system was assembled with *mCherry* gene in a shuttle vector plasmid (pXZ198). The *araBAD* promoter system, *mCherry* gene and a terminator sequence were amplified from pKD46, pET20b-*mCherry* and pKLi028 with a pair of primers Ara-F and Ara-R, mCherry-F and mCherry-R, T1T2-F and T1T2-R, respectively. Yeast transformants grown from agar plates without uracil source were pooled by resuspension in water, and plasmid DNA was isolated using Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research, Orange, CA),

yielding plasmid pXZ198_*WT_araC*-P_{BAD}-*mCherry*. Plasmid pXZ198_*WT_araC*-P_{BAD}-*mCherry* was PCR-amplified with a pair of primers Sc_F and Sc_R. pET28a-P_{tac}-*tm_uxaE* was PCR-amplified with a pair of primers ScV_F and ScV_R. The two PCR products were digested with HindIII and SalI, and ligated by NEB Quick Ligation Kit. The ligated plasmid was transformed into *E. coli* TOP10, yielding plasmid pScV-*uxaE*. Plasmid pET28a-P_{tac}-*tm_uxaE-Mut_araC*-P_{BAD}-*mCherry* (pHScV-*uxaE*) harboring a truncated *araC* gene was constructed by using a different pair of primers Sc_F and HSc_F during the PCR-amplification of pXZ198_*WT_araC*-P_{BAD}-*mCherry*.

Random mutagenesis and mutant library construction

Mutant libraries were constructed as previously described (Kim et al., 2016). The inserted TM0440 (*tm_uxaE*) gene was amplified by error-prone PCR (epPCR) from pET28a-*tm_uxaE* as a template DNA with a pair of primers epF_TM0440 and R_TM0440. The primer epF_TM0440 was ordered as a 5'-end phosphorylated primer. The last 26 bp of the 5' terminus of R_TM0440 were designed to introduce complementary sequences to the vector fragment. The conditions used for random mutagenesis by epPCR were optimized by changing the concentrations of $MnCl_2$ and $MgCl_2$. The reaction mixture with a total volume of 50 µl contained 5 µl of 10× standard *Taq* reaction buffer, 0.2 mM $MnCl_2$, 5 mM $MgCl_2$, 0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP, 1 mM dTTP, 0.4 µM of each primer, 5 ng of template plasmid, and 2.5 U of NEB *Taq* polymerase. The epPCR conditions were as followings: an initial denaturation (120 s at 94 °C), 25 cycles of denaturation (30 s at 94 °C), annealing (15 s at 54 °C), and extension (87 s at 68 °C), and a final extension step (5 min at 68 °C). The linear vector fragment was amplified from either pScV-*uxaE* or pHScV-*uxaE* by using NEB high-fidelity Q5 DNA polymerase with a

pair of primers F_pET28a and epR_pET28a. The first 26 bp of the 5' terminus of F_pET28a were designed to introduce complementary sequences to the insert fragment. The primer epR_pET28a was ordered as a 5'-end phosphorylated primer. The PCR products were recovered from 0.8% w/v agarose gel by using Zymoclean Gel DNA Recovery Kit (Irvine, CA, USA).

The gel-purified PCR products for insert and vector fragments were spliced through OE-PCR. The reaction solution with a total volume of 50 μl contained dNTPs (0.2 mM each), 0.5 μM of each primer (epF_TM0440 and epR_pET28a), the insert fragment (10 ng/μl), the vector fragment (equimolar with the insert fragment), and NEB high-fidelity Q5 DNA polymerase (0.02 U/μl). OE-PCR was conducted as followings: an initial denaturation (60 s at 98 °C), 15 cycles of denaturation (20 s at 98 °C), annealing (20 s at 60 °C), and extension (155 s at 72 °C), and a final extension step (5 min at 72 °C). The OE-PCR product was recovered from 0.8% w/v agarose gel by using Zymoclean Gel DNA Recovery Kit. The purified linear DNA was self-ligated by using NEB Quick Ligation Kit at 25 °C for 5 min. The ligation product was transformed into an engineered *E. coli* TOP10 strain for subsequent high-throughput screening.

Development of in vivo genetic screening system in *E. coli*

Gene knock-out and insertion were accomplished simultaneously by replacing the *xylB* gene with the *araA* gene as previously described (Datsenko and Wanner, 2000). The *araA* gene was amplified from *E. coli* K-12 MG1655 genomic DNA using a pair of primers AK_IF and AK_IR. A backbone vector was amplified from pET28a using a pair of primers AK_VF and AK_VR. Two PCR products were assembled into multimeric plasmids by prolonged overlap extension PCR (You et al., 2012). The PCR product was directly transformed into *E. coli* TOP10, yielding

pET28-AK. The sequence containing *araA* gene and kanamycin resistance gene was amplified from pET28a-AK using a pair of primers KO_xylB_F and KO_xylB_R. The PCR product with flanking DNA homologous to the target chromosome was electroporated into *E. coli* TOP10 expressing the λ Red recombinase from plasmid pKD46. Kanamycin-resistant colonies arising from the homologous recombination were selected and tested by a simple PCR. The resulting strain was named JK01. The kanamycin resistance gene of JK01 was replaced by a FRT-flanked chloramphenicol resistance gene amplified from pKD4, resulting in strain JK02. The antibiotic marker was removed from strain JK02 by using a FLP helper plasmid expressing the FLP recombinase. The FLP helper plasmid was easily cured by growth at 37 °C afterward. The final resulting was named JK03. The phenotype of deleted *xylB* was verified by disability of growing on xylose as a sole carbon source and accumulation of xylulose. The phenotype of inserted *araA* was verified by conversion ability between L-arabinose and L-ribulose.

High-throughput screening of mutant UxaE library

Chemically competent cells of strain JK03 were prepared as described elsewhere (Sambrook and Russell, 2006). The strain JK03 cells transformed with the mutant libraries of UxaE were grown on LB-agar plates supplemented with 1% xylose and 50 μM kanamycin. After 12 hours of incubation at 37 °C, colors of colonies were checked every 12 hours. Positive colonies were identified based on the development of mCherry color. The concentration of xylose and incubation time to take for the color development were adjusted as the desired activity of UxaE increases over rounds of screening. After two rounds of screening, the mutant libraries constructed from pScV-*uxaE* harboring wild-type *araC* gene were used to decrease overall sensitivities of screening. Positive colonies showing relatively higher fluorescence intensities

than others were collected and grown in 4 mL of LB medium supplemented with kanamycin. After overnight incubation at 37 °C, 3 mL of cell cultures were used to extract plasmids for DNA sequencing. 500 μL of cell cultures were used to inoculate freshly prepared 50 mL LB media supplemented with kanamycin for overexpression and purification of mutant UxaEs. Inoculated cultures were grown at 37 °C. The protein overexpression was induced with 0.1 mM IPTG when A_{600} reaches about 0.5, and then continuously cultivated at 37 °C for 9 hours. After harvesting *E. coli* cells by centrifugation, cell pellets were washed twice by re-suspending into a sonification buffer containing 10 mM Tris-HCl buffer (pH 7.5) with 10 g/L NaCl. Final volume of 10 ml of the sonification buffer was added for cell lysis by sonification in an ice bath (Fisher Scientific sonic dismembrator Model 500; 3 s pulse on and 6 s off, total 120 s at 30% amplitude). After sonification, 1 ml of cell lysates was heat-treated at 60 °C for 20 minutes. Soluble cell lysates containing thermostable mutant UxaEs were collected by centrifugation at the maximum speed for 10 min, then used for SDS-gel analysis and enzyme assay.

Overexpression and purification of enzymes

The DNA sequences of UDP-glucose 4-epimerase (UGE) from *Xanthomonas campestris pv. campestris* and *Marinithermus hydrothermalis* were synthesized by GenScript (Piscataway, NJ) with codon optimization, yielding plasmid pTRC99A-Xcc*_galE* and pTRC99A-Mh*_galE*, respectively. Plasmid pET20b-Tm*_galE* encoding UGE from *T. maritima* was prepared and given by You Chun. Plasmids pET20b-Sm*_xylE* encoding UDP-xylose 4-epimerase (UXE) from *Sinorhizobium meliloti* 1021 and pET20b-Pv*_yeeZ* encoding probable UXE from *Pantoea vagans* C9-1 were acquired from other lab. For protein overexpression, *E. coli* BL21 cells transformed with a plasmid harboring a target protein were grown at 37 °C in a 250 mL

Erlenmeyer flask containing 150 mL of LB medium with appropriate antibiotics. The protein expression was induced with 0.1 mM IPTG when the culture A_{600} reached about 0.6, and then cultivated at 18 °C for 16 hours. The *E. coli* cells were harvested by centrifugation and washed with 50 mM of HEPES buffer (pH 7.2) containing 0.5 M NaCl. The harvested cell pellets were re-suspended in the same buffer and lysed by sonification in an ice bath (Fisher Scientific sonic dismembrator Model 500; 3 s pulse on and 6 s off, total 180 s at 50% amplitude). After sonification, soluble cell lysates were collected for purification using ThermoFisher HisPur Ni-NTA Agarose Resin. The target protein was eluted with 100 mM HEPES buffer (pH 7.5) containing 300 mM imidazole after several washing steps with 100 mM HEPES buffer (pH 7.5) containing 10 mM imidazole.

Enzyme activity and product assays using HPLC and a commercial kit

D-xylulose was prepared as a xylose-xylulose mixture by a batch enzymatic reaction of xylose isomerase with D-xylose at 70 °C. D-xylose, D-xylulose, L-ribulose, and L-arabinose were detected by the SHIMADZU high performance liquid chromatography (HPLC) equipped with a refractive index detector. Samples were separated on a Bio-Rad Aminex HPLC organic acid column (HPX-87H 300 mm × 7.8 mm) at 60 °C with a mobile phase of 5 mM sulfuric acid solution at a rate of 0.6 mL/min. In order to specifically measure L-arabinose concentrations, Megazyme L-arabinose/D-galactose assay kit (Product code: K-ARGA, Chicago, IL) was used following its microplate assay procedure.

One-pot biosynthesis of L-arabinose via a novel enzymatic pathway

L-arabinose was produced by one-pot enzymatic reaction conducted in 1 mL of the 50 mM HEPES buffer (pH 8.0) containing 0.2 mM $NiCl_2$, 1 mM $MnCl_2$, 100 mM xylose-xylulose mixture. The xylose-xylulose mixture prepared as described above was used as a substrate. Two enzymes, Xu4E and AI, were added into the reaction solution with the final concentration of 0.5 U/mL for each enzyme. After 3 hours of reaction at 70 ℃, the production of L-arabinose was measured by withdrawing samples for HPLC analysis.

Results and discussion

Numerous 4-epimerase enzymes and a 3-epimerase (UxaE) were tested of their promiscuous activities on D-xylulose as a substrate. UDP-based 4-epimerases, UGEs and UXEs cloned from six different microbial sources, were successfully expressed in *E. coli* BL21 and purified to homogeneity. The different cell sources were *Aquifex aeolicus*, *Xanthomonas campestris pv. Campestris*, *Thermotoga Maritima*, *Marinithermus hydrothermalis*, *Pantoea vagans* and *Sinorhizobium meliloti*. All UDP-based epimerases showed no activity of 4-epimerization with D-xylulose as a substrate, regardless of host organisms. The phosphate-based epimerase, L-ribulose 5-phosphate 4-epimerase, showed only a marginal activity of about 0.0018 U/mg measured at 70 °C. Among epimerases tested in this study, the C3 epimerase, recently discovered tagaturonate–fructuronate epimerase from *T. maritima*, showed the highest activity of about 0.017 U/mg measured at 70 °C. Engineering thermostable proteins are believed to be more robust and evolvable than their marginally stable wild-type counterparts, because amino acid changes are likely to decrease thermostability by destabilizing the tertiary structure (Bloom et al., 2006; Matsumura and Ivanov, 2009). Therefore, UxaE from *T. maritima* showing hyper-

thermostability and the highest activity of Xu4E was used as a starting enzyme for directed evolution to increase the desired activity of 4-epimerization on D-xylulose.

Directed evolution is a powerful method used in protein engineering that mimics the process of natural selection to evolve proteins toward a user-defined goal. The most challenging task for directed evolution is how to rapidly identify the desired mutants from a large-size library of mutants. A novel HTS method was designed to identify more active Xu4E mutants using a transcription-factor-driven biosensor (**Figure 2a**). The key idea of our HTS system is that Xu4E activity is related to the strength of fluorescence signal. The reporter gene, mCherry fluorescent (*mCherry*), is under the control of arabinose promoter (PBAD). Thus, the enhanced activity of Xu4P producing L-arabinose leads to a higher level of mCherry expression, resulting in stronger fluorescence signals (**Figure 2b**). The assimilated xylose is converted to xylulose by an endogenous isomerase (*xylA*). Then, Xu4E continuously expressed under P_{tac} promoter converts D-xylulose to L-ribulose. Finally, L-ribulose is converted to L-arabinose by another endogenous isomerase (*araA*). Once L-arabinose is accumulated high enough to combine with AraC, they form an activator to the P_{BAD} and induce the expression of mCherry protein according to intracellular L-arabinose concentrations.

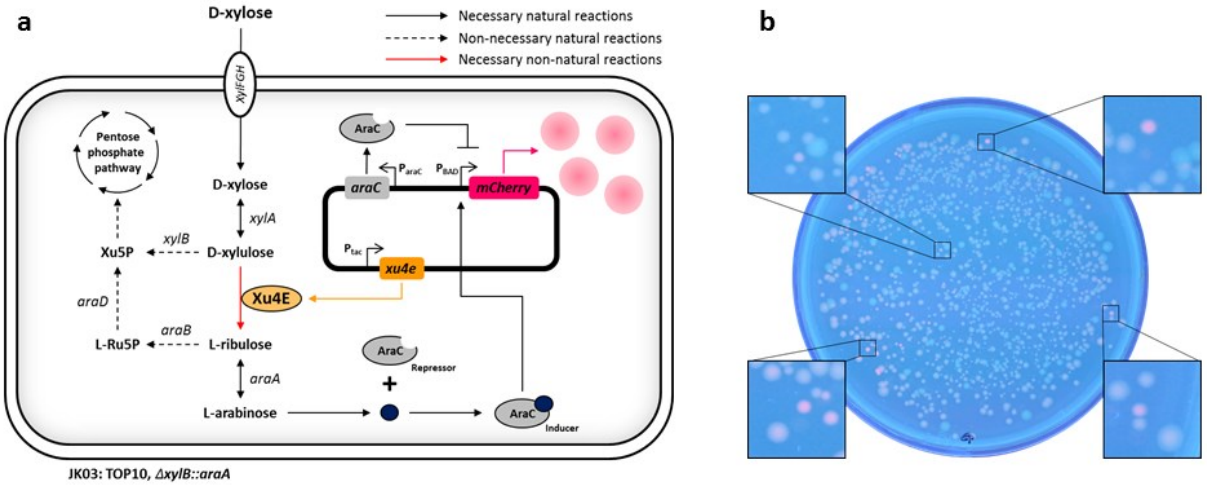

Figure 2. Schematic diagram of the in vivo HTS system to identify mutant Xu4E with increased activity (a). Engineered cells transformed with mutant libraries of Xu4E showed different intensities of mCherry color according to different activities of Xu4E (b).

To accumulate enough intracellular L-arabinose to induce the P_{BAD} promoter, it was necessary to knock out two genes, *xylB* and *araB*, responsible for the metabolism of assimilated D-xylose or its derived sugars (**Figure 3a**). A typical strain *E. coli* TOP10 was chosen for chromosomal gene modification to obtain an engineered strain for the HTS system, because TOP10 strain lacks of the entire L-arabinose operon (*ΔaraABCD*) while retaining the xylose operon (Durfee et al., 2008). *xylB* gene in the chromosome was replaced with *araA* gene amplified from *E. coli* K-12 MG1655. The bacterial chromosome modification was done using the λ Red recombinase system. The *E. coli* strain obtained for the HTS system, JK03, was verified by its disability of growing on xylose or L-arabinose as a sole carbon source (**Figure 3b**) and conversion ability between L-arabinose and L-ribulose (**Figure 3c**).

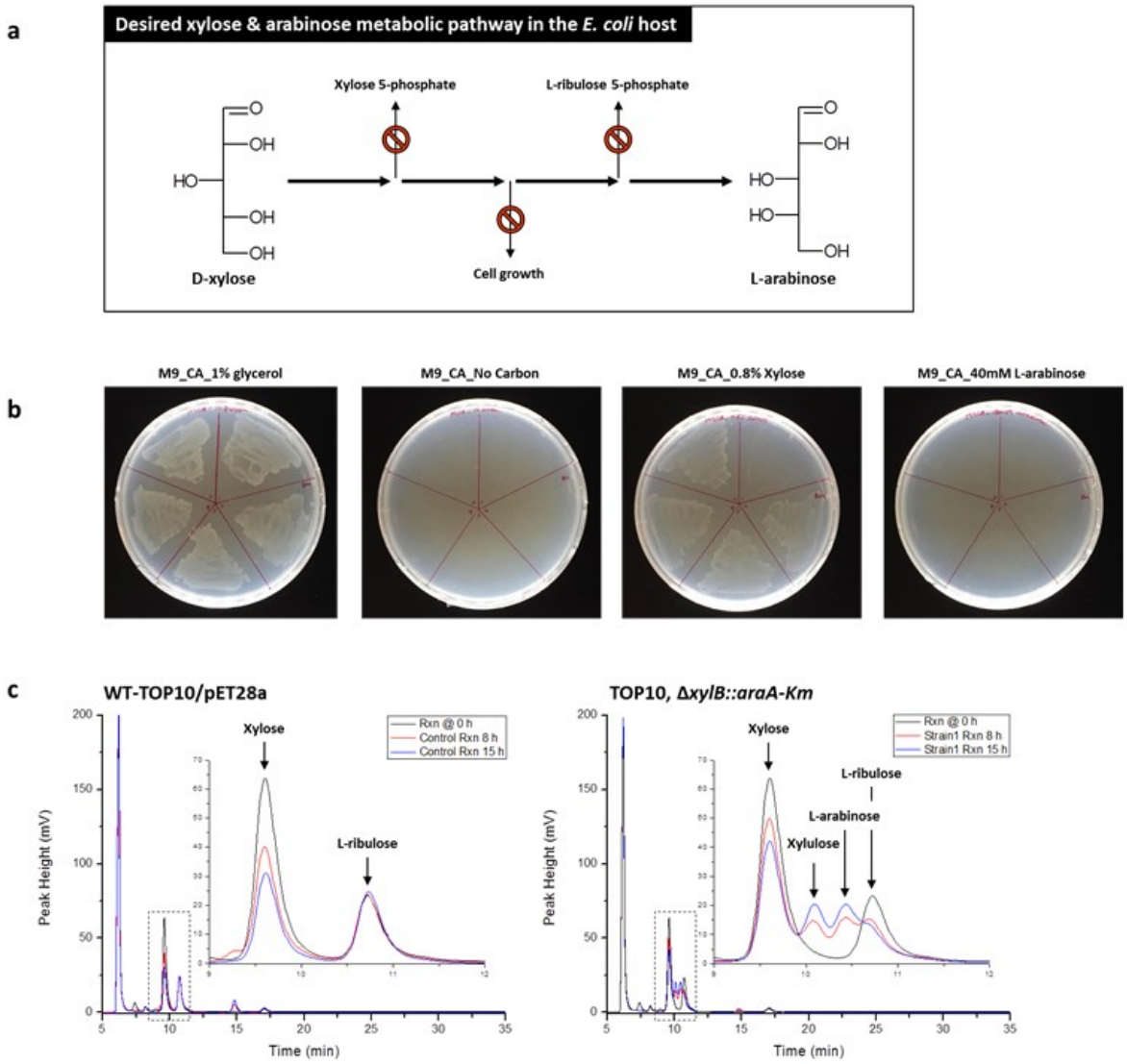

Figure 3. Scheme of desired D-xylose and L-arabinose metabolic pathways in the *E. coli* host for the HTS system (a). Deletion of *xylB* gene was verified by cells unable to grow on D-xylose as a sole carbon source (b) and the accumulation of xylulose (c). The phenotype of inserted *araA* gene was verified by cells able to convert between L-arabinose and L-ribulose (c).

Two screening plasmids with different sensitivities to L-arabinose were successfully constructed (**Figure 4**). Sensitivities of the HTS system was adjusted by using the different screening plasmids; one harboring wild-type AraC (**Figure 4a**) and the other harboring a truncated AraC with increased sensitivity to L-arabinose (**Figure 4b**). The *araBAD* promoter system is regulated by the transcriptional regulatory protein AraC, and the C terminus of AraC was truncated to

different lengths to improve the sensitivity of the system to the inducer (Lee et al., 2007). Sensitivities of the *araBAD* promoter system with wild-type or a truncated AraC were compared on different concentrations of L-arabinose. Production of mCherry was under the control of P_{BAD} promoter. The system with wild-type AraC showed almost no expression of mCherry with 10 μM (**Figure 5a**), while the system with the truncated AraC showed significant levels of mCherry production with the same concentration of L-arabinose (**Figure 5b**). Relative mCherry expression levels of cells with different AraC as a function of L-arabinose concentrations showed that not only the system with the truncated AraC initiates the signal induction with lower L-arabinose concentrations, also generates higher signal intensities at the same L-arabinose concentrations (**Figure 5c**). The increased sensitivity of the system to L-arabinose was also observed from cells grown on LB-agar plates with different L-arabinose concentration (**Figure 5d**).

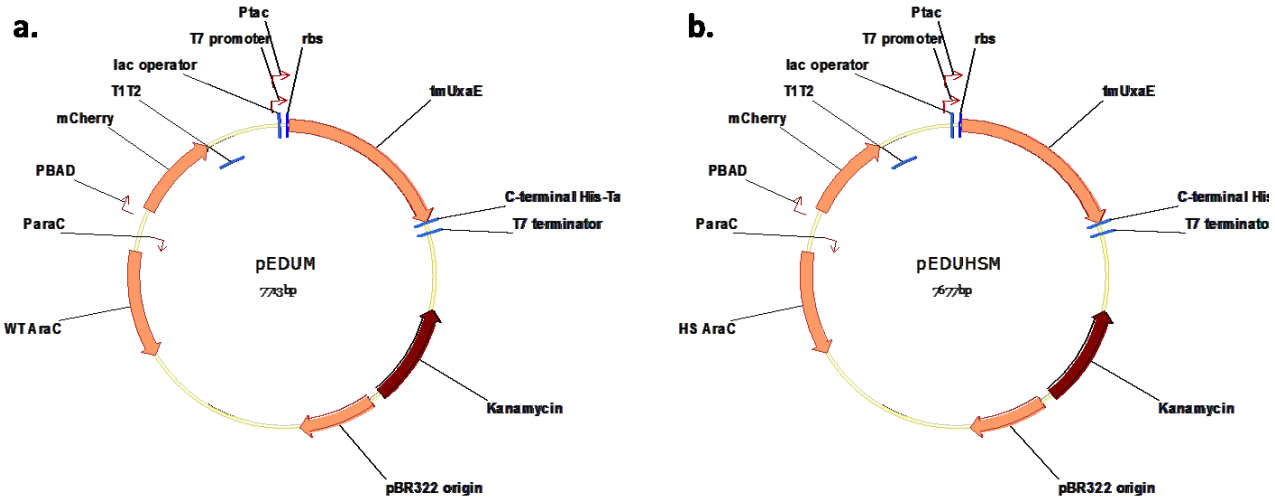

Figure 4. Screening plasmids with different sensitivities to L-arabinose. pEDUM and pEDUHSM include wild-type AraC and truncated AraC as a regulatory protein, respectively, for P_{BAD} system.

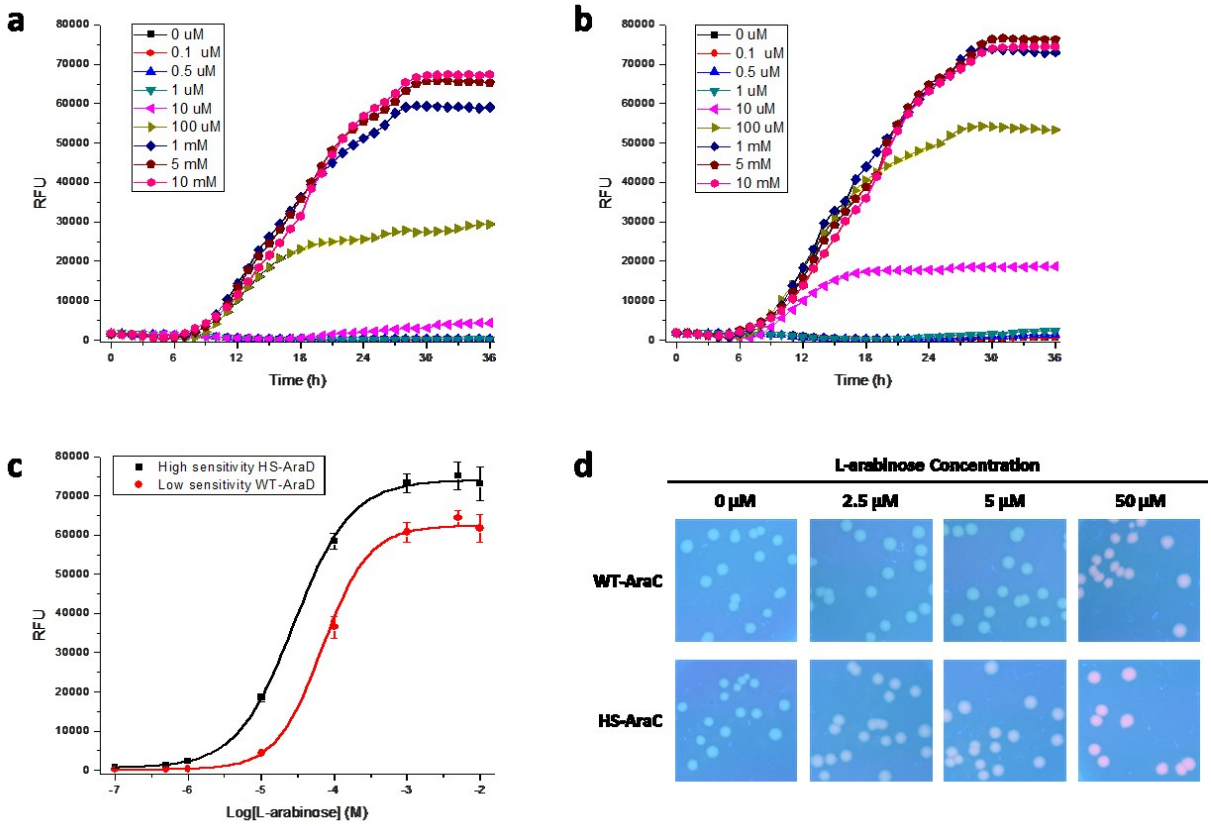

Figure 5. Relative mCherry expression from JK03 strain expressing WT-AraC, wild-type AraC (a) or HS-AraC, truncated AraC with higher sensitivity (b) over time. The expression levels were compared as a function of L-arabinose concentration (c). Different sensitivities were also observed from JK03 cells growing on LB-agar plates supplemented with different L-arabinose concentrations (d).

The HTS with the increased sensitivity was initially used to identify mutant Xu4Es with increased activity of 4-epimerization on D-xylulose. Colonies appeared with higher intensity of the fluorescence color were collected for in vitro enzyme assays. From each round of directed evolution, nearly 1×10^4 colonies were screened. After the second round, the sensitivity of HTS system was decreased by using wild-type AraC for the *araBAD* promoter system. Three positive mutants, DM1-1, DM1-2 and DM1-3, were identified from the first round of directed evolution using a CJ mutant as a template (**Figure 6**). Among mutants from the first round DM1-2 exhibited about 3-fold increased activity of 0.054 U/mg, and it was used as a template for the

next round. From the second round of directed evolution, three mutants were identified with the highest activity of 0.210 U/mg. The third round of mutagenesis was conducted using the DM2-1 as a template. The best mutant, DM3-1, through three consecutive rounds of directed evolution showed the activity of 0.440 U/mg increased by more than 25-fold from the wild-type protein (**Figure 6**). From the three rounds of random mutagenesis and screening, total eight mutants were obtained as positive mutants, and six of them were identified of their changes in amino acid sequences (Table 3).

Table 3. List of mutations in engineered Xu4Es

Enzyme	Active mutations (amino acid)
WT	Wild-type UxaE
CJ-mut	S125D, V163A, D183N, F263I, D311G
DM1-1	T104A, S125D, V163A, N183D, D186N, F263I, D311G, T357A, G441C, E458G
DM1-2	S125D, V163A, N183D, D186N, F263I, D311G
DM1-3	K46E, S125D, V163A, N183D, D186N, F263I, D311G, F464L
DM2-1	S125D, V163A, M179I, N183D, D186N, K257I, F263I, N297Y, D311G, S402F
DM3-1	D72V, G102S, S125D, V163A, M179I, N183D, D186N, K257I, F263I, N297Y, D311G, D394E, S402F
DM3-2	R78K, S125D, V163A, M179I, N183D, D186N, K257I, F263I, N297Y, D311G, F319Y, S402F

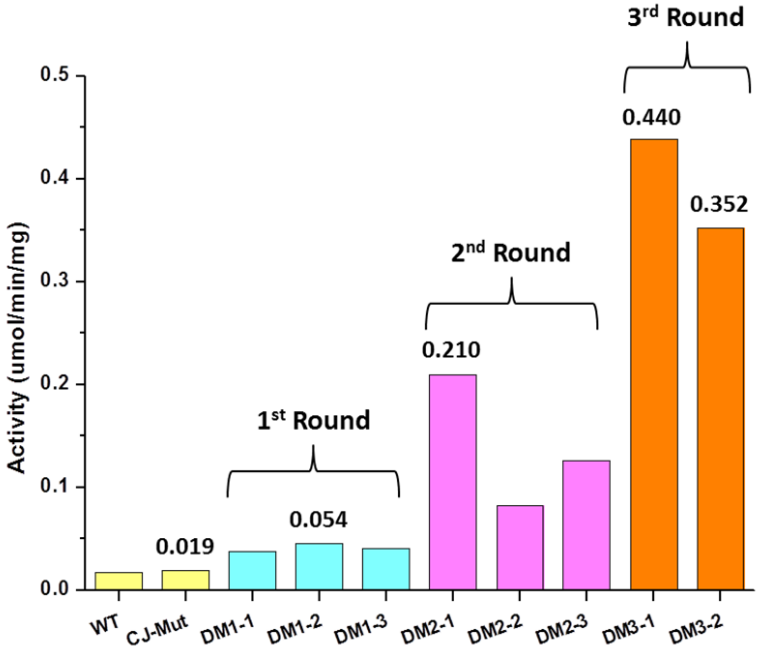

Figure 6. Evolution of UxaE for the activity of 4-epimerization on D-xylulose as a substrate.

A novel three-enzyme pathway comprised of xylose isomerase (XI), Xu4E, and L-arabinose isomerase (AI) was constructed to produce L-arabinose from D-xylose. The production L-arabinose was detected by HPLC (**Figure 7**). The reaction was prepared with DM3-1, the best Xu4E so far, and conducted at 70 °C for 3 hours. The control reaction prepared without DM3-1 showed no peak appearing at the reaction time for L-arabinose. From the reaction with DM3-1, the production of L-arabinose was observed by the peak appearing at the retention time of 10.4 min. The peak of D-xylulose was decreased as converted to L-arabinose.

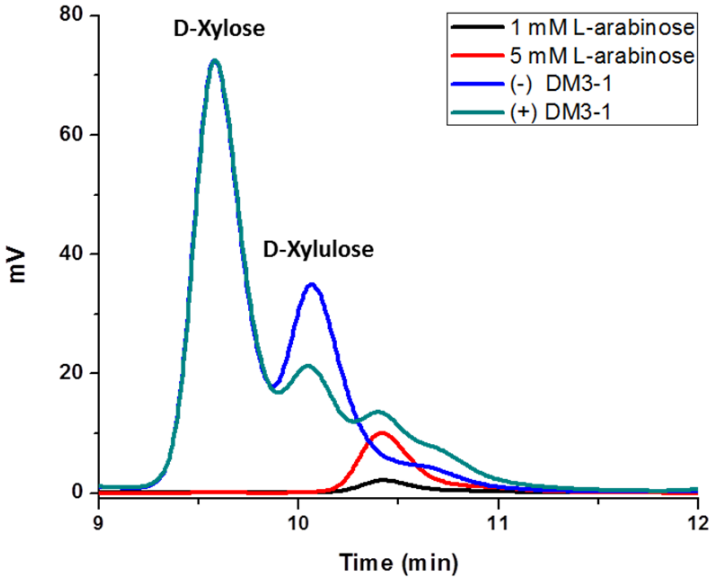

Figure 7. HPLC of reaction mixtures with different conditions. Two standard samples of L-arabinose (1 mM and 5 mM) were shown as references. The control reaction prepared without DM3-1 showed no peak appearing at the retention time of L-arabinose (10.4 min). Production of L-arabinose was observed from the reaction prepared with DM3-1, showing a peak appearing at 10.4 min.

For the first in this study, D-xylulose 4-epimerase was discovered through mining various epimerases for their promiscuous activities of 4-epimerization on D-xylulose as a substrate. Using tagaturonate-fructuronate epimerase as a template, the activity of Xu4E was enhanced by more than 25-fold via multiple-rounds of directed evolution. A transcription-factor-driven

biosensor that specifically senses cytosolic L-arabinose and converts this signal into an observable phenotype was developed for a high-throughput screening method. Enhanced Xu4E was used to construct a novel enzymatic pathway converting biomass sugar D-xylose to zero calorie sweetener L-arabinose. This pathway would enable to produce high-value L-arabinose in a much more cost-effective way than the current L-arabinose production way, expecting to bring great economic benefits to biorefineries and health benefits to patients suffering from high sucrose and fructose diets.

Reference

Alper H, Stephanopoulos G. 2009. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? Nat Rev Microbiol 7(10):715-723.

Beerens K, Van Overtveldt S, Desmet T. 2017. The "epimerring" highlights the potential of carbohydrate epimerases for rare sugar production. Biocatal Biotransformation:1-8.

Bloom JD, Labthavikul ST, Otey CR, Arnold FH. 2006. Protein stability promotes evolvability. Proc Natl Acad Sci USA 103(15):5869-5874.

Burbank F, Fraumeni JF. 1970. Synthetic sweetener consumption and bladder cancer trends in United States. Nature 227(5255):296-297.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. P Natl Acad Sci USA 97(12):6640-6645.

Durfee T, Nelson R, Baldwin S, Plunkett G, Burland V, Mau B, Petrosino JF, Qin X, Muzny DM, Ayele M. 2008. The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. J Bacteriol 190(7):2597-2606.

Gibson GR, Roberfroid MB. 1995. Dietary modulation of the human colonic microbiota - Introducing the concept of prebiotics. J Nutr 125(6):1401-1412.

Goran MI, Dumke K, Bouret SG, Kayser B, Walker RW, Blumberg B. 2013. The obesogenic effect of high fructose exposure during early development. Nat Rev Endocrinol 9(8):494-500.

Ha SJ, Galazka JM, Kim SR, Choi JH, Yang XM, Seo JH, Glass NL, Cate JHD, Jin YS. 2011. Engineered Saccharomyces cerevisiae capable of simultaneous cellobiose and xylose fermentation. Proc Natl Acad Sci U S A 108(2):504-509.

Hao L, Lu XL, Sun M, Li K, Shen LM, Wu T. 2015. Protective effects of L-arabinose in high-carbohydrate, high-fat diet-induced metabolic syndrome in rats. Food Nutr Res 59:28886.

Kearns CE, Schmidt LA, Glantz SA. 2016. Sugar industry and coronary heart disease research: a historical analysis of internal industry documents. JAMA Intern Med 176(11):1680-1685.

Kim JE, Huang R, Chen H, You C, Zhang YHP. 2016. Facile Construction of Random Gene Mutagenesis Library for Directed Evolution Without the Use of Restriction Enzyme in Escherichia coli. Biotechnol J 11(9):1142-1150.

Krog-Mikkelsen I, Hels O, Tetens I, Holst JJ, Andersen JR, Bukhave K. 2011. The effects of L-arabinose on intestinal sucrase activity: Dose-response studies in vitro and in humans. Am J Clin Nutr 94(2):472-478.

Kroger M, Meister K, Kava R. 2006. Low-calorie sweeteners and other sugar substitutes: A review of the safety issues. Compr Rev Food Sci Food Saf 5(2):35-47.

Lanaspa MA, Ishimoto T, Li NX, Cicerchi C, Orlicky DJ, Ruzicky P, Rivard C, Inaba S, Roncal-Jimenez CA, Bales ES and others. 2013. Endogenous fructose production and metabolism in the liver contributes to the development of metabolic syndrome. Nat Commun 4:2434.

Lee SK, Chou HH, Pfleger BF, Newman JD, Yoshikuni Y, Keasling JD. 2007. Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. Appl Environ Microb 73(18):5711-5715.

Lyssiotis CA, Cantley LC. 2013. Metabolic syndrome: F stands for fructose and fat. Nature 502(7470):181-182.

Malik VS, Schulze MB, Hu FB. 2006. Intake of sugar-sweetened beverages and weight gain: A systematic review. Am J Clin Nutr 84(2):274-288.

Matsumura I, Ivanov AA. 2009. How evolving enzymes can beat the heat and avoid defeat. Nat Chem Biol 5(8):539-540.

Nestle M. 2016. Food industry funding of nutrition research: The relevance of history for current debates. JAMA Intern Med 176(11):1685-1686.

Osaki S, Kimura T, Sugimoto T, Hizukuri S, Iritani N. 2001. L-arabinose feeding prevents increases due to dietary sucrose in lipogenic enzymes and triacylglycerol levels in rats. J Nutr 131(3):796-799.

Philippe RN, De Mey M, Anderson J, Ajikumar PK. 2014. Biotechnological production of natural zero-calorie sweeteners. Curr Opin Biotech 26:155-161.

Sambrook J, Russell DW. 2006. The Inoue method for preparation and transformation of competent *E. coli*: "Ultra Competent" cells. CSH Protoc 2006(1).

Seri K, Sanai K, Matsuo N, Kawakubo K, Xue CY, Inoue S. 1996. L-arabinose selectively inhibits intestinal sucrase in an uncompetitive manner and suppresses glycemic response after sucrose ingestion in animals. Metabolism 45(11):1368-1374.

Shao ZY, Zhao H, Zhao HM. 2009. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucleic Acids Res 37(2).

Soffritti M, Belpoggi F, Degli Esposti D, Lambertini L, Tibaldi E, Rigano A. 2006. First experimental demonstration of the multipotential carcinogenic effects of aspartame administered in the feed to Sprague-Dawley rats. Environ Health Perspect 114(3):379-385.

Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, Israeli D, Zmora N, Gilad S, Weinberger A and others. 2014. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature 514(7521):181-186.

You C, Zhang X-Z, Zhang Y-HP. 2012. Simple cloning via direct transformation of PCR product (DNA multimer) to *Escherichia coli* and *Bacillus subtilis*. Appl Environ Microbiol 78(5):1593-1595.

Zhang HR, Pereira B, Li ZJ, Stephanopoulos G. 2015. Engineering *Escherichia coli* coculture systems for the production of biochemical products. Proc Natl Acad Sci U S A 112(27):8266-8271.

Zhang Y-HP. 2011. What is vital (and not vital) to advance economically-competitive biofuels production. Process Biochem 46(11):2091-2110.

Chapter 7. Conclusions

The production of biochemicals and bioenergy based on *in vitro* synthetic biosystems offers a new feasible and promising biomanufacturing platform for biorefinery industries. Hydrogen gas is an alternative transportation fuel and a high-density electricity storage carrier with a potential application for fuel-cell vehicles. Hydrogen as a fuel is a low-value bulk biocommodity where substrate costs contribute a major portion of the production cost (Rollin et al., 2013). Therefore, the complete utilization of substrates for hydrogen production is essentially important for cost-effective production of hydrogen. The *in vitro* synthetic biosystem enabled to achieve the theoretical yield of hydrogen production from starch, one of the abundant and inexpensive renewable resources (i.e., $0.30/kg). Remaining challenges for the commercialization of *in vitro* enzymatic hydrogen production include replacing the expensive and labile coenzyme, $NADP^+$, and further improving enzyme thermal stability and decreasing enzyme production costs. The cost and stability of natural coenzymes could be addressed by replacing them with more stable and less costly (biomimetic) coenzymes (Huang et al., 2016; Rollin et al., 2013; Zhang, 2015), while more thermostable enzymes might be discovered and/or obtained by enzyme engineering and immobilization. With the first industrial example of *in vitro* enzymatic myo-inositol production from starch (You et al., 2017), it is highly anticipated that more examples of commercial success from the *in vitro* synthetic biology will take place within next few years through close collaboration with industrial partners in this emerging area.

In the studies presented here, production of a sugar phosphate and highly marketable zero-calorie sweetener was suggested as a high-return investment to subsidize biofuel production for biorefineries. D-Xylulose 5-phosphate was biosynthesized from D-xylose and polyphosphate

catalyzed by a minimized two-enzyme pathway comprised of xylose isomerase and polyphosphate-utilizing xylulokinase. D-xylulose 5-phosphate as a sugar phosphate is the most suitable product for using *in vitro* synthetic biosystems, because it cannot be fermented efficiently by regular microorganisms due to the negatively charged and hydrophilic phosphate groups. Approximately 32mM Xu5P was produced from 50mM xylose, achieving 64% conversion. This simple cascade reaction without involvement of labile metabolites could provide a cost-effective way to product Xu5P under the broad reaction conditions.

A monosaccharide 4-epimerase, Xu4E, as a new enzyme with very weak promiscuous activity was discovered for the first time in this study. Using directed evolution, the catalytic function of Xu4E from *T. maritima* was enhanced by more than 25-fold. This monosaccharide 4-epimase enabled to construct a novel enzymatic pathway for producing L-arabinose, a FDA-approved zero-calorie sweetener, from D-xylose. Additional efforts should be made for further improvement on the catalytic function of Xu4E with the goal of obtaining the specific activity of more than 1 U/mg at 70 °C. Using Xu4E as a building block of *in vitro* synthetic biosystems, it is possible to design additional enzymatic pathways for production of various derivatives, such as L-ribose.

References

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

Rollin JA, Tam TK, Zhang Y-HP. 2013. New biotechnology paradigm: cell-free biosystems for biomanufacturing. Green Chem 15(7):1708-1719.

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.

Zhang Y-HP. 2015. Production of biofuels and biochemicals by in vitro synthetic biosystems: Opportunities and challenges. Biotechnol Adv 33(7):1467-83.

Appendix a. Supplementary materials for Complete enzymatic phosphorylation of starch with application to green hydrogen gas production at theoretical yield

Kim, J.-E., Kim, E.J., Chen, H., Huang, R., Adams, M.W.W and Zhang, Y.H.P. (2017). Complete enzymatic phosphorylation of starch for green hydrogen gas production at theoretical yield by in vitro metabolic engineering. (*Submitted*)

Supplemental Information

Complete enzymatic phosphorylation of starch with application to green hydrogen gas production at theoretical yield

Jae-Eung Kim,[1] Eui-Jin Kim,[1] Hui Chen,[1] Chang-Hao Wu,[2] Michael W.W. Adams,[2]

Y.-H. Percival Zhang[1,3 *]

1 Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall, Blacksburg, Virginia 24061, USA

2 Department of Biochemistry and Molecular Biology University of Georgia, Athens, Georgia, 30602, USA

3 Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, 32 West 7th Avenue, Tianjin Airport Economic Area, Tianjin 300308, China

*Corresponding authors: YPZ, Email: ypzhang@vt.edu

Tel: (540) 231-7414 [O]; Fax: (540) 231-3199

SI Materials and Methods

Chemicals and strains

All chemicals were reagent-grade or higher and purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), unless otherwise noted. *Escherichia coli* BL21 (DE3) (Invitrogen, Carlsbad, CA) was used to conduct recombinant protein expression. Luria-Bertani (LB) medium (Miller formula) with either 100 μg/L ampicillin or 50 μg/L kanamycin was used for *E. coli* cell culture and recombinant protein expression.

Preparation of recombinant enzymes

All enzymes used in this study are listed in **Table S1**. *Pyrococcus furiosus* hydrogenase SHI was provided from Michael W. W. Adams (Chandrayan et al., 2015). *E. coli* BL21 transformed with expression vectors were grown on agar plates with the appropriate antibiotic overnight. Colonies were picked for seed cultures grown in 3 mL LB medium at 37 °C overnight. Freshly prepared 200 mL of the LB medium containing 100 μg/mL ampicillin or 50 μg/mL kanamycin was inoculated with the seed culture at a ratio of 1% (v/v). Cell cultures were incubated in a rotary shaking rate of 250 rpm at 37 °C until absorbance at 600 nm reached *ca.* 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. Cell cultures were incubated at 37 °C for 6 h or at 18 °C for 20 h for protein expression. Harvested cells by centrifugation at 4 °C were washed twice with 50 mM Tris-HCl buffer (pH 7.5), and re-suspended in 30 mL of 20 mM HEPES buffer (pH 7.5) containing 0.5 M NaCl and 1 mM EDTA. Re-suspended cells were lysed using sonification in an ice bath (Fisher Scientific Sonic Dismembrator Model 500; 3 s pulse on and 6 s off, total 300 s at 50% amplitude). Supernatants containing target proteins were centrifuged at 4 °C. The target proteins were purified by several methods: His-tag purification, CBM-intein self-cleavage, ethylene glycol elution of CBM-tagged enzyme, or heat precipitation (Rollin et al., 2014). The His-tagged proteins G6PDH, 6PGL, 6PGDH, TK, and TAL were purified by affinity adsorption on Ni-charged resins (Thermo Scientific, Rockford, IL). PGM, PGI, and NROR were purified by CBM adsorption followed by intein self-cleavage method. CBM-tagged PPGK was prepared as immobilized on regenerated amorphous cellulose. αGP, IA, 4GT, RPI, RuPE, TIM, ALD and FBP were purified by heat precipitation. For additional purification details, refer to **Table S1**.

Enzyme activity assays

The activity of *Thermotoga maritima* αGP was measured in a 50 mM phosphate buffer (pH 7.2) containing 30 mM maltodextrin (DE 4.0-7.0) (Zhou et al., 2016). The generation of glucose 1-phosphate was assayed by using a glucose assay kit (Sigma-Aldrich, St. Louis, MO).

The activity of *Thermococcus litoralis* 4GT was measured in a 50 mM mM phosphate buffer (pH 7.2) containing 20 mM maltose. The liberated glucose was assayed by using a glucose assay kit (Sigma-Aldrich, St. Louis, MO) (Jeon et al., 1997; Zhou et al., 2016).

The activity of *Thermobifida fusca* PPGK immobilized on CBM was measured by the generation of glucose 6-phophate from polyphosphate and glucose in a 50 mM HEPES buffer (pH 7.5) containing 4 mM $MgCl_2$, 5 mM glucose, and 1 mM polyphosphate for 5 min (Liao et al., 2012). The specific activity of immobilized PPGK on CBM was 48.1 U/mg at 50 ^{o}C.

The activity of *Clostridium thermocellum* PGM was measured by the generation of glucose 6-phosphate from glucose 1-phosphate in a 50 mM HEPES buffer (pH 7.5) containing 5mM glucose 1- phosphate, 5 mM $MgCl_2$ and 0.5 mM $MnCl_2$ for 5 min (Wang and Zhang, 2010). The specific activity of PGM was 410 U/mg at 50 ^oC.

The activity of *Geobacillus stearothermophilus* G6PDH was measured in a 100 mM HEPES buffer (pH 7.5) containing 5 mM $MgCl_2$ and 0.5 mM $MnCl_2$, 2 mM glucose 6-phosphate and 0.67 mM $NADP^+$ for 5 min. The generation of NADPH was measured by increased absorbance at 340 nm (Martin del Campo et al., 2013). The specific activity of G6PDH was 35 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *T. maritima* 6PGL was measured in a 100 mM HEPES buffer (pH 7.5) containing 100 mM NaCl, 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 1 mM glucose 1-phosphate, and 5 mM NAD^+ (Zhu and Zhang, 2017).

The activity of *Moorella thermoacetica* 6PGDH was measured in a 50 mM HEPES buffer (pH 7.5) containing 2 mM 6-phosphogluconate, 1 mM $NADP^+$, 5 mM $MgCl_2$, 0.5 mM $MnCl_2$ for 5 min (Wang and Zhang, 2009). The specific activity of 6PGDH was 16 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *T. maritima* RPI was measured by a modified Dische's cysteine–carbazole method. The specific activity was 290 U/mg at 50 °C (Sun et al., 2012).

The activity of *T. maritima* Ru5PE was measured in a 50 mM Tris/HCl buffer (pH 7.5) containing 2 mM D-ribulose 5-phosphate, 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 0.5 mg/ml BSA for 5 min (Martin del Campo et al., 2013; Wang et al., 2011). The specific activity of Ru5PE was 66 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *Thermus thermophilus* TK was measured in a 50 mM Tris/HCl buffer (pH 7.5) containing 0.8 mM D-xylulose 5-phosphate, 0.8 mM D-ribose 5-phosphate, 5 mM $MgCl_2$, 0.5 mM Thiamine pyrophosphate, 0.15 mM NADH, 60 U/mL of triose-phosphate isomerase, and 20 U/mL of glycerol 3-phosphate dehydrogenase (Myung et al., 2014; Wang et al., 2011). The specific activity of TK was 5.3 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *T. maritima* TAL was measured in a 50 mM glycine–NaOH buffer (pH 9.0) containing 0.6 mM D-fructose 6-phosphate, and 0.5 mM D-erythrose 4-phosphate for 60 s (Huang et al., 2012). The specific activity of TAL was 3.9 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *T. thermophilus* TIM was measured in a 100 mM HEPES buffer (pH 7.5) containing 10 mM $MgCl_2$, 0.5 mM $MnCl_2$ for 5 min (Wang et al., 2011; You et al., 2012). The specific activity of TIM was 450 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *T. thermophilus* ALD was measured in a 100 mM HEPES buffer (pH 7.5) containing 10 mM $MgCl_2$, 0.5 mM $MnCl_2$, 2 mM D-glyceraldehyde 3-phosphate with TIM, FBP, and PGI for 5min. The specific activity of ALD was 16 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *T. maritima* FBP was measured in a 50 mM HEPES buffer (pH 7.5) containing 10 mM $MgCl_2$ and 20 mM fructose 1,6-bisphosphate (Myung et al., 2010). The generation of inorganic phosphate was measured by using the Sahiki method. The specific activity of ALD was 6 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *C. thermocellum* PGI was measured in a 100 mM HEPES buffer (pH 7.5) containing 10 mM $MgCl_2$, 0.5 mM $MnCl_2$ and 5 mM D-fructose 6-phosphate (Myung et al., 2011). The generation of glucose 6-phosphate was measured by using glucose assay kit (Sigma-Aldrich, St. Louis, MO). The specific activity of PGI was 950 U/mg at 50 ^oC (Rollin et al., 2014).

The activity of *P. furiosus* NROR was measured in a degassed 100 mM HEPES buffer (pH 7.5) containing 0.5 mM NADPH and 0.5 mM oxidized benzyl viologen. The increase in absorbance at 578 nm was detected during the assay (Kim et al., 2016). The specific activity of NROR was 20.8 U/mg at 50 ^oC.

System for continuous hydrogen detection

Measurement of hydrogen production was conducted in a continuous flow system purged with 50 mL/min of ultrapure nitrogen (Airgas, Christiansburg, VA) as described previously (Rollin et al., 2015). Hydrogen evolution was detected with a tin oxide thermal conductivity sensor (Figaro TGS 821, Osaka, Japan) which was previously calibrated with in-line flow controllers and ultrapure hydrogen (Airgas, Christiansburg, VA) (Kim et al., 2016). The volume of the reaction was maintained constant by humidifying the carrier gas and controlling the rate of condensation. The temperature of reactor and condenser were controlled at 50 °C and 21.5 °C, respectively, by refrigerated/heated circulating baths (NESLAB RTE7, Thermo Scientific; Isotemp 3016D, Fisher Scientific). Data collection was conducted with a NI USB-6210 (National Instruments, Austin, TX) and analyzed by Lab-View SignalExpress 2009 (National Instruments).

References

Chandrayan SK, McTernan PM, Hopkins RC, Sun JS, Jenney FE, Adams MWW. 2012. Engineering Hyperthermophilic Archaeon Pyrococcus furiosus to Overproduce Its Cytoplasmic [NiFe]-Hydrogenase. J Biol Chem 287(5):3257-3264.

Chandrayan SK, Wu CH, 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.

Huang SY, Zhang Y-HP, Zhong JJ. 2012. A thermostable recombinant transaldolase with high activity over a broad pH range. Appl Microbiol Biotechnol 93(6):2403-2410.

Jeon BS, Taguchi H, Sakai H, Ohshima T, Wakagi T, Matsuzawa H. 1997. 4-alpha-glucanotransferase from the hyperthermophilic archaeon *Thermococcus litoralis* - Enzyme purification and characterization, and gene cloning, sequencing and expression in *Escherichia coli*. Eur J Biochem 248(1):171-178.

Kim EJ, Wu CH, Adams MWW, Zhang Y-HP. 2016. Exceptionally high rates of biological hydrogen production by biomimetic in vitro synthetic enzymatic pathways. Chemistry 22(45):16047-16051.

Liao HH, Myung SW, Zhang Y-HP. 2012. One-step purification and immobilization of thermophilic polyphosphate glucokinase from *Thermobifida fusca* YX: glucose-6-phosphate generation without ATP. Appl Microbiol Biotechnol 93(3):1109-1117.

Martin del Campo JS, Rollin J, Myung S, You C, Chandrayan S, Patino R, Adams MWW, Zhang Y-HP. 2013. High-yield production of dihydrogen from xylose by using a synthetic enzyme cascade in a cell-free system. Angew Chem Int Ed Engl 52(17):4587-4590.

Moustafa HMA, Kim EJ, Zhu ZG, Wu CH, Zaghloul TI, Adams MWW, Zhang Y-HP. 2016. Water splitting for high-yield hydrogen production energized by biomass xylooligosaccharides catalyzed by an enzyme cocktail. ChemCatChem 8(18):2898-2902.

Myung S, Rollin J, You C, Sun FF, Chandrayan S, Adams MWW, Zhang Y-HP. 2014. In vitro metabolic engineering of hydrogen production at theoretical yield from sucrose. Metab Eng 24:70-77.

Myung S, Wang YR, Zhang Y-HP. 2010. Fructose-1,6-bisphosphatase from a hyper-thermophilic bacterium Thermotoga maritima: Characterization, metabolite stability, and its implications. Process Biochem 45(12):1882-1887.

Myung S, Zhang XZ, Zhang Y-HP. 2011. Ultra-Stable Phosphoglucose Isomerase Through Immobilization of Cellulose-Binding Module-Tagged Thermophilic Enzyme on Low-Cost High-Capacity Cellulosic Adsorbent. Biotechnol Progr 27(4):969-975.

Rollin JA, del Campo JM, Myung S, Sun F, You C, Bakovic A, Castro R, Chandrayan SK, Wu C-H, Adams MW. 2015. High-yield hydrogen production from biomass by in vitro metabolic engineering: Mixed sugars coutilization and kinetic modeling. Proc Natl Acad Sci USA 112(16):4964-4969.

Rollin JA, Ye X, del Campo JM, Adams MW, Zhang Y-HP. 2014. Novel hydrogen bioreactor and detection apparatus. Adv Biochem Eng Biotechnol 274:1-17.

Sun FF, Zhang XZ, Myung S, Zhang Y-HP. 2012. Thermophilic Thermotoga maritima ribose-5-phosphate isomerase RpiB: Optimized heat treatment purification and basic characterization. Protein Expres Purif 82(2):302-307.

Wang YR, Huang WD, Sathitsuksanoh N, Zhu ZG, Zhang Y-HP. 2011. Biohydrogenation from Biomass Sugar Mediated by In Vitro Synthetic Enzymatic Pathways. Chem Biol 18(3):372-380.

Wang YR, Zhang Y-HP. 2009. Overexpression and simple purification of the Thermotoga maritima 6-phosphogluconate dehydrogenase in Escherichia coli and its application for NADPH regeneration. Microb Cell Fact 8.

Wang YR, Zhang Y-HP. 2010. A highly active phosphoglucomutase from Clostridium thermocellum: cloning, purification, characterization and enhanced thermostability. J Appl Microbiol 108(1):39-46.

Ye XH, Wang YR, Hopkins RC, Adams MWW, Evans BR, Mielenz JR, Zhang Y-HP. 2009. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. ChemSusChem 2(2):149-152.

You C, Myung S, Zhang Y-HP. 2012. Facilitated Substrate Channeling in a Self-Assembled Trifunctional Enzyme Complex. Angew Chem Int Edit 51(35):8787-8790.

Zhang Y-HP, Evans BR, Mielenz JR, Hopkins RC, Adams MWW. 2007. High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. PLoS One 2(5):e456.

Zhou W, You C, Ma HW, Ma YH, Zhang Y-HP. 2016. One-pot biosynthesis of high-concentration alpha-glucose 1-phosphate from starch by sequential addition of three hyperthermophilic enzymes. J Agric Food Chem 64(8):1777-1783.

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

Zhu ZG, Sun FF, Zhang XZ, Zhang Y-HP. 2012. Deep oxidation of glucose in enzymatic fuel cells through a synthetic enzymatic pathway containing a cascade of two thermostable dehydrogenases. Biosens Bioelectron 36(1):110-115.

SI Figures

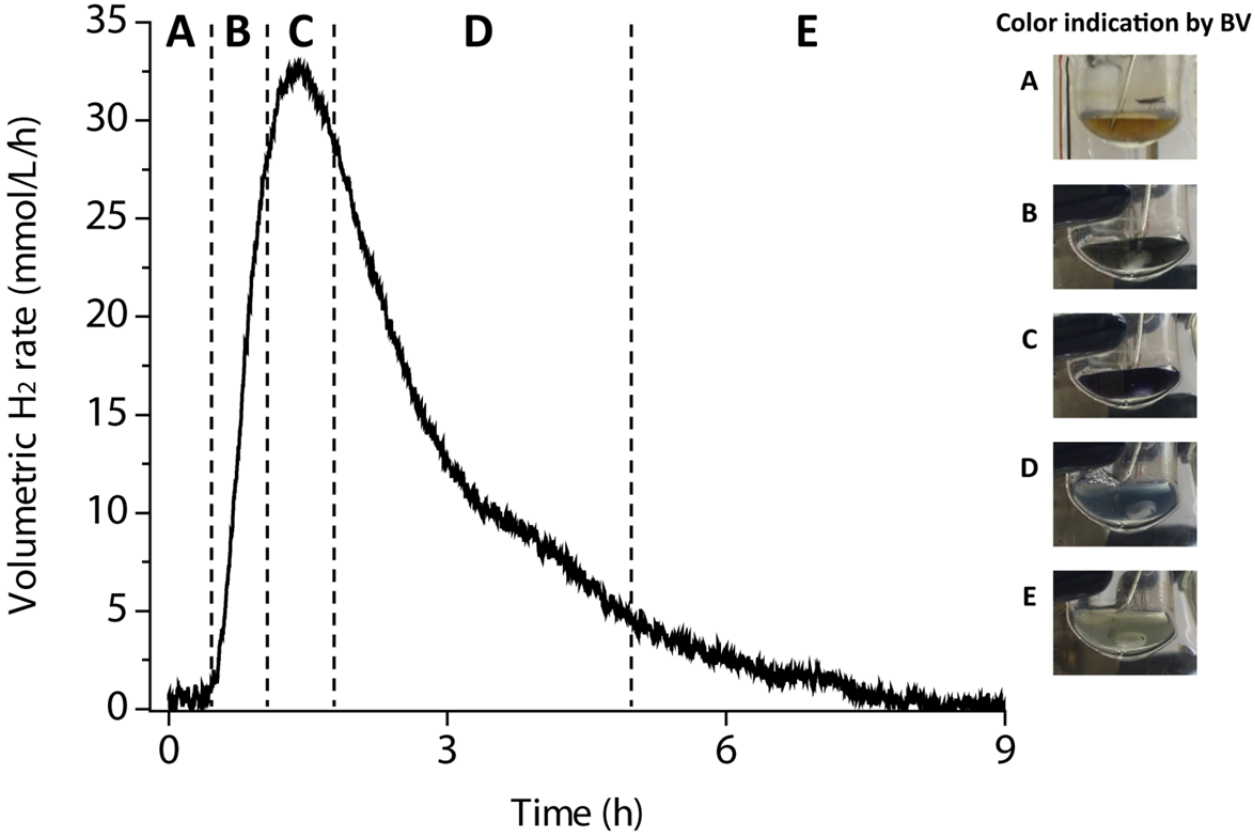

Figure S1. Hydrogen production and redox status of the reactants indicated by BV. When oxidized BV (colorless) was initially added, the reaction solution was nearly transparent (A). As the reaction started, NADPH was generated, and oxidized BV was converted to reduced BV (blue color) by developing a blue color (B). The reaction developed a deeper blue color as more NADPH generated and more oxidized BV converted to reduced BV. The strongest blue color was in consistent with the maximum H_2 volumetric productivity (C). Reducing power decreased as substrate concentration decreases, lightening the blue color (D). With a depletion of substrate, the reaction stopped producing hydrogen while leaving most BV in the oxidized form with no color (E).

Figure S2. The formation of bubbles as an indicator of the high-speed hydrogen production. When the reaction was run with a high substrate concentration (up to 125 g/L), the formation of bubbles was observed.

SI Tables

Table S1. The list of enzymes and their properties for hydrogen generation at theoretical yield from starch.

Enzyme	Abbreviation	E.C. #	Source	Purification	Sp. Act. at 50 °C (U/mg)	Units (U/mL)	Ref.
Isoamylase	IA	3.2.1.68	*Sulfolobus tokodaii*	Heat precipitation	-	[S]/[E] = 1:1000*	(Zhou et al., 2016)
α-glucan phosphorylase	αGP	2.4.1.1	*Thermotoga maritima*	Heat precipitation	20†	5	(Zhou et al., 2016)
4-α-glucanotransferase	4GT	2.4.1.25	*Thermococcus litoralis*	Heat precipitation	3.2†	2	(Jeon et al., 1997)
Polyphosphate glucokinase	PPGK	2.7.1.63	*Thermobifida fusca*	CBM immobilization	48.1	5	(Liao et al., 2012)
Phosphoglucomutase	PGM	5.4.2.2	*Clostridium thermocellum*	CBM adsorption-intein cleavage	410	5	(Wang and Zhang, 2010)
Glucose 6-phosphate dehydrogenase	G6PDH	1.1.1.49	*G. stearothermophilus*	His-tag adsorption on NTA	35	5	(Zhu et al., 2012)
6-phosphogluconolactonase	6PGL	3.1.1.31	*T. maritima*	His-tag adsorption on NTA	230‡	5	(Zhu and Zhang, 2017)
6-phosphogluconate dehydrogenase	6PGDH	1.1.1.44	*Moorella thermoacetica*	His-tag adsorption on NTA	16	2	(Wang and Zhang, 2009)
Ribose 5-phosphate isomerase	RPI	5.3.1.6	*T. maritima*	Heat precipitation	290	2	(Sun et al., 2012)
Ribulose 5-phosphate 3-epimerase	RuPE	5.1.3.1	*T. maritima*	Heat precipitation	66	2	(Myung et al., 2014)
Transketolase	TK	2.2.1.1	*Thermus thermophilus*	His-tag adsorption on NTA	5.3	2	(Myung et al., 2014)
Transaldolase	TAL	2.2.1.2	*T. maritima*	His-tag adsorption on NTA	3.9	2	(Huang et al., 2012)
Triose phosphate isomerase	TIM	5.3.1.1	*T. thermophiles*	Heat precipitation	450	1	(Wang et al., 2011)
Aldolase	ALD	4.1.2.13	*T. thermophiles*	Heat precipitation	16	1.5	(Myung et al., 2014)
Fructose-1,6-bisphosphatase	FBP	3.1.3.11	*T. maritima*	Heat precipitation	6	2	(Myung et al., 2010)
Phosphoglucose isomerase	PGI	5.3.1.9	*C. thermocellum*	CBM adsorption-intein cleavage	950	2	(Myung et al., 2011)
NADPH rubredoxin oxidoreductase	NROR	1.18.1.4	*Pyrococcus furiosus*	CBM adsorption-intein cleavage	20.8§	1	(Kim et al., 2016)
[NiFe]-hydrogenase SHI	H_2ase	1.12.1.3	*P. furiosus*	-	2¶	5	(Chandrayan et al., 2012)

* [S]/[E] is subtrate/enzyme weight ratio.
† Specific activity was measured at 70 °C.
‡ Specific activity was measured at 23 °C.
§ Specific activity was measured in the presence of benzyl viologen.
¶ NADPH-based activity at 80 ^{o}C. Reduced methyl viologen-based specific activity at 80 ^{o}C is 121 U/mg (Chandrayan et al., 2015).

Table S2. Comparison of hydrogen production rates and yield from different less-costly carbohydrates

Low-cost substrate	Concentration (mM)	Temperature (°C)	Volumetric productivity (mmol H_2/L/h)	Yield (%)	Ref.
Starch	2*	30	0.4	43	(Zhang et al., 2007)
Cellobiose	2*	32	0.5	93	(Ye et al., 2009)
Cellodextrins	8*	32	3.9	68	(Ye et al., 2009)
Xylose	2	50	2.2	96	(Martin del Campo et al., 2013)
Sucrose	2*	37	3.0	97	(Myung et al., 2014)
Sucrose	50*	37	8.1	-	(Myung et al., 2014)
Biomass sugars	2* + 2	40	2.3	100	(Rollin et al., 2015)
Xylooligosaccharides	2	50	4.7	95	(Moustafa et al., 2016)
Starch	5*	50	36.1	99	This study
Starch	125*	50	90.2	-	This study

* Glucose equivalent.

Table S3. Comparison of sources of recombinant enzymes used for in vitro hydrogen production.

Module	Enzyme (abbreviation)	E.C. #	Enzyme origin			
			Zhang, 2007 (Glycogen)	**Myung, 2014 (Sucrose)**	**Rollin, 2015 (Biomass sugars)**	**This study (Starch)**
Substrate phosphorylation	Isoamylase (IA)	3.2.1.68	-	-	-	*S. tokodaii*
	α-glucan phosphorylase (αGP)	2.4.1.1	Rabbit muscle	-	-	*T. maritima*
	4-α-glucanotransferase (4GT)	2.4.1.25	-	-	-	*T. litoralis*
	Sucrose phosphorylase (SP)	2.4.1.7	-	*L. mesenteroides*	-	-
	Polyphosphate glucokinase (PPGK)	2.7.1.63	-	*T. fusca*	*T. fusca*	*T. fusca*
	Phosphoglucomutase (PGM)	5.4.2.2	Rabbit muscle	*C. thermocellum*	-	*C. thermocellum*
	Xylose isomerase (XI)	5.3.1.5	-	*S. murinus*	*S. murinus*	-
	Polyphosphate xylulokinase (PPXK)	2.7.1.17	-	-	*T. maritima*	-
NADPH generation	Glucose 6-phosphate dehydrogenase (G6PDH)	1.1.1.49	*S. cerevisiae*	*G. stearothermophilus*	*G. stearothermophilus*	*G. stearothermophilus*
	6-phosphogluconolactonase (6PGL)	3.1.1.31	-	-	-	*T. maritima*
	6-phosphogluconic dehydrogenase (6PGDH)	1.1.1.44	*S. cerevisiae*	*M. thermoacetica*	*M. thermoacetica*	*M. thermoacetica*
G6P regeneration (PPP and gluconeogenesis)	Ribose 5-phosphate isomerase (RPI)	5.3.1.6	Spinach	*T. maritima*	*T. maritima*	*T. maritima*
	Ribulose-5-phosphate 3-epimerase (Ru5PE)	5.1.3.1	*S. cerevisiae*	*T. maritima*	*T. maritima*	*T. maritima*
	Transketolase (TK)	2.2.1.1	*E. coli*	*T. thermophilus*	*T. thermophilus*	*T. thermophilus*
	Transaldolase (TAL)	2.2.1.2	*S. cerevisiae*	*T. maritima*	*T. maritima*	*T. maritima*
	Triose-phosphate isomerase (TIM)	5.3.1.1	Rabbit muscle	*T. thermophilus*	*T. thermophilus*	*T. thermophilus*
	Aldolase (ALD)	4.1.2.13	Rabbit muscle	*T. thermophilus*	*T. thermophilus*	*T. thermophilus*
	Fructose-1,6-bisphosphatase (FBP)	3.1.3.11	*E. coli*	*T. maritima*	*T. maritima*	*T. maritima*
	Phosphoglucose Isomerase (PGI)	5.3.1.9	*S. cerevisiae*	*C. thermocellum*	*C. thermocellum*	*C. thermocellum*
Hydrogenation	NADPH rubredoxin oxidoreductase (NROR)	1.18.1.4	-	-	-	*P. furiosus*
	Hydrogenase SHI (H_2ase)	1.12.1.3	*P. furiosus*	*P. furiosus*	*P. furiosus*	*P. furiosus*