

**Development of Building Blocks - Thermostable Enzymes for Synthetic  
Pathway Biotransformation (SyPaB)**

Fangfang Sun

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

Master of Science

In

Biological System Engineering

Y.-H. Percival Zhang, Chair

Chenming (Mike) Zhang

Daniel G. S. Capelluto

April 25<sup>th</sup>, 2012

Blacksburg, Virginia

Keywords: biofuel, thermostable enzymes, ribose-5-phosphate isomerase, RpiB,  
glucose-6-phosphate dehydrogenase, diaphorase, substrate channeling,  
synthetic pathway biotransformation (SyPaB)

# **Development of Building Blocks - Thermostable Enzymes for Synthetic Pathway Biotransformation (SyPaB)**

Fangfang Sun

## **ABSTRACT**

Hydrogen production from abundant renewable biomass would decrease reliance on crude oils, achieve nearly zero net greenhouse gas emissions, create more jobs, and enhance national energy security. Cell-free synthetic pathway biotransformation (SyPaB) is the implementation of complicated chemical reaction by the in vitro assembly of numerous enzymes and coenzymes that microbes cannot do. One of the largest challenges is the high cost and instability of enzymes and cofactors. To overcome this obstacle, strong motivations have driven intensive efforts in discovering, engineering, and producing thermostable enzymes.

In this project, ribose-5-phosphate isomerase (RpiB), one of the most important enzymes in the pentose phosphate pathway, was cloned from a thermophile *Thermotoga maritima*, and heterologously expressed in *Escherichia coli*, purified and characterized. High-purity RpiB was obtained by heat pretreatment through its optimization in buffer choice, buffer pH, as well as temperature and duration of pretreatment. This enzyme had the maximum activity at 80°C and pH 6.5 – 8.0. It had a half lifetime of 71 h at 60°C, resulting in its turn-over number of more than  $2 \times 10^8$  mol of product per mol of enzyme. Another two thermostable enzymes glucose-6-phosphate dehydrogenase (G6PDH) and diaphorase (DI) and their fusion proteins G6PDH-DI

and DI-G6PDH were cloned from *Geobacillus stearothermophilus*, heterologously expressed in *E. coli* and purified through its His-tag. The individual proteins G6PDH and DI have good thermostability and reactivity. However, the presence of DI in fusion proteins drastically decreased G6PDH activity. However, a mixture of G6PDH and a fusion protein G6PDH-DI not only restored G6PDH activity through the formation of heteromultimeric network but also facilitated substrate channeling between DI and G6PDH, especially at low enzyme concentrations.

My researches would provide important building blocks for the on-going projects: high-yield hydrogen production through cell-free enzymatic pathways and electrical energy production through enzymatic fuel cells.

## ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude and appreciation to my advisor, Dr. Y.-H. Percival Zhang, for his guidance and financial support throughout this research. I would also like to thank Dr. Chenming (Mike) Zhang and Dr. Daniel G. S. Capelluto for serving on my graduate committee. Through committee meetings and discussions, they have provided me with invaluable advice and guidance.

I owe thanks to everyone in my lab for their help throughout my thesis. Many thanks go to Dr. Xiao-Zhou Zhang, Dr. Chun You, Dr. Hui Ma, Dr. Hongge Chen, Dr. Xing Zhang, Dr. Noppadon Sathitsuksanoh, Dr. Xinhao Ye, Zhiguang Zhu, Suwan Myung, Joe Rollin and Hehuan Liao for numerous helps in my project and the remaining members in Biofuels lab at VT for our time working together. Additional thanks are extended to Dr. Hong Zheng and Yun Hu, Wei Huang in Dr. Mike Zhang's lab. They gave me many helps in using the Malvern Zetasizer Nano ZS system. Besides, I am grateful to all the graduate student friends, especially those from China, and many staffs and faculties in my department for all of their help in my research.

Most importantly, none of this work would have been done without the support from my beloved husband and son, Xiao-Zhou Zhang and Daniel Zhang. I would also like to thank my family, including my parents, parents-in-law, brother and sister, who have all supported me wholeheartedly throughout my life.

## Table of Contents

1. Background and goals.....	1
1.1 Introduction of biofuels.....	1
1.2 Cell-free Synthetic Pathway Biotransformation (SyPaB).....	3
1.3 Thermostable enzymes.....	7
1.5 Goals and significance.....	9
2. Thermophilic <i>Thermotoga maritima</i> ribose-5-phosphate isomerase RpiB: Optimized heat treatment purification and basic characterization.....	10
2.1 Abstract.....	10
2.2 Introduction.....	10
2.3 Materials and methods.....	16
2.3.1 Chemicals and strains.....	16
2.3.2 Protein expression plasmid construction.....	16
2.3.3 Expression and purification of recombinant protein.....	17
2.3.4 SDS-PAGE and determination of protein concentration.....	17
2.3.5 Enzyme activity assay.....	18
2.3.6 Optimization of RpiB reaction conditions.....	18
2.3.7 Thermostability.....	19
2.4 Results.....	19
2.4.1 Overexpression of RpiB.....	19
2.4.2 Optimized purification of RpiB.....	20
2.4.3 Basic properties.....	22
2.4.4 Thermostability.....	23
2.5 Discussion.....	27
3. Spatial organization of heteromultimeric complex is important to restore multimeric glucose-6-phosphate dehydrogenase activity and channel substrate between two cascade enzymes.....	30
3.1 Abstract.....	30
3.2 Introduction.....	30

3.2 Materials and methods .....	35
3.2.1 Chemicals and strains .....	35
3.2.2 Construction of recombinant plasmids .....	36
3.2.3 Recombinant protein expression and purification .....	38
3.2.4 Enzyme activity assays .....	39
3.2.5 Assay of particle size of enzymes .....	41
3.2.6. Other assays .....	41
3.2.7. Native gel .....	41
3.3 Results .....	42
3.3.1 Plasmid construction and protein purification .....	42
3.3.2 Substrate channeling .....	45
3.3.3 Particle size in aqueous solution .....	50
3.4 Discussion .....	53
4. Conclusions and perspectives .....	55
4.1 Conclusions .....	55
4.2 Perspectives .....	55
References .....	57
Appendix .....	66
Fig A1 .....	67
Fig A2 .....	68

## LIST OF FIGURES

<b>Fig. 1-1.</b> The cycle of cell-free synthetic pathway biotransformation (SyPaB) .....	4
<b>Fig. 1-2.</b> The synthetic pathway for conversion of starch and water to hydrogen and carbon dioxide.....	6
<b>Fig. 2-1.</b> Isomerization reactions of ribose-5-phosphate (R5P) to ribulose 5-phosphate (Ru5P). 15	
<b>Fig. 2-2.</b> SDS-PAGE analysis of RpiB purification in 50 mM Tris-HCl buffer (pH 7.0) and 50 mM HEPES buffer (pH 7.0) by heat treatment at 80 °C for 40 min.....	21
<b>Fig. 2-3.</b> The profiles of pH (A) and temperature (B) of the purified RpiB .....	25
<b>Fig. 2-4.</b> Thermo-inactivation of the purified RpiB (0.1mg/mL) at 60, 70 and 80 °C in 50 mM Tris-HCl buffer (pH 7.0).....	26
<b>Fig. 3-1.</b> Construction of the protein expression plasmids .....	43
<b>Fig. 3-2.</b> 12% SDS-PAGE analysis of the purified free enzymes and fusion enzymes (A) and 10% Native gel analysis of free enzymes, fusion enzymes and the mixture of G6PDH and fusion enzymes (B) .....	44
<b>Fig. 3-3.</b> Schematic representation of the catalytic reactions in the mixture of free G6PDH and DI system (A) and the possible substrate channeling in the mixture of free G6PDH and fusion enzyme G6PDH/DI system (B) .....	46
<b>Fig. 3-4.</b> Enzyme activity assays of free enzymes, fusion enzymes and the mixture of G6PDH and fusion enzymes G6PDH-DI and DI-G6PDH .....	48
<b>Fig. 3-5.</b> The effects of enzyme concentration on the degree of substrate channeling of the mixture of G6PDH/G6PDH-DI relative to the mixture of 2 × G6PDH/DI at the same enzyme loading.....	49
<b>Fig. 3-6.</b> Particle sizes of the individual enzyme components and fusion enzymes determined by the Malvern Zetasizer Nano ZS system .....	52

## LIST OF TABLES

<b>Table 2-1.</b> The RpiB purification from 250 mL of the <i>E. coli</i> cell culture .....	22
<b>Table 2-2.</b> The kinetic characteristics of RpiB based on ribose-5-phosphate as the substrate ....	23
<b>Table 2-3.</b> Comparison of ribose-5-phosphate isomerases from different sources in the literature .....	28
<b>Table 3-1.</b> Strain, plasmids and primers used in this study .....	40
<b>Table 3-2.</b> Specific activities of G6PDH, DI, as well as two fusion proteins G6PDH-DI and DI-G6PDH.....	45

## **1. Background and goals**

### **1.1 Introduction of biofuels**

A major challenge that mankind is facing in this century is the gradual and inescapable exhaustion of the earth's fossil energy resources. The combustion of these fossil energy materials lavishly used as heating, electricity generation and transportation fuel is one of the key factors responsible for climate change due to large-scale carbon dioxide emissions. Thus, developments in alternative renewable energy sources are urgently needed. High energy prices, increasing energy imports, concerns about crude oil supplies, and greater recognition of the negative environmental consequences of the use of fossil fuels also have driven more interests in developing transportation biofuels.

Biofuels can be defined as any solid, liquid, or gaseous fuels that are produced from biomass or made from biological approaches. The switch from nonrenewable fossil fuels to renewable resources provides a means to decrease the reliance on fossil fuels as well as to decrease global net emissions of greenhouse gases into the environment (Demain et al. 2005; Lynd et al. 2002; Zhang 2009). To be a viable substitute for a fossil fuel, an alternative fuel should not only have superior environmental benefits over the fossil fuels it displaces, be economically competitive with it, and be producible in sufficient quantities to meet energy demands, but it should also provide a net energy gain over the energy sources used to produce it (Hill et al. 2006).

Based on their production technology, renewable biofuels can be classified as first generation biofuels, second generation biofuels, third generation biofuels and fourth generation biofuels (Demirbas 2009). First generation biofuels refer to biofuels made from sugar, starch, vegetable

oils, or animal fats using conventional technology. First generation biofuels have already resulted in reduced vehicular emissions of greenhouse (Hill et al. 2006). However, first generation biofuels produced from food crops such as grains, sugar beet, and oil seed are limited in their ability to achieve targets for oil-product substitution, climate change mitigation, and economic growth (Demirbas 2009; Yeoman et al. 2010). Moreover, first generation biofuels produced from corn starch and sugar cane is neither economically nor ecologically sustainable, because both crops require large land area for their cultivation and compete with food crops indispensable for human consumption. Second generation biofuels made from nonfood feedstocks from lignocellulosic materials include cereal straw, wheat straw, corn, wood, forest residues and some short rotation forests are more desirable to tackle the looming environmental and social crisis (Yeoman et al. 2010). Second and third generation biofuels are also called advanced biofuels. Third generation biofuels includes algae fuel. Due to many advantages of algae, several biofuels such as biodiesel, bioethanol, hydrogen can be produced from algae with most advanced technology. Fourth generation biofuels are created using revolutionary processes that defies any other category of biofuels, such as enzymatic hydrogen.

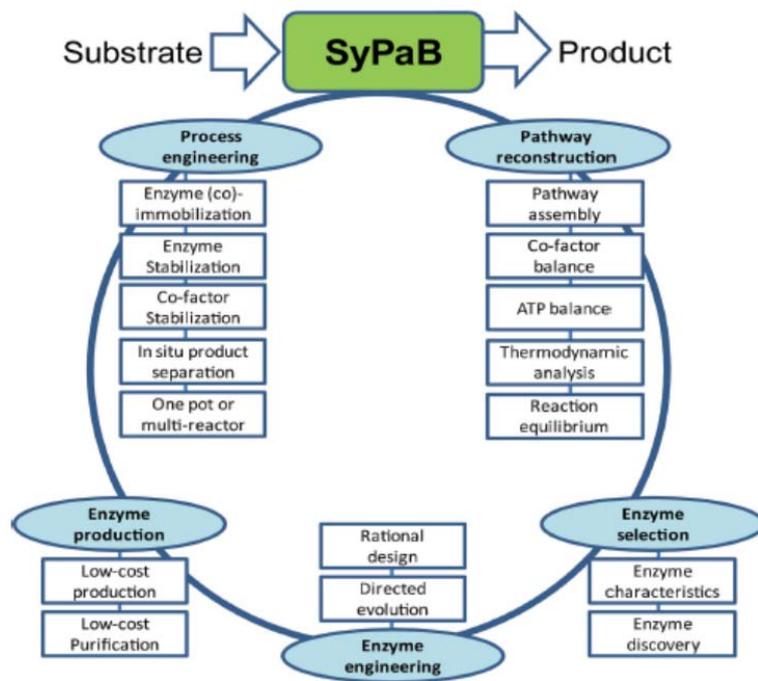
Because of some special requirements for the transportation sector, such as high power density and high-energy storage capacity in a small volume, potential solutions for the transportation sector are limited. Biomass, the most abundant renewable bioresource, is the only low-cost renewable resource that can be utilized for the production in large scale transportation fuels and renewable materials (Zhang et al. 2010). Starting from biomass sugars, numerous biocatalytic approaches have been proposed and investigated for the production of various biofuels, such as ethanol (Shaw et al. 2008), hydrogen (Ye et al. 2009), high-chain alcohols (Atsumi et al. 2008),

fatty acids or their esters (Kalscheuer et al. 2006). Liquid biofuels work well with current infrastructure and internal combustion engines, meaning a short-term or middle-term solution to the transportation sector. In future, hydrogen fuel system affords a long-term solution for transportation sector because hydrogen fuel cells have higher energy efficiencies and produce less pollution than internal combustion engines (Zhang et al. 2010).

### **1.2 Cell-free Synthetic Pathway Biotransformation (SyPaB)**

Cell-free synthetic pathway biotransformation (SyPaB) is the *in vitro* assembly of a number of purified enzymes and coenzymes for the production of desired end products through complicated biochemical reaction networks that a single enzyme cannot implement (Zhang 2010). Synthetic pathway assembly is usually designed based on natural metabolic pathways with necessary modifications with ATP and coenzyme balanced (Zhang et al. 2010). SyPaB and microbial fermentation can implement similar level complicated biochemical reactions. In principle, microbes contain thousands of proteins responsible for metabolism regulation, self-duplication, and formation of a desired product, where a small fraction of cellular enzymes are usually responsible for transforming the substrate into the desired end product. However, in the SyPaB process, it contains only the enzymes that responsible for the production of the desired end product without any other side pathways. Cell-free SyPaB, as compared to microbial fermentation, has several distinctive advantages, such as high production yield, good engineering flexibility, high product titer, fast reaction rate, and simplified process control (Zhang 2010). The development cycle of SyPaB involves five parts (Fig. 1-1) (Zhang 2010): (i) pathway reconstruction, which is the central point of SyPaB, (ii) enzyme selection, which is in principle based on its catalytic reaction and substrate specificity, (iii) enzyme engineering, by which to

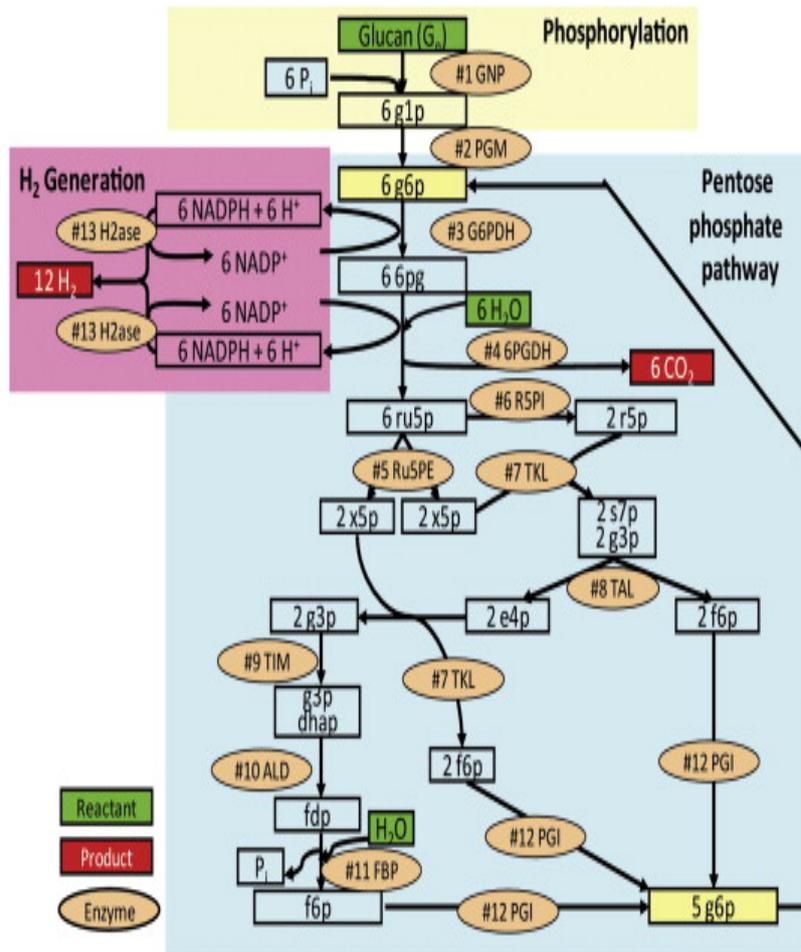
enhance enzyme's catalytic efficiency for industrial applications such as thermostability (Giver et al. 1998; Liu et al. 2009), cofactor preference, (iv) enzyme production, which means producing low-cost recombinant protein in high cell density fermentation (Barnard et al. 2004; Hartley 2006) and low-cost enzyme purification by scalable purification methods such as precipitation (Banki et al. 2005) or adsorption/desorption (Hong et al. 2008), and (v) process engineering, which includes enzyme immobilization, cofactor stabilization, in situ product separation, reactor engineering, and so on.



**Fig. 1-1.** The cycle of cell-free synthetic pathway biotransformation (SyPaB) (Zhang, 2012c).

SyPaB has been designed successfully for generating high-yield hydrogen (Zhang 2010) (Fig. 1-2). Hydrogen is a promising future clean energy carrier with high-energy utilization efficiency. The production of hydrogen from low-costly abundant biomass is a shortcut for producing low-

cost hydrogen without net carbon emissions (Adams and Stiefel 1998; Salge et al. 2006; Zhang 2010). Fig. 1-2 shows these cell-free synthetic enzymatic pathways generate nearly 12 moles of hydrogen per mol of glucose unit of starch or cellulosic materials. Moreover, SyPaB has also been designed for other applications such as the production of ethanol (Zhang 2010), polyols, the generation of electricity, the biological fixation of CO<sub>2</sub> (Zhang 2011c). One of the biggest challenges facing us is the high cost and instability of enzymes and cofactors. To overcome these obstacles, strong motivations have driven intensive efforts in discovering, engineering, and producing thermostable enzymes. Especially after the concept of SyPaB is accepted, more and more research will focus on cloning, expression, purification, characterizing and producing thermostable building blocks associated with catabolic pathways (Wang and Zhang 2009) and development biomimic cofactors analogs and their associated enzymes (Ansell et al. 1999; Ryan et al. 2008). In general, cell-free SyPaB can be applied to produce not only low-value biofuels but also high-value products, such as fine chemicals, non-natural carbohydrates, proteins, and glycoproteins. Therefore, SyPaB will have many significant applications in biofuel production in the near future.



**Fig. 1-2.** The synthetic pathway for conversion of starch and water to hydrogen and carbon dioxide. The abbreviations are: PPP, pentose phosphate pathway; glp, glucose 1-phosphate; g6p, glucose 6-phosphate; 6pg, 6-phosphogluconate; ru5p, ribulose 5-phosphate; r5p, ribose 5-phosphate; x5p, xylose 5-phosphate; s7p, sedoheptulose 7-phosphate; e4p, erythrose 4-phosphate; g3p, glyceraldehyde 3-phosphate; dhap, dihydroxyacetonephosphate; f6p, fructose 6-phosphate; fdp, fructose 1,6-bisphosphate; and Pi, inorganic phosphate. The enzymes are: #1 glucan phosphorylase, #2 phosphoglucomutase; #3 g6p dehydrogenase, #4 6-phosphogluconate dehydrogenase, #5 ribulose 5-phosphate epimerase, #6 phosphoribose isomerase, #7 transketolase, #8 transaldolase, #9 triose phosphate isomerase, #10 aldolase, #11 fructose 1,6-bisphosphatase, #12 phosphoglucose isomerase, and #13 hydrogenase (Zhang et al. 2007).

### **1.3 Thermostable enzymes**

The application of thermostable enzymes as biocatalysts is receiving more and more industrial and biotechnological interests. An enzyme is called thermostable when a high defined unfolding (transition) temperature ( $T_m$ ), or a long half-life at a selected high temperature, is observed (Turner et al. 2007). Thermostable enzyme's strong thermostabilities make them more suitable for the harsh conditions required in the industry production. The thermostability of these enzymes is an intrinsic property determined by many factors, such as: (i) amino acid composition, (ii) hydrophobic interactions, (iii) aromatic interactions, ion pairs, and salt bridges, (iv) metal binding, (v) substrate stabilization. Thermostable enzymes are receiving significant attention and they have many applications already, because they possess a number of important advantages over their mesophilic counterparts: (i) these enzymes have higher specific activity and higher stability; (ii) these enzymes have higher reactivity at high temperature; (iii) microbial contamination risks are significantly decreased; (iv) these enzymes can be stored at room temperature without inactivation.

Various bacterial, archaea, and fungi have received considerable attention as potential sources for thermostable enzymes. However, thermostable enzymes are produced both by thermophilic and mesophilic microorganisms. Besides tons of researches focusing on thermostable enzymes expression, purification and characterization, there are many researches regarding the structural basis of thermostability. Enzyme engineering strategies that have been used to introduce disulfide bridges, aromatic residues, or salt bridges to mesophilic or moderately thermophilic enzymes have succeeded in enhancing thermostability (Georis et al. 2000; Yeoman et al. 2010). One of the early successful commercialized examples was analytical use of a thermostable

enzyme, *Taq*-polymerase, in polymerase chain reactions (PCR) for amplification of DNA, and a number of other DNA modifying enzymes from thermophilic sources have been commercialized in this area (Podar and Reysenbach 2006; Satyanarayana et al. 2005). Efficient cellulose hydrolysis requires the concerted actions of three different classes of enzymes, including endoglucanases and exoglucanases, which operate on the solid: liquid interface and  $\beta$ -glucosidases which operate on the soluble degradation products of cellulose. Thermostable endoglucanases have been isolated from a number of thermophilic bacteria, and archaea as well as mesophilic and moderately thermophilic filamentous fungi, such as GH family 12 endoglucanase from *Pyrococcus furiosus* (Bauer et al. 1996), endoglucanase from *Thermotoga neapolitana* (Bok et al. 1998). Thermostable exoglucanases and  $\beta$ -glucosidases have been reported to be isolated and characterized from numerous strains such as *Clostridium thermoocellum* (Zverlov et al. 2002), *Aspergillus phoenicis* (Zeng and Zhang 1989), respectively.

In general, thermophiles and especially thermophilic enzymes have to date gained a great deal of interest both as analytical tools, and as biocatalysts for application in large scale. Especially with a paradigm shift in industry moving from fossils towards renewable resource utilization, the need of biocatalysts is predicted to increase, and certainly there will be a continued and increased need of thermostable biocatalysts in the future.

## 1.5 Goals and significance

This study contains two parts:

- (1) To purify a recombinant ribose 5-phosphate isomerase from *Thermotoga maritima* MSB8 with a low-cost purification method and to characterize its properties.

The recombinant ribose 5-phosphate isomerase was prepared easily with one-step heat treatment method and its robust thermostability made it better suited for the hydrogen production through the *in vitro* cell-free synthetic pathway (SyPaB).

- (2) To express and purify two fusion proteins containing Glucose-6-phosphate dehydrogenase (G6PDH) and Diaphorase (DI) from *Geobacillus stearothermophilus* 10 with two different spatial orientations and to study substrate channeling of two fusion proteins.

The recombinant fusion proteins G6PDH-DI and DI-G6PDH were purified through his-tag and the substrate channeling between G6PDH and fusion protein G6PDH-DI was observed.

## **2. Thermophilic *Thermotoga maritima* ribose-5-phosphate isomerase RpiB:**

### **Optimized heat treatment purification and basic characterization**

#### **2.1 Abstract**

The open reading frame TM1080 from *Thermotoga maritima* encoding ribose-5-phosphate isomerase type B was cloned and over-expressed in *E. coli* BL21 (DE3). After optimization of cell culture conditions, more than 30% of intracellular proteins were soluble recombinant RpiB. High-purity RpiB was obtained by heat pretreatment through its optimization in buffer choice, buffer pH, as well as temperature and duration of pretreatment. This enzyme had the maximum activity at 80 °C and pH 6.5-8.0. Under its suboptimal conditions (60°C and pH 7.0),  $k_{cat}$  and  $K_m$  values were 540 s<sup>-1</sup> and 7.6 mM, respectively; it had a half lifetime of 71 h, resulting in its turnover number of more than  $2 \times 10^8$  mol of product per mol of enzyme. This study suggests that it is highly feasible to discover thermostable enzymes from exploding genomic DNA database of extremophiles with the desired stability suitable for *in vitro* synthetic biology projects and produce high-purity enzymes at very low costs.

#### **2.2 Introduction**

Enzyme-based biocatalysis has become an attractive alternative to chemical catalysis because of its higher reaction selectivity and more modest reaction conditions (Schmid et al., 2001; Blumer-Schuette et al., 2008). But most enzymes are not suitable for industrial applications due to their relatively poor stability and biocatalysis re-use. The former can be addressed by protein engineering (Liu et al., 2009; Giver et al., 1998), enzyme immobilization (Betancor et al., 2008; De Cordt et al., 1998), utilization of stable enzymes from extremophilic microorganisms, or their

combinations. The latter can be solved through enzyme immobilization. For example, immobilized thermostable glucose isomerase has been used in the food industry to convert glucose to fructose at ~60°C for several months before its conversion (Bhosale et al., 1996).

*E.coli* is a common prokaryotic microorganism for genetic manipulation and for the production of recombinant proteins because of its fast cell growth in inexpensive media, rapid accumulation of cellular mass, amenability to high cell-density fermentation, simple scale-up, and relatively simple protein purification (Baneyx et al., 2004). But *E.coli* often produces recombinant protein in the form of insoluble, inactive inclusion bodies. A number of approaches have been explored to improve the expression of soluble recombinant proteins in *E.coli*. With regard to the expression vector, the heterologous protein could be fused with a protein-folding partner (e.g. thioredoxin, cellulose-binding module) or with a secretory protein fragment (e.g. outer-membrane protein A) that aids protein folding in a less-reducing periplasmic environment (LaVallie et al., 1993; Baneyx et al., 2004). Expression hosts can be chosen according to different approaches, such as (i) mitigating codon bias in a host containing a second plasmid expressing the *E.coli* rare tRNA genes (Gustafsson et al., 2004), (ii) enhancing protein folding in a host co-expressing folding modulators, such as chaperons (Martinez-Alonso et al., 2007), (iii) decreasing formation of disulfide bond in some special host's cytoplasm (Sorensen et al., 2005), and (iv) repressing basal expression of a toxic protein in a host with a repressor (Hassan et al., 2009). In addition, cultivation conditions, such as expression temperature, medium composition, timing of induction, inducer concentration, and inducer type, can be optimized for over-expression of a soluble protein. But over-expression of soluble heterologous proteins, especially

for hyperthermophilic ones, in *E.coli* still remains on a trial-and-test stage (Esposito et al., 2006; Adams et al., 1998).

The production of biofuels and biochemical from renewable sugars would decrease net greenhouse gas emissions, create new biomanufacturing jobs, and enhance national energy security (Zhang 2011b). Numerous biofuels have been produced based on different key metabolites through natural or non-natural biochemical pathways, for example, ethanol production from pyruvate (Shaw et al. 2008), non-fermentative alcohols from keto acids (Atsumi et al. 2008), fatty acid-derived hydrocarbons from malonyl-CoA (Liu et al. 2010), isoprenoid-derived hydrocarbons from isopentenyl pyrophosphate (Klein-Marcuschamer et al. 2007). On a basis of the pentose phosphate pathway, theoretical yield hydrogen (i.e., 12 mol of dihydrogen per mol of glucose and water) was accomplished by in vitro enzyme cocktails (Ye et al. 2009). Hydrogen produced from renewable sugar is widely believed to be the best future biofuel based on its production energy efficiency (Huang and Zhang 2011) and utilization efficiency (Zhang 2011b). However, its large-scale production requires low-cost production of stable enzymes.

In vitro synthetic pathway biotransformation (SyPaB), a new direction of synthetic biology, is an emerging low-cost biomanufacturing platform due to its unique advantages, such as high product yield, fast reaction rate, easy access and control, broad reaction condition (Eric Hodgman and Jewett 2012; Swartz 2011; Zhang 2010; Zhang 2011c). To produce low-value biofuels by cascade enzyme factories, one of the prerequisites is highly-stable enzyme building blocks with total turn-over number (TTN) values of more than  $10^7$ - $10^8$  mol of product per mol of enzyme (Wang et al. 2011; Zhang 2010; Zhang 2011c). Therefore, the discovery of stable enzymes from

extremophiles, enzyme engineering, and enzyme immobilization are among top priorities for the preparation of build blocks for in vitro synthetic biology projects (Zhang 2010; Zhang 2011c).

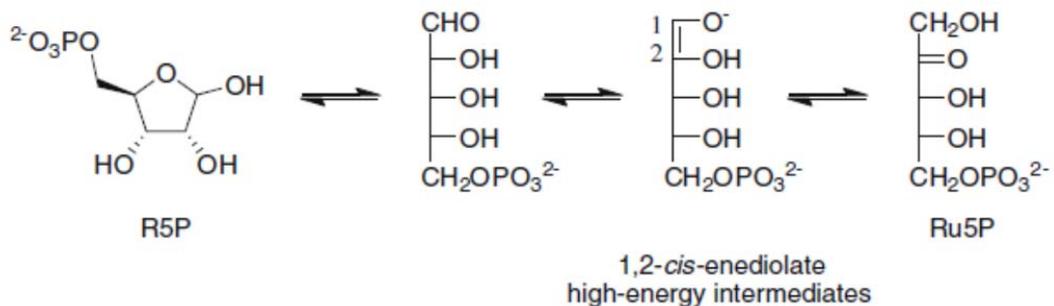
Ribose-5-phosphate isomerase (Rpi, EC 5.3.1.6) is responsible for catalyzing the reversible conversion between an aldose phosphate -- ribose-5-phosphate (R5P) and a ketose phosphate -- ribulose-5-phosphate (Ru5P) in the non-oxidative pentose phosphate pathway, the Calvin cycle, and in the process of photosynthesis (Ishikawa et al. 2002; Mariano et al. 2009). The isomerization reaction catalyzed is thought to proceed through a proton transfer mechanism between the two carbon atoms C-1 and C-2 of the substrates, concomitant with a proton transfer between the two oxygen atoms O-1 and O-2, and thus involves a 1,2-*cis*-eniol (ate) high-energy intermediate as depicted in Fig. 2-1 (Mariano et al. 2009). This enzyme was first described from yeast (Horecker et al. 1951) and then further studied in *Escherichia coli* (Hove-Jensen and Maigaard 1993). Mutant studies in *E. coli* have shown that Rpi is important to bacterial growth (Sorensen and Hove-Jensen 1996). Another medical research has revealed that a deficiency in human Rpi can cause the destruction of myelin sheaths, which in turn leads to brain abnormalities (Huck et al. 2004; Roos et al. 2005). Two completely unrelated RpiA and RpiB have been identified. RpiA is the most common, occurring in all three kingdoms of life, including most eukaryotic organisms, fungi and some bacteria; while RpiB has so far only been found in the genomes of some bacterial and protozoa (Zhang et al. 2003). RpiA is a highly conserved enzyme with no primary sequence similarities to other known protein families (Hove-Jensen and Maigaard 1993). RpiB belongs to the RpiB-LacAB family of enzymes, including RpiB as well as both LacA and LacB subunits of the galactose 6-phosphate isomerase (Sorensen and Hove-Jensen 1996). The open reading frame (ORF) TM1080 was designated to be a ribose-

5-phosphate isomerase B (RpiB) and its crystalline structure has been determined (Xu et al. 2004).

In addition to high-level expression of recombinant proteins in *E. coli*, it is vital to purify recombinant proteins at low costs. Although affinity chromatographic techniques are the most popular protein purification technology in laboratories, they are too costly for the production of industrial bulk enzymes with selling prices of tens of dollars per kg of dry weight enzyme. Several low-cost scalable protein purification technologies have been developed, such as heat treatment, ammonia precipitation, self-cleaving aggregation (Banki et al. 2005), one-step adsorption and immobilization (Myung et al. 2011). Among them, heat treatment may be the most promising because it does not involve any chemical reagent, protein tag, and adsorbent. Its application, however, was limited to thermostable enzyme production in mesophilic hosts like *E. coli*. During the process, crude cell extract containing the heat-resistant recombinant protein was heated. The denatured *E. coli* cellular proteins can be removed by centrifugation. For example, a recombinant *Thermus aquaticus* DNA polymerase (*Taq* polymerase) produced in *E. coli* was often purified by heat treatment in academic laboratories. However, there lacks detailed study of heat treatment purification for thermostable proteins.

Hyperthermophiles and thermophiles are great resources for isolation and discovery of thermostable enzymes for potential applications in molecular biology and industrial biocatalysis (Myung et al. 2010; Vieille and Zeikus 2001; Wang and Zhang 2009b). *Thermotoga maritima* is an anaerobic, rod-shaped bacterium, originally isolated from geothermally heated marine sediment. *Thermotoga maritima* is a rod-shaped bacterium, originally isolated from geothermal

heated marine sediments or hot spring. This thermophilic organism has an optimum growth temperature of  $\sim 80^{\circ}\text{C}$  and can utilize many simple and complex carbohydrates, including glucose, sucrose, starch, and xylan (Myung et al, 2010; Liu L et al, 2009). *T. maritima* is regarded as an invaluable source of intrinsically thermostable enzymes (Wang and Zhang 2009). The open reading frame (ORF) TM1080 was designated to be a ribose-5-phosphate isomerase B (RpiB) and its crystalline structure has been determined (Xu et al. 2004), However, its basic biochemical properties were not characterized yet. In this study, the ORF TM1080 encoding RpiB was cloned and over-expressed in *E. coli*. High-purity RpiB was obtained through optimized heat treatment. Its basic properties were characterized for the first time.



**Fig. 2-1.** Isomerization reactions of ribose-5-phosphate (R5P) to ribulose 5-phosphate (Ru5P). (Mariano et al, 2009).

## 2.3 Materials and methods

### 2.3.1 Chemicals and strains

All chemicals were reagent grade, purchased from Sigma-Aldrich (St. Louis, MO, USA) and Fisher Scientific (Pittsburgh, PA, USA), unless otherwise noted. Ribose-5-phosphate and ribulose-5-phosphate were purchased from Sigma-Aldrich. Restriction enzymes, Taq DNA polymerase, T4 ligase, and a broad range protein marker (2-212 kDa) were purchased from New England Biolabs (Ipswich, MA, USA). The *T. maritima* MSB8 genomic DNA was purchased from the American Type Culture Collection (Manassas, VA, USA). *E. coli* DH5 $\alpha$  was used as a host for recombinant DNA manipulation; *E. coli* BL21 Star (DE3) (Invitrogen, Carlsbad, CA, USA) was used as a host strain for recombinant protein expression. The Luria-Bertani (LB) medium was used for *E. coli* cell growth and recombinant protein expression. Ampicillin (100  $\mu$ g/mL) was added in the *E. coli* media. The oligonucleotides were synthesized by Integrated DNA Technologies (Coraville, IA, USA).

### 2.3.2 Protein expression plasmid construction

The DNA fragment containing the ORF TM1080 (GenBank accession No AE0005120) was amplified by PCR from genomic DNA of *T. maritima* MSB8 by using a forward primer (5'-GCA TAC CAT ATG ATG AAG ATC GCT ATT GCA TCG-3', NdeI site underlined) and a reverse primer (5'-CCT CAG CTC GAG TTA AAC CTC ATC GAT CTT TCT G -3', XhoI site underlined). PCR conditions were as the following: initial denaturation (5 min at 94°C), 30 cycles of denaturation (30 s at 94°C, annealing 30 s at 55°C, and elongation 60 s at 72°C), and a final extension step (10 min at 72°C). After *Nde*I and *Xho*I digestion of the PCR product and

plasmid pET20b, the ligated product was transformed to *E. coli* DH5 $\alpha$ , yielding plasmid pET20b-rpiB. The DNA sequence of pET20b-rpiB was validated by sequencing by Virginia Bioinformatics Institute (Blacksburg, VA).

### **2.3.3 Expression and purification of recombinant protein**

The protein expression plasmid pET20b-rpiB was transformed into the strain *E. coli* BL21 (DE3) for yielding *E. coli* BL21 (pET20b-rpiB). The freshly-prepared strain *E. coli* BL21 (pET20b-rpiB) was cultivated in two hundred and fifty milliliters of the LB medium supplemented with 100  $\mu$ g/mL ampicillin in 1-L Erlenmeyer flasks at a rotary shaking rate of 220 rpm at 37°C. The recombinant protein expression was induced by adding isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (0.1 mM, final) when A<sub>600</sub> reached 0.6~0.8. The cell cultures were incubated at 18°C for 16 h. The cells were harvested by centrifugation at 4°C, washed once in a 50mM Tris-HCl buffer (pH 7.0), re-suspended by ~30 mL of 50 mM Tris-HCl buffer (pH 7.0). The cell pellets were lysed in an iced bath by ultra-sonication by Fisher Scientific Sonic Dismembrator Model 500 (3-s pulse, total 90 s, at 50 % amplitude). After centrifugation at 10,000 rpm for 20 min, the supernatant of the cell lysate was used for protein purification by heat treatment. The pellets of the cell lysate were washed in 10 mL of 50 mM Tris-HCl buffer (pH 7.0). After centrifugation and re-suspension in 100  $\mu$ L of the SDS-PAGE protein loading buffer, 10  $\mu$ L of the boiled sample was analyzed by SDS-PAGE. RpiB was purified by heating the supernatant which was recovered after the centrifugation of the cell culture. The samples were heated 40 min at 80°C and then centrifuged at 14,000 rpm for 8 min. The supernatant containing high-purity RpiB was analyzed by SDS-PAGE.

### **2.3.4 SDS-PAGE and determination of protein concentration**

The protein samples were analyzed by SDS-PAGE with 12 % separate gel and 5% stacking gel. The gel was stained by the Bio-Rad Coomassie Blue 250 kit. Protein concentration was measured by the Bio-Rad Bradford protein kit with bovine serum albumin as a reference protein.

### **2.3.5 Enzyme activity assay**

One unit (U) of enzyme activity was defined as the amount of enzyme required to produce one  $\mu\text{mol}$  of ribulose 5-phosphate per min from ribose-5-phosphate. Ribose phosphate isomerase activity was assayed by a modified Dische's cysteine-carbazole method (Dische and Borenfreund 1951; Hove-Jensen and Maigaard 1993). The enzyme (20  $\mu\text{L}$ , the final enzyme concentration in the reaction mixtures ranged between 2.5 and 8  $\text{ng}/\mu\text{L}$ ) was mixed with 180  $\mu\text{L}$  of a pre-warmed substrate solution containing 5 mM ribose 5-phosphate in 50 mM Tris-HCl buffer (pH 7.0). After 5 min incubation at 65°C, 20  $\mu\text{L}$  of 2 M HCl was added (final concentration is 0.2 M) to stop the reaction and then the mixture was put in an iced water bath. One mL of 66 % (vol/vol)  $\text{H}_2\text{SO}_4$  containing 35  $\mu\text{L}$  of 0.12 % ethanol-dissolved carbazole and 35  $\mu\text{L}$  of 1.5 % (wt/vol) cysteine chloride was added into each tube and mixed by vortexing. After incubation at 37°C for 30 min, the absorbency was measured as 540 nm.

### **2.3.6 Optimization of RpiB reaction conditions**

The optimum pH of RpiB was determined in a series of buffers at 65°C for 5 min. The buffers were 100 mM sodium dihydrogen phosphate-citric acid (pH 4.6, 4.8, 5.0, 5.2, 6.4, 7.0, 7.6 and 8.0), Tris-HCl (pH 7.5, 8.0, 8.5, and 9.0), as well as glycine-NaOH (pH 8.8, 9.6 and 10.6). The optimum temperature was tested at different temperatures (30 to 99 °C) in 50 mM Tris-HCl buffer (pH 7.0) for 5 min. The kinetic parameters of RpiB were determined based on initial

reaction rate at the first 2 min. The reactions were conducted in a 50 mM Tris-HCl buffer (pH 7.0).

### **2.3.7 Thermostability**

Thermostability of the purified RpiB was studied by incubating the different concentration enzyme (0.1, 0.01, and 0.001 mg/mL) in 50 mM Tris-HCl (pH 7.0) at 60°C, 70°C, and 80°C. After heat incubation, the residual RpiB activity was measured according to the RpiB activity assay as described above. Each result was an average of at least three repetitions.

## **2.4 Results**

### **2.4.1 Overexpression of RpiB**

The 456-bp ORF TM1080 from *T. maritima* encoding a putative RpiB was amplified by PCR. After digestion with *NdeI* and *XhoI*, the PCR product was inserted into pET20b vector, yielding a recombinant protein expression vector pET20b-rpiB. The sequence of the cloned RpiB gene was confirmed by DNA sequencing.

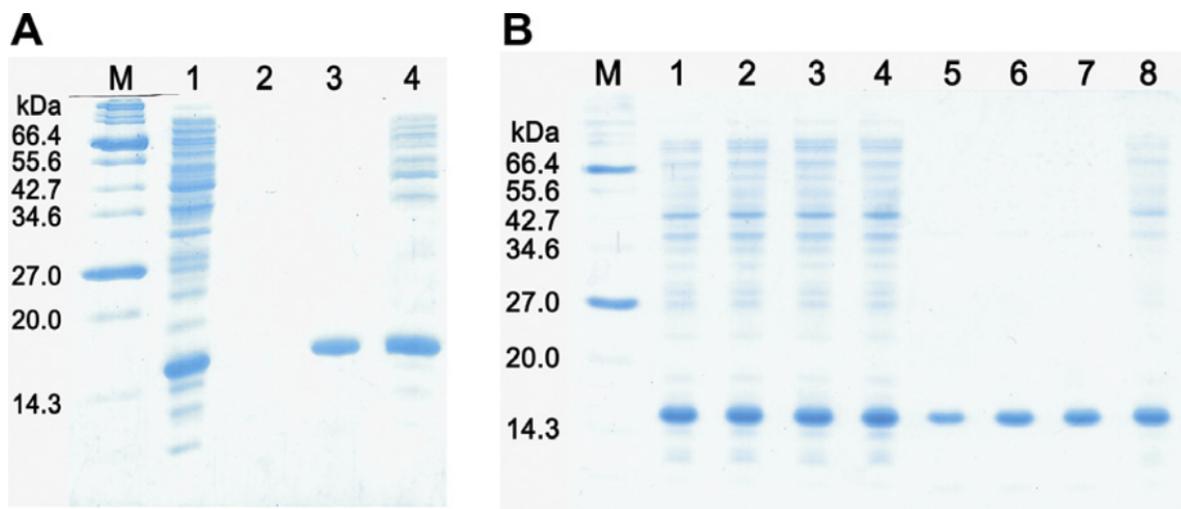
Expression conditions of RpiB in *E. coli* BL21 (DE3) were optimized by varying several factors: cultivation temperature (18 and 37°C), IPTG concentration (0.01, 0.1, and 1 M), and different induction time. It was found that the optimal protein expression conditions was 0.1 M IPTG, induction at  $A_{600} = 0.6-0.8$ , and 18°C for cultivation for 16 h. Under this condition, the target recombinant protein was highly soluble (Fig. 2-2A, lane 1), accounting for approximately 30%

of the overall soluble cellular protein. No obvious protein bands were observed for the pellets of the *E. coli* cell lysate (Fig. 2-2A, lane 2), suggesting that no inclusion body of RpiB was formed.

#### **2.4.2 Optimized purification of RpiB**

In our initial attempts, we planned to purify this protein based on its His-tag. Due to its high expression levels and thermostability, we investigated the feasibility of simple protein purification through heat treatment. Heat treatment has been widely-used as the first step for purifying thermostable enzymes produced in mesophilic hosts because most mesophilic enzymes denatured and aggregated during the heat treatment, for example, *Taq* polymerase. To our limited knowledge, there are no reports expanding its optimization by testing different buffers and pHs. First, it was found that the buffer used had large influences on the purity of the target protein. As shown in Fig. 2-2A, heat treatment was conducted in 50 mM Tris-HCl buffer (pH 7.0) at 80°C for 40 min, yielding high-purity RpiB (Fig. 2-2A, lane 3), much better than in 50 mM HEPES buffer at the same pH (Fig. 2-2A, lane 4). Second, it was found that pH of Tris buffer from 6.5, 7.0, 7.5 to 8.5 greatly influenced the purity of the targeted protein (Fig. 2-2B, lanes 5-8). When pH was high (e.g., 8.5), the supernatant protein contained numerous impurities (Fig. 2-2B, lane 8). The optimal pH of the Tris-HCl buffer was 7.0. Also, heat treatment temperatures from 60 to 95°C and treatment duration were optimized (data not shown). In general, longer heat treatment time and/or higher pretreatment temperature resulted in higher purity of the targeted protein, examined by SDS-PAGE. However, too long treatment time and/or too high temperature also deactivated RpiB, resulting in lower specific activity (data not shown). The optimal purification conditions were the preparation of the cell lysate in 50 mM Tris-HCl buffer (pH 7.0) followed by heat treatment at 80°C for 40 min. After centrifugation, the purity of

the purified RpiB in the supernatant was more than 95 % (Fig. 2-2A, lane 3). Molecular mass of RpiB without any tag was 15,867 based on its deduced amino acid sequence, close to the estimated value from the SDS-PAGE analysis. Approximately 11.2 mg of RpiB was purified from 250 mL of the cell culture with a purification yield of 31 % (Table 1). This enzyme had a specific activity of 290 U/mg on ribose-5-phosphate at pH 7.0 and 50°C.



**Fig. 2-2.** SDS-PAGE analysis of RpiB purification in 50 mM Tris-HCl buffer (pH 7.0) and 50 mM HEPES buffer (pH 7.0) by heat treatment at 80 °C for 40 min. M, marker; lane 1, the supernatant of the soluble cell lysate containing RpiB; lane 2, the pellets of the cell lysate; lane 3, the supernatant of heat-treated cell lysate in 50 mM Tris-HCl buffer (pH 7.0); lane 4, the supernatant of heat treated cell lysate in 50 mM HEPES buffer (pH 7.0) (A). SDS-PAGE analysis of RpiB purification in 50 mM Tris-HCl buffer with different pH. M, marker; lane 1-4, the cell lysate containing RpiB at pH 6.5, 7.0, 7.5, and 8.5; lane 5-8, the supernatant containing RpiB at pH 6.5, 7.0, 7.5 and 8.5 (B).

**Table 2-1.** The RpiB purification from 250 mL of the *E. coli* cell culture.

<b>Fraction (mg)</b>	<b>Vol. (mL)</b>	<b>Protein (mg/mL)</b>	<b>Total Protein (mg)</b>	<b>Sp. act. (U/mg)</b>	<b>Total act. (U)</b>	<b>Yield (%)</b>
Cell lysate	35	1.57	55	187	10296	100
Purified RpiB	34	0.33	11.2	290	3254	31.7

Sp. act., specific activity was measured in 50 mM Tris-HCl (pH 7.0) at 50°C.

### 2.4.3 Basic properties

Previous research suggested that RpiB was the metal-independent aldose-ketose isomerase (Stern et al. 2007). Therefore, the effect of metal ions on its activity was not investigated. The pH profile of RpiB activity at 50°C is shown in Fig. 2-3A. It had the maximum activities in the sodium dihydrogen phosphate-citric acid buffer at pH 6.5-8.0 and in 50 mM Tris-HCl buffer pH 7.0-9.0. Because pH of the Tris-HCl buffer decreases with increasing temperature (i.e., ~0.03 per °C), this enzyme had optimal pH ranging from 6.5-8.0. This enzyme remained ~65% activity from pH 8.8 to 11.0 in the glycine-NaOH buffer. At low pH, it exhibited very low activities.

The temperature effect on RpiB activity from 30 to 99°C was measured in a 50 mM Tris-HCl buffer (pH 7.0). Its activity increased significantly when the temperature increased until 80°C (Fig. 2-3B). The optimum temperature was 80°C. The activity at 30°C was 25 % of that at 80°C. The enzyme activity decreased after 80°C. It was still highly active when the temperature reached 99°C, 90% activity of that of 80°C. This enzyme showed relatively high activity (above 50% of its maximum activity) over a large temperature range of 40-99°C.

The kinetic parameters of RpiB were examined at different temperature of 60, 70 and 80°C (Table 2-2). The  $k_{cat}$  value increased from  $540 \pm 6$  to  $1192 \pm 88 \text{ s}^{-1}$  when the temperatures increased from 60 to 80°C. The  $K_m$  value ranged around 7.6 to 12.1 mM (Table 2-2).

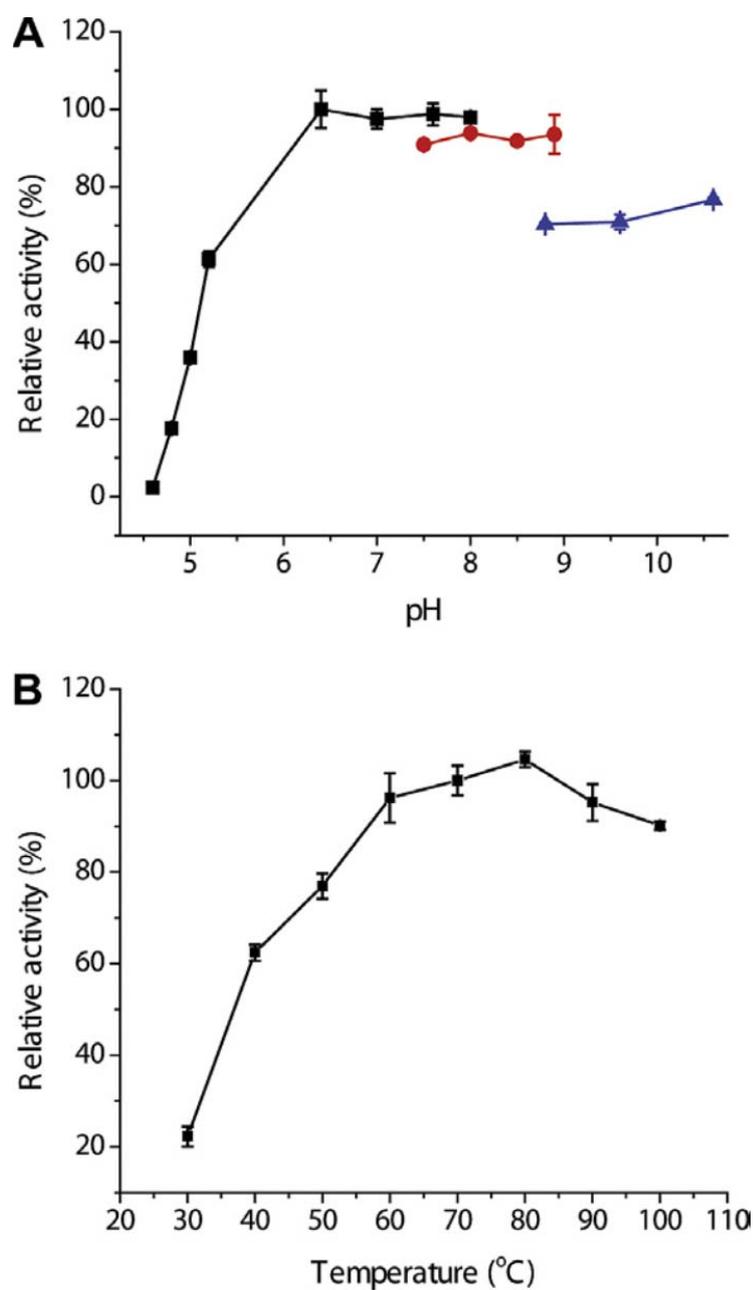
**Table 2-2.** The kinetic characteristics of RpiB based on ribose-5-phosphate as the substrate.

Temp. (°C)	Conc. (mg/mL)	$k_{cat}$ ( $\text{s}^{-1}$ )	$K_m$ (mM)	$k_{cat}/K_m$ ( $\text{mM}^{-1} \text{ s}^{-1}$ )	$k_d$ ( $\text{h}^{-1}$ )	$T_{1/2}$ (h)	TTN (mol/mol)
60	0.10	$540 \pm 6$	$7.6 \pm 0.2$	71.1	0.009	71	$2.2 \times 10^8$
70	0.10	$767 \pm 35$	$10.3 \pm 1.0$	74.5	0.013	53	$2.1 \times 10^8$
	0.010				0.110	6.3	$2.5 \times 10^7$
80	0.10	$1192 \pm 88$	$12.1 \pm 1.9$	98.5	0.108	6.4	$3.9 \times 10^7$
	0.010				0.189	3.6	$2.3 \times 10^7$

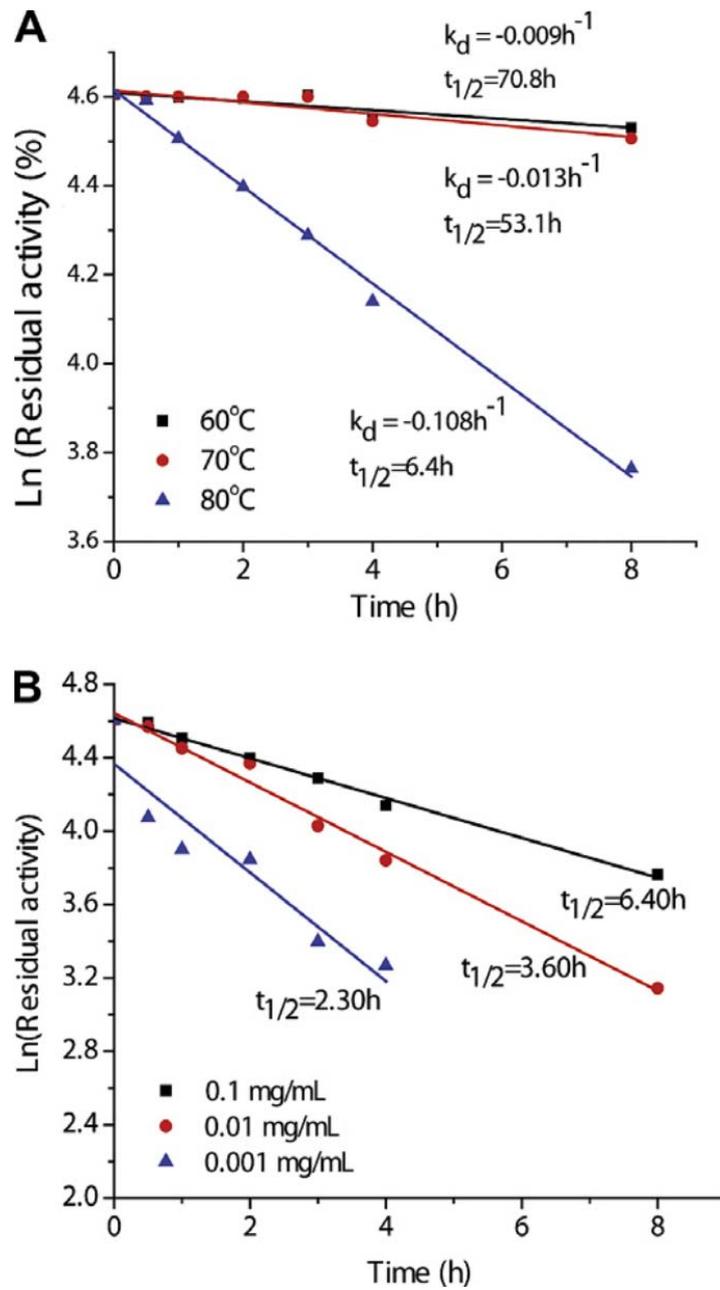
#### 2.4.4 Thermostability

Thermostability of RpiB in 50 mM of Tris-HCl buffer (pH 7.0) strongly depended on temperature (Fig. 2-4A). RpiB at 0.1 mg/mL was very stable at 60 and 70°C but it deactivated at 80°C. The degradation constants were -0.009, -0.013 and -0.108  $\text{h}^{-1}$ , at 60, 70 and 80°C, respectively. That is, it had half lifetimes of 70.8, 53.1 and 6.4 h at 60, 70 and 80°C, respectively.

Also, the half lifetime of RpiB at 80°C greatly depended on its mass concentration from 0.1, 0.01 to 0.001 mg/mL (Fig. 2-4B). The half-life time of thermo-inactivation was about 6.4 h at 0.1 mg/mL but decreased to only 2.3 h at a low concentration of 0.001mg/mL. Under its suboptimal condition (pH 7.0, 60°C, and 0.10 mg/mL), this enzyme had the turn-over number (TTN) of  $2.2 \times 10^8$  mol of product per mol of enzyme, where TTN value can be calculated as  $TTN = k_{cat}/k_d$  (Myung et al. 2011; Rogers and Bommarius 2010).



**Fig. 2-3.** The profiles of pH (A) and temperature (B) of the purified RpiB. Black square symbols represent activities measured in the sodium dihydrogen phosphate–citric acid at pH 4.6, 4.8, 5.0, 5.2, 6.4, 7.0, 7.6 and 8.0; red cycle symbols represent activities in the Tris–HCl buffer at pH 7.5, 8.0, 8.5 and 9.0; and blue triangle symbols represent activities in the glycine–NaOH buffer at pH 8.8, 9.6 and 10.6.



**Fig. 2-4.** Thermo-inactivation of the purified RpiB (0.1mg/mL) at 60, 70 and 80 °C in 50 mM Tris-HCl buffer (pH 7.0). (A), and different concentration RpiB from 0.001 to 0.1 mg/mL at 80 °C in 50 mM Tris-HCl buffer (pH 7.0) (B).  $k_d$ , degradation of constant;  $T_{1/2}$ , half lifetime.

## 2.5 Discussion

Thermostable enzymes from thermophilic microorganisms are playing important roles in molecular biology and industrial applications. The production of recombinant proteins from thermophilic organisms in mesophilic hosts (e.g. *Escherichia coli*) can take the advantage of easy fermentation, over-expression and simple purification.

Heat treatment could be widely-used as the first step for protein purification. However, our study demonstrated to obtain high-purity RpiB through the optimization of heat treatment conditions – choice of buffer, buffer pH, treatment temperature, and treatment duration (Fig. 2-2). Therefore, it was feasible to obtain high-purity thermostable enzymes at very low costs.

The development in basic building blocks was urgently needed for the commercialization of *in vitro* synthetic biology projects (Eric Hodgman and Jewett 2012; Swartz 2011; Zhang 2010; Zhang 2011c). Economic analysis suggested that enzyme costs in cascade enzyme factories could be minimal when all of the enzymes have TTN values of more than  $10^7$ - $10^8$  mol of product per mol of enzyme (Huang and Zhang 2011; Wang et al. 2011; Zhang 2011c). In this study, another enzyme suitable for *in vitro* synthetic biology projects was obtained. As compared to other reported Rpis (Table 2-2), the *T. maritima* RpiB seemed to be a good building block, which had high activities at a broad temperature range and had great stability even without enzyme immobilization.

In summary, another thermostable building block for cell-free SyPaB projects- *T. maritima* RpiB was simply purified by optimized heat treatment and it exhibited high TTN values under its

**Table 2-3.** Comparison of ribose-5-phosphate isomerases from different sources in the literature.

<b>Organism</b>	<b>T<sub>opt</sub></b> <b>(°C)</b>	<b>Sp. act. (U/mg)</b> <b>(Condition)</b>	<b>Half lifetime</b> <b>(Condition)</b>	<b>Ref.</b>
<i>Arabidopsis thaliana</i>	37	NA	1 h (50°C)	(Xiong et al. 2009)
<i>Bacillus caldolyticus</i>	NA	52.7 (65°C, pH 6.5)	1 h (91°C)	(Middaugh and MacElroy 1976)
<i>Clostridium thermocellum</i>	65	25690 (65°C, pH 7.5)	96 h (50°C) 4.7 h (65°C)	(Yoon et al. 2009)
<i>Escherichia coli</i>	37	NA	2.2 min (60°C)	(Essenberg and Cooper 1975)
<i>Mycobacterium tuberculosis</i>	37	NA	NA	(Mariano et al. 2009)
<i>Pichia jadinii</i>	50	356 (50°C, pH 7.5)	NA	(Domagk et al. 1973)
<i>Thermotoga maritima</i>	80	290 (50°C, pH 7.0)	70.8 (60°C) 53.1 h (70°C) 6.4 h (80°C) 1.2 h (90°C)	This study
<i>Thiobacillus thioparus</i>	NA	544 (40°C, pH 6.5)	1 h (57°C)	(Middaugh and MacElroy 1976)

T<sub>opt</sub>, optimal temperature of the enzyme; Sp. act., specific activity; NA, not available.

suboptimal conditions suggesting great potential for decreasing protein costs associated with its production, separation, and use. Along with the previously purified thermostable enzymes *Geobacillus stearothermophilus* Glucose-6-phosphate dehydrogenase and *Geobacillus stearothermophilus* diaphorase (Zhu et al. 2012), it suggested that discovery of thermostable building blocks from (hyper) thermophilis are very operative.

### **Acknowledgment**

This work was supported by the College of Agriculture and Life Sciences Bioprocessing and Biodesign Research Center at Virginia Tech.

### **3. Spatial organization of heteromultimeric complex is important to restore multimeric glucose-6-phosphate dehydrogenase activity and channel substrate between two cascade enzymes**

#### **3.1 Abstract**

Glucose-6-phosphate dehydrogenase (G6PDH) and diaphorase (DI) were linked together into fusion proteins for the study of their activities and substrate channeling between the cascade enzymes. The presence of large-size DI greatly impaired G6PDH activity in two fusion proteins, especially when DI was located in the N-terminus because G6DPH must form multimers so to be highly active. The addition of a monomeric G6PDH greatly restored G6PDH activity in the fusion proteins. Furthermore, the mixture of G6PDH-DI and G6PDH exhibited substrate channeling between cascade enzymes while not for the mixture of DI-G6PDH and G6PDH, suggesting that the special spatial organization of enzyme network determined by using dynamic light scattering was vital to substrate channeling. The degree of substrate channeling decreased when the enzyme concentration increased because of longer traveling distance of the metabolite at lower concentration free enzyme mixture. The above study suggested that whether substrate channeling occurred to the fusion of two cascade enzymes were complicated, especially when multimeric enzymes were involved.

#### **3.2 Introduction**

Substrate channeling is a phenomenon that the product of one enzyme is rapidly utilized by another spatially proximate cascade enzyme without full equilibrium of the first enzyme's product in the bulk phase (Zhang 2011a). It often occurs when metabolic enzymes are packed tightly in the cytoplasm, e.g., pyruvate dehydrogenase complex. Due to proximate existence of

cascade enzymes and other biomacromolecules, reaction rates among enzyme complexes can be accelerated greatly, leading to enhanced cellular efficiency. Most of such intracellular enzyme complexes (metabolons) formed through non-covalent protein-protein interactions are transient and weak (Meyer et al. 2011). When cells are broken, such enzyme complexes are often disrupted, resulting in difficulties in identifying the existence of substrate channeling in vitro. The simplest way to facilitate substrate channeling between two cascade enzymes is the fusion of two cascade enzymes through a short peptide linker (Bulow et al. 1985; Iturrate et al. 2009; Ljungcrantz et al. 1989; Orita et al. 2007; Winter et al. 2012). For example, a direct fusion of galactosidase and galactose dehydrogenase was reported to facilitate galactose channeling (Ljungcrantz et al. 1989), while a contradictory report suggested that substrate channeling did not take place (Pettersson and Pettersson 2001). Numerous experimental factors, such as enzyme concentration, pH, substrate concentration, could influence whether substrate channeling occurs or is observed in vitro (Ljungcrantz et al. 1989). Substrate channeling has more potential advantages in addition to accelerating reaction rate. These potential advantages include: (1) protecting unstable intermediates and/or stabilize labile factors (Geck and Kirsch 1999; Rudolph and Stubbe 1995) (2) circumventing unfavorable equilibrium and kinetics imposed by bulk phase metabolite concentrations (Ovadi et al. 1994; Ushiroyama et al. 1992); (3) isolating intermediates from competing reactions; (4) conserving the scarce solvation capacity of the cell (Atkinson et al. 1969); (5) enhancing biocatalysis by avoiding unfavorable energetics of desolvating substrates (Dewar and Storch 1985); (6) reducing lag transients (times to reach steady-state response to a change in substrate concentration upstream in a coupled reaction path) (Easterby 1981; Westerhoff and Welch 1992); and (7) providing new means of metabolic regulation by modulation of enzyme association and increased sensitivities to regulatory signals (Kholodenko

et al. 1993); (8) conserving bond energy stored in oligosaccharides through substrate phosphorylation for extra ATP generation (Zhang and Lynd 2004; Zhang and Lynd 2005). Thus, studies of substrate channeling are important for a better understanding of metabolism.

The simple fusion strategy by linking two cascade enzymes is hypothesized not to work when one enzyme in the fusion protein cannot form its active multimeric form due to the adjacent other enzyme (Conrado et al. 2008; Iturrate et al. 2009). To our limited knowledge, however, there lack such reports. In-depth understanding of substrate channeling in fusion enzymes would be of importance to multi-enzyme one pot and synthetic biology projects for the circumvention of unfavorable equilibrium and kinetics imposed, protection of unstable substrates, forestallment of substrate competition among different pathways, regulation of metabolic fluxes, mitigation of toxic metabolite inhibition, and so on (Zhang 2011a).

Discovery and utilization of thermoenzyme from (hyper) thermophilic microorganisms is of great interest from numerous applications. *Geobacillus stearothermophilus* is a gram positive thermophilic (heat loving) bacteria characterized by an inner cell membrane and a thick cell wall. *G. stearothermophilus* is a rod shaped anaerobic found in thermophilic habitats like thermal vents. Many heat stable enzymes like xylanase for pulp treatment and thermolysin-like protease for production of artificial aspartame have been isolated from these thermophilic bacteria. *Geobacillus stearothermophilus* strain 10 is an isolated strain that was found in a hot spring in Yellowstone National Park and has been used in comparative analysis of thermophiles and mesophiles. It has an optimum growth temperature of 30 ~70°C. It is first described in 1920 as *Bacillus stearothermophilus*, but following a reclassification, it is officially a member of the

genus *Geobacillus*. *G. stearothermophilus* is constantly used in the biotech industry to test the success of sterilization cycles of equipment. Due to the bacterium's high resistance to heat, it is suitable Biological Indicator of microbe life after sterilization cycle.

Glucose-6-phosphate dehydrogenase (G6PDH) is the first enzyme in the pentose phosphate pathway that can convert glucose-6-phosphate (G6P) to 6-phosphogluconate (6PG) and cogenerate NADPH from the oxidized  $\text{NADP}^+$ . This enzyme is widely distributed and has been isolated from humans, animals, plants, and microorganisms. G6PDH in eukaryotes and many prokaryotes are either  $\text{NAD}^+$ - or  $\text{NADP}^+$ -dependent, generating reduced NADH or NADPH, respectively (Cosgrove et al. 1998). In cells NADH provides reducing power for catabolism, which NADPH is utilized in biosynthetic pathways. G6PDH can be classified into five categories based on coenzyme specificity (Levy 1979) as follows: (i)  $\text{NADP}^+$  specific G6PDHs, which can react only with  $\text{NADP}^+$ ; (ii)  $\text{NADP}^+$ -preferring G6PDHs, which can react with  $\text{NAD}^+$  under unusual conditions, but do not appear to utilize  $\text{NAD}^+$  physiologically; (iii) G6PDHs of dual nucleotide specificity, which have comparable  $\text{NAD}^+$ - and  $\text{NADP}^+$ - linked activities under physiological conditions; (iv)  $\text{NAD}^+$ -preferring G6PDHs, which can react with both coenzyme, but would be expected to react only with  $\text{NAD}^+$  under physiological conditions; and (v)  $\text{NAD}^+$ -specific G6PDHs, which can react only with  $\text{NAD}^+$ . In some organisms, however, a single dual nucleotide specific enzyme can catalyze both  $\text{NAD}^+$ - and  $\text{NADP}^+$ - linked reaction (Cosgrove et al. 1998). *Leuconostoc mesenteroides* glucose-6-phosphate dehydrogenase, for example, catalyzes the oxidation of glucose 6-phosphate by either  $\text{NAD}^+$  or  $\text{NADP}^+$  in vitro and in vivo (Olive et al. 1971). G6PDHs from *E. coli* reacts exclusively with  $\text{NADP}^+$ . G6PDHs from animal tissues are  $\text{NADP}^+$ -preferring, reacting only poorly with  $\text{NAD}^+$ . The  $K_m$  values of *Bacillus*

*stearothermophilus* G6PDH for  $\text{NADP}^+$  and  $\text{NAD}^+$  are 0.016 and 1.64 mM, respectively. From the  $K_m$  values, this enzyme should be classified as (ii), and perhaps under physiological conditions,  $\text{NADP}^+$  is a good electron acceptor for this enzyme (Okuno et al. 1985). Most G6PDHs form homo-multimers, with a predominance for dimers and tetramers, determined by electron microscopic observation, native gel, glutaraldehyde covalent linking in the aqueous solution followed by matrix-assisted laser desorption mass spectrometry (Farmer and Caprioli 1991). Multiple oligomeric forms of G6PDH are influenced by numerous experimental factors (Scott 1971; Sundaram et al. 1998).

Diaphorase (DI) is a ubiquitous class of flavin-bound enzyme capable of oxidizing NAD(P)H and transferring two electrons to an oxidized synthetic dye, such as 2,6-dichlorophenolindophenol (DCPIP, blue color) to its reduced form (colorless) (Chakraborty et al. 2008a). Based on the original functions, enzymes with diaphorase activity are classified into various enzyme categories: NAD(P)H dehydrogenase (quinone) (EC 1.6.5.2), NADPH dehydrogenase (EC 1.6.99.1), NADH dehydrogenase (EC 1.6.99.3), dihydrolipoyl dehydrogenase (EC 1.8.14), or nitric oxide synthase (EC 1.14.13.39) (Chakraborty et al. 2008a). Diaphorases are applicable in the colorimetric determination of NAD(P)H, and hence different dehydrogenase activities can be assayed when coupled with various dyes that act as hydrogen acceptors from NAD(P)H. Actually, various diaphorases are commercially available on the market, such as diaphorase enzyme from *Clostridium kluyveri* (Chakraborty et al. 2008b). Several bacterial species contain enzymes that utilize either NADH or NADPH to reduce a tightly or noncovalently bound FMN or FAD cofactor in their active site. A 636-bp DNA

fragment from the pyruvate dehydrogenase multi-enzyme complex of *G. stearothermophilus* was found to encode a thermophilic DI (Borges et al. 1990)

In this study, we studied substrate channeling between *G. stearothermophilus* G6PDH and DI by using the fusion strategy with its application in enzymatic fuel cells. Although the presence of DI in fusion proteins clearly disrupted the formation of multimeric G6PDH, resulting in drastic decreases in G6DPH activity. However, a mixture of G6PDH and a fusion protein G6PDH-DI not only restored G6PDH activity through the formation of heteromultimers but also facilitated substrate channeling between DI and G6PDH, especially at low enzyme concentrations.

### **3.2 Materials and methods**

#### **3.2.1 Chemicals and strains**

All chemicals were reagent grade, purchased from Sigma-Aldrich (St. Louis, MO, USA) and Fisher Scientific (Pittsburgh, PA, USA), unless otherwise noted. Restriction enzymes, *Taq* DNA polymerase, T4 ligase, and a broad range protein marker were purchased from New England Biolabs (Ipswich, MA, USA). The oligonucleotides were synthesized by Integrated DNA Technologies (Coraville, IA). *E. coli* DH5 $\alpha$  was used as a host for cell for DNA manipulation; *E. coli* BL21 Star (DE3) was used as a host strain for recombinant protein expression. The Luria-Bertani (LB) medium supplemented with 100  $\mu$ g/mL ampicillin was used for *E. coli* cell growth and recombinant protein expression.

### 3.2.2 Construction of recombinant plasmids

The sequences of primers for PCR are listed in Table 3-1. The 1485-bp DNA fragment containing the open reading frame (ORF) of the glucose-6-phosphate dehydrogenase (g6pdh) gene (GenBank accession number JQ040549) was amplified by the polymerase chain reaction (PCR) from the genomic DNA of *Geobacillus stearothermophilus* 10 (Bacillus Genetic Stock Center accession number 9A21) with a primer pair of g6pdh-IF and g6pdh-IR. The pET20b vector backbone was amplified with a primer pair of g6pdh-VF and g6pdh-VR by PCR. Both PCR products were cleaned using the Zymo Research DNA Clean & Concentrator Kit (Irvine, CA). Via the newly developed restriction enzyme-free, ligase-free and sequence-independent Simple Cloning technique (You et al. 2012b), the insertion fragment and vector backbone were assembled to DNA multimers by a prolonged overlap extension PCR (POE-PCR), and the PCR product was directly transformed into *E. coli* DH5 $\alpha$  cells, yielding the chimeric plasmid pET20b-gsg6pdh.

The 636-bp DNA fragment encoding diaphorase (DI, GenBank accession number JQ040550) was amplified by PCR by a primer pair of di-IF and di-IR based on the genomic DNA of *G. stearothermophilus* 10 as the template. The pET20b vector backbone of was amplified by a primer pair of di-VF and di-VR by PCR based on plasmid pET20b. Plasmid pET20b-Gsdi was assembled through Simple Cloning PCR (You et al. 2012b).

For the construction of plasmid pET20b-g6pdh-di expressing the fusion protein G6PDH-DI, by using plasmid pET20b-Gsg6pdh as the template, the *G. stearothermophilus* 10 glucose-6-phosphate dehydrogenase-encoding gene Gsg6pdh was amplified with PCR by using forward

primer g6pdh-di-IF and reverse primer g6pdh-di-Insrt-R. A ten amino acids linker (GGGSGGGGSG) was introduced by the reverse primer g6pdh-di-Insrt-R and in frame fused at the 3' terminus of the cloned Gsg6pdh. For simple cloning (You et al. 2012b) of Gsg6pdh into plasmid pET20b-Gsdi for the fusion expression of glucose-6-phosphate dehydrogenase and diaphorase, the vector backbone was amplified with plasmid pET20b-Gsdi as the template using primers g6pdh-di-VF and g6pdh-di-VR and the insert was amplified with the above linker added Gsg6pdh PCR product as the template using primers g6pdh-di-IF and g6pdh-di-IR. The above two PCR products were further subjected to the formation of multimers (You et al. 2012b) and transformation of *E. coli* to get the recombinant vector of pET20b-g6pdh-di. In the recombinant vector pET20b-g6pdh-di, Gsg6pdh was in frame fused at the 5' terminus of Gsdi and with a ten amino acids-encoding DNA linker between them. A His-tag was also in-frame fused at the 3'-terminus of the fusion construct.

For the construction of plasmid pET20b-di-g6pdh expressing the fusion protein DI-G6PDH, by using plasmid pET20b-Gsg6pdh as the template, the *G. stearothermophilus* 10 glucose-6-phosphate dehydrogenase-encoding gene Gsg6pdh was amplified with PCR by using forward primer di-g6pdh-Insert-F and reverse primer di-g6pdh-IR. A ten amino acids linker (GGGSGGGGSG) was introduced by the forward primer di-g6pdh-Insert-F and in frame fused at the 5' terminus of the cloned Gsg6pdh. For simple cloning of Gsg6pdh into plasmid pET20b-Gsdi for the fusion expression of diaphorase and glucose-6-phosphate dehydrogenase, the vector backbone was amplified with plasmid pET20b-Gsdi as the template using primers di-g6pdh-VF and di-g6pdh-VR and the insert was amplified with the above linker added Gsg6pdh PCR product as the template using primers di-g6pdh-IF and di-g6pdh-IR. The above two PCR

products were further subjected to the formation of multimers and transformation of *E. coli* to get the recombinant vector of pET20b-di-g6pdh. In the recombinant vector pET20b-di-g6pdh, Gsg6pdh was in frame fused at the 3' terminus of Gsdi and with a ten amino acids-encoding DNA linker between them. A His-tag was also in-frame fused at the 3'-terminus of the fusion construct.

All inserted sequences in the recombinant protein expression plasmids were validated by sequencing.

### **3.2.3 Recombinant protein expression and purification**

The plasmids were transformed into *E. coli* BL21 (DE3) for the expression of recombinant proteins. Two hundred and fifty mL of the LB medium in 1-L Erlenmeyer flasks was inoculated with the transformed *E. coli* cells and incubated at a rotary shaking rate of 220 rpm at 37°C until the absorbance at 600 nm reached between ~0.6 and 0.8. The expression of the recombinant protein was induced by adding isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (0.01mM, final concentration). The cultures were incubated at decreased temperature of 18°C for 16 h. The cells were harvested by centrifugation at 4°C. The cells were lysed by ultra-sonication. After centrifugation at 12,000 rpm for 15 min, 10  $\mu$ L of the supernatant was applied to SDS-PAGE for checking protein expression level. The supernatant of the cell lysate was applied to a column containing 2 mL of the Promega nickel-NTA resin (Madison, WI) and then the resin was pre-washed with 10 mL of 50 mM HEPES buffer (pH 7.5) containing 10 mM imidazole and 50 mM NaCl. The bound proteins were eluted by 5 mL of 50 mM HEPES buffer (pH 7.5) containing 50

mM imidazole and 50 mM NaCl. The enzyme was desalted by dialysis in a 50 mM HEPES buffer (pH 7.5) containing 5% glycerol.

#### **3.2.4 Enzyme activity assays.**

The activity of G6PDH was measured in a 50 mM HEPES buffer (pH 7.5), containing 100 mM NaCl, 2.5 mM glucose-6-phosphate (G6P), 2.5 mM NAD<sup>+</sup>, 5 mM MgCl<sub>2</sub>, and 0.5 mM MnCl<sub>2</sub> at 23 °C. An increase in the absorbance at 340 nm due to the formation of NADH was recorded using a UV spectrometer after 5 minutes reaction (Wang et al. 2011). One unit of activity is defined as the amount of G6PDH that forms 1 μmol of NADH per minute at 23°C. The activity of DI was measured in a 50 mM HEPES buffer (pH 7.5) containing 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.5 mM MnCl<sub>2</sub>, 0.16 mM NADH, and 0.1 mM dichlorophenolindophenol (DCPIP) at 23 °C. A decrease in the absorbance at 600 nm due to the reduction of DCPIP was recorded using a photospectrometer after 1 minute reaction (Chakraborty et al. 2008b). The activities of the fusion DI-G6PDH and G6PDH-DI proteins were assayed in a 50 mM HEPES buffer (pH 7.5) containing 100 mM NaCl, 2.5 mM NAD<sup>+</sup>, 5 mM MgCl<sub>2</sub>, 0.5 mM MnCl<sub>2</sub>, 2.5 mM G6P, and 0.1 mM dichlorophenolindophenol (DCPIP) at 23 °C. A decrease in the absorbance at 600 nm due to the reduction of DCPIP was measured using a photospectrometer after 1 minute reaction (Chakraborty et al. 2008b).

**Table 3-1.** Strain, plasmids and primers used in this study.

<b>Description</b>	<b>Content</b>	<b>Ref.</b>
<b>Strain</b>		
<i>E. coli</i> BI21 <sup>star</sup> (DE3)	B F- <i>ompT hsdSB</i> (rB <sup>m</sup> B <sup>b</sup> ) <i>gal dcm rne131</i> (DE3)	Invitrogen
<b>Plasmids</b>		
<b>pET20b</b>		Epoch Biolabs
<b>pET20b-di</b>	<i>di</i> gene	This study
<b>pET20b-g6pdh</b>	<i>g6pdh</i> gene	This study
<b>pET20b-g6pdh-di</b>	<i>g6pdh-di</i> chimeric gene	This study
<b>pET20b-di-g6pdh</b>	<i>di-g6pdh</i> chimeric gene	This study
<b>Primers</b>		
<b>di-VF</b>	5'-GGCGA AAGAC TTGGC GCACA CGTTT CTCGA GCACC ACCAC CACCA CCACT-3'	For the construction of pET20b-di
<b>di-VR</b>	5'-GGGCG GTGAT GTACA ATACT TCGTC ATATG TATAT CTCCT TCTTA AAGT-3'	
<b>di-IF</b>	5'-ACTTT AAGAA GGAGA TATAC ATATG ACGAA AGTAT TGTAC ATCAC CGCCC-3'	
<b>di-IR</b>	5'-AGTGG TGGTG GTGGT GGTGC TCGAG AAACG TGTGC GCCAA GTCTT TCGCC-3'	
<b>g6pdh-VF</b>	5'-GTGGC CGATC GAGCA TCCGC GTTCG CTCGA GCACC ACCAC CACCA CCACT-3'	For the construction of pET20b-g6pdh
<b>g6pdh-VR</b>	5'-AAATG ACGAT GATCG ATTTT GGGTT CATAT GTATA TCTCC TTCTT AAAGT-3'	
<b>g6pdh-IF</b>	5'-ACTTT AAGAA GGAGA TATAC ATATG AACCC GAAAT CGATC ATCGT CATT-3'	
<b>g6pdh-IR</b>	5'-AGTGG TGGTG GTGGT GGTGC TCGAC GAACG CGGAT GCTCG ATCGG CCAC-3'	
<b>g6pdh-di-VF</b>	5'-AGGAT CCGGA GGTGG AGGAT CCGGC ACGAA AGTAT TGTAC ATCAC CGCCC-3'	For the construction of pET20b-g6pdh-di
<b>g6pdh-di-VR</b>	5'-AAATG ACGAT GATCG ATTTT GGGTT CATAT GTATA TCTCC TTCTT AAAGT-3'	
<b>g6pdh-di-IF</b>	5'-ACTTT AAGAA GGAGA TATAC ATATG AACCC GAAAT CGATC ATCGT CATT-3'	
<b>g6pdh-di-Insrt-R</b>	5'-GCCGG ATCCT CCACC TCCGG ATCCT CCACC CGAAC GCGGA TGCTC GATCG-3'	
<b>g6pdh-di-IR</b>	5'-GGGCG GTGAT GTACA ATACT TTCGT GCCGG ATCCT CCACC TCCGG ATCCT-3'	
<b>di-g6pdh-VF</b>	5'-GTGGC CGATC GAGCA TCCGC GTTCG CACCA CCACC ACCAC CACTG AGATC-3'	For the construction of pET20b-di-g6pdh
<b>di-g6pdh-VR</b>	5'-ATCCT CCACC TCCGG ATCCT CCACC CTCGA GAAAC GTGTG CGCCA AGTCT-3'	
<b>di-g6pdh-Insert-F</b>	5'-GGTGG AGGAT CCGGA GGTGG AGGAT CCGGC AACCC GAAAT CGATC ATCG-3'	
<b>di-g6pdh-IF</b>	5'-AGACT TGGCG CACAC GTTTC TCGAG GGTGG AGGAT CCGGA GGTGG AGGAT-3'	
<b>di-g6pdh-IR</b>	5'-GATCT CAGTG GTGGT GGTGG TGGTG CGAAC GCGGA TGCTC GATCG GCCAC-3'	

### **3.2.5 Assay of particle size of enzymes.**

The particle sizes of individual proteins, the fusion proteins, and their mixtures in the 50 mM HEPES buffer (pH 7.5) were estimated at room temperature by the Malvern Zetasizer Nano ZS system (Worcestershire, UK). Prior to the assay, the enzyme solutions were diluted to different concentrations with 50 mM HEPES buffer (pH 7.5).

### **3.2.6. Other assays**

Protein mass concentration was measured by the Bio-Rad Bradford protein dye reagent method (Bio-Rad, Hercules, CA) with bovine serum albumin as a reference. The protein mass based on the Bradford method was calibrated by their absorbance (280 nm) in 6 M guanidine hydrochloride (You et al. 2012a). The purity of protein samples was examined by 12% SDS-PAGE and the formation of protein multimers was examined by 10% native gel, SDS-PAGE and native gel were stained by the Bio-Rad Bio-Safe Colloidal Coomassie Blue G-250.

### **3.2.7. Native gel**

The native gel included the separating gel and stacking gel. The 10% acrylamide/bisacrylamide separating gel contained 2.2 mL of H<sub>2</sub>O, 1.0 mL of 40% acrylamide/bisacrylamide, 0.75 mL of 2 M Tris-HCl (pH 8.8), 40  $\mu$ L of 10% ammonium persulfate, and 3  $\mu$ L of tetramethylethylenediamine (TEMED). The 5% acrylamide/bisacrylamide stacking gel contained 1.0 mL of H<sub>2</sub>O, 190  $\mu$ L of 40% acrylamide/bisacrylamide, 94  $\mu$ L of 2 M Tris-HCl (pH8.8), 15  $\mu$ L of 10% APS, and 1.5  $\mu$ L of TEMED. The protein samples were mixed at the following ratio

at room temperature in a 50 mM HEPES buffer (pH 7.5) containing 50 mM NaCl and 5 fold non-SDS loading buffer. After 5 min, the protein samples were loaded to a non-denaturing PAGE gel.

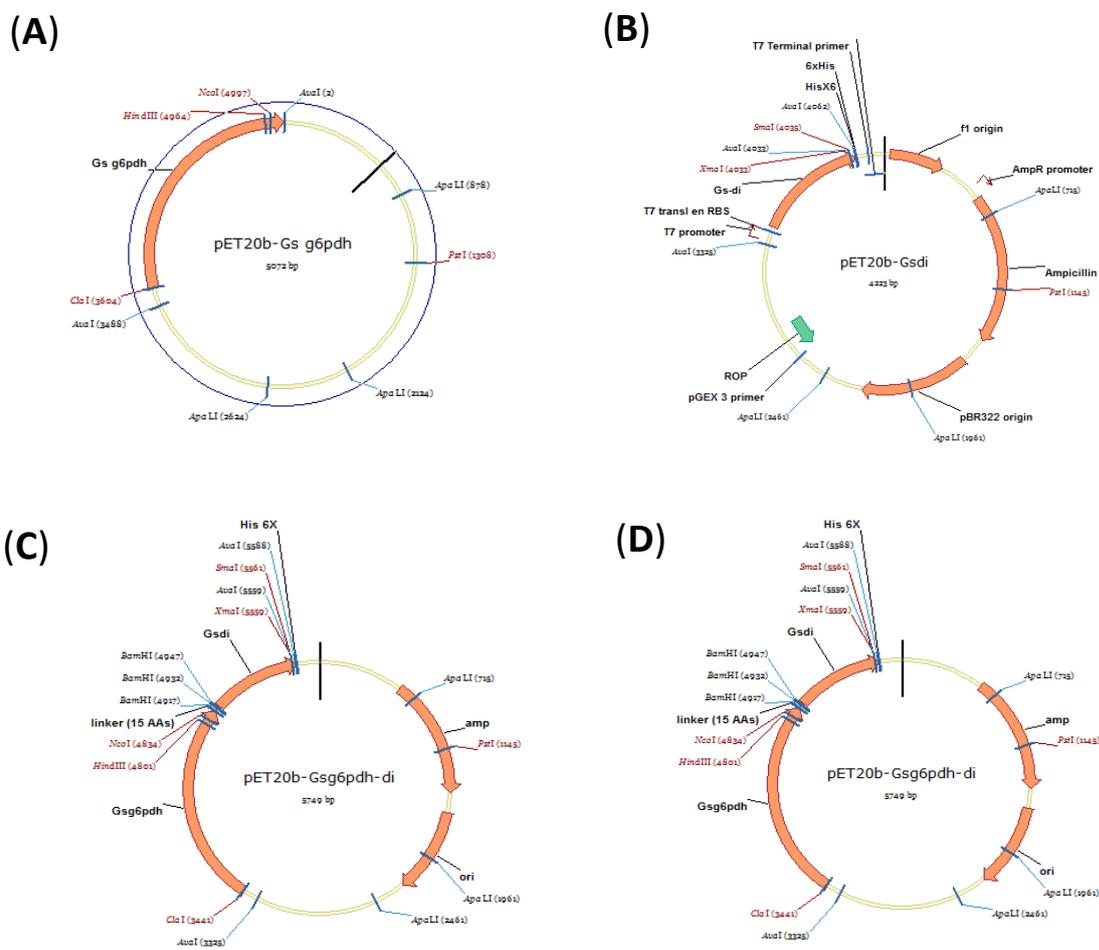
### **3.3 Results**

#### **3.3.1 Plasmid construction and protein purification**

The DNA fragment containing a 1485-bp *G. stearothersophilus g6pdh* gene with both 25-bp 5' and 25-bp 3' termini of the vector pET20b backbone sequence was amplified by PCR based on the genomic DNA of *G. stearothersophilus* 10. The pET20b vector backbone with both 25-bp 5' and 25-bp 3' termini of the *g6pdh* gene was amplified by PCR. Both PCR products were assembled to DNA multimers by prolonged overlap extension PCR (POE-PCR) and then transformed into *E. coli* DH5 $\alpha$  cells, yielding the chimeric plasmid pET20b-Gsg6pdh, which can express the recombinant G6PDH with a His tag at the C terminus under the control of T7 promoter (Fig. 3-1A). Via the same strategy, plasmid pET20b-Gsdi was constructed for expressing recombinant *G. stearothersophilus* diaphorase (Fig. 3-1B).

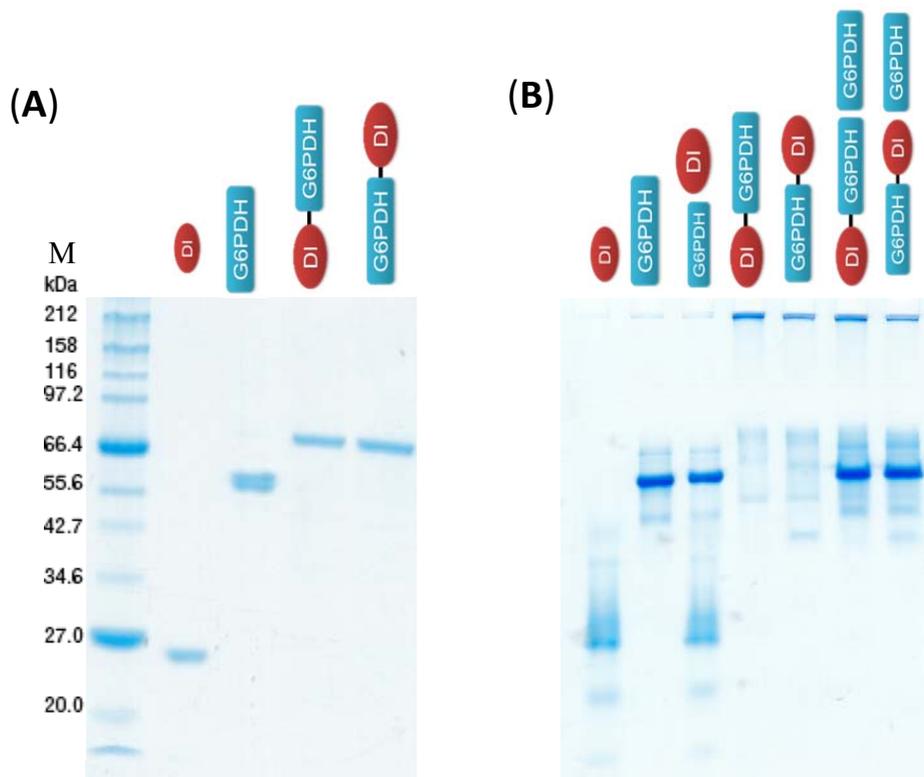
Plasmid pET20b-g6pdh-di was constructed for expressing the fusion protein G6PDH-DI, which had a 10-amino acid linker (GGGSG GGGSG) between G6PDH and DI and a C-terminal His tag. First, the *g6pdh* gene was amplified with PCR with a primer pair of g6pdh-di-IF and g6pdh-di-Insrt-R based on pET20b-Gsg6pdh. Second, the g6pdh gene fused with a 30-bp DNA fragment encoding the linker at the 3' terminus was generated with a primer pair of g6pdh-di-IF and g6pdh-di-Insrt-R by PCR based on the first round PCR product. Third, the backbone plasmid containing the DI gene was amplified with a primer pair of g6pdh-di-VF and g6pdh-di-VR based

on pET20b-Gsdi. Fourth, the second and third round PCR products were assembled by POE-PCR and transformed into *E. coli* DH5 $\alpha$ , yielding plasmid pET20b-g6pdh-di (Fig. 3-1C). Via the same strategy, plasmid pET20b-di-g6pdh was constructed for expressing the fusion protein DI-G6PDH, which had a 10-amino acid linker (GGGSG GGGSG) between DI and G6PDH and a C-terminal His-tag (Fig. 3-1D).



**Fig. 3-1.** Construction of the protein expression plasmids (A) pET20b-Gsg6pdh; (B) pET20b-Gsdi; (C) pET20b-Gsg6pdh-di; (D) pET20b-Gsdi-g6pdh

These recombinant plasmids were transformed into *E. coli* BL21 (DE3). After IPTG induction, the soluble recombinant proteins in the *E. coli* cell lysate were purified by using Ni-NTA resins to homogeneity and examined by SDS-PAGE analysis (Fig. 3-2A). The molecular mass of G6PDH, DI, G6PDH-DI, and DI-G6PDH were 25,000, 58,000, 80,000 and 80,000 kDa, in consistence with their deduced molecular mass based on amino acid sequences (57,972, 24,812, 82 101, and 82 102), respectively.



**Fig. 3-2.** 12% SDS-PAGE analysis of the purified free enzymes and fusion enzymes (A) and 10% Native gel analysis of free enzymes, fusion enzymes and the mixture of G6PDH and fusion enzymes (B).

The specific G6PDH activities of G6PDH, DI-G6PDH and G6PDH-DI were 710, 76, and 108 units/ $\mu$ mol of enzyme, respectively (Table 3-2), suggesting that the DI addition in fusion proteins

severely impaired activity of G6PDH. On the contrary, the specific DI activities of DI, DI-G6PDH and G6PDH-DI was 5,066, 5,120, and 5,933 units/ $\mu\text{mol}$  of enzyme, respectively (Table 3-2), suggesting that DI activity was not significantly affected with the presence of G6PDH.

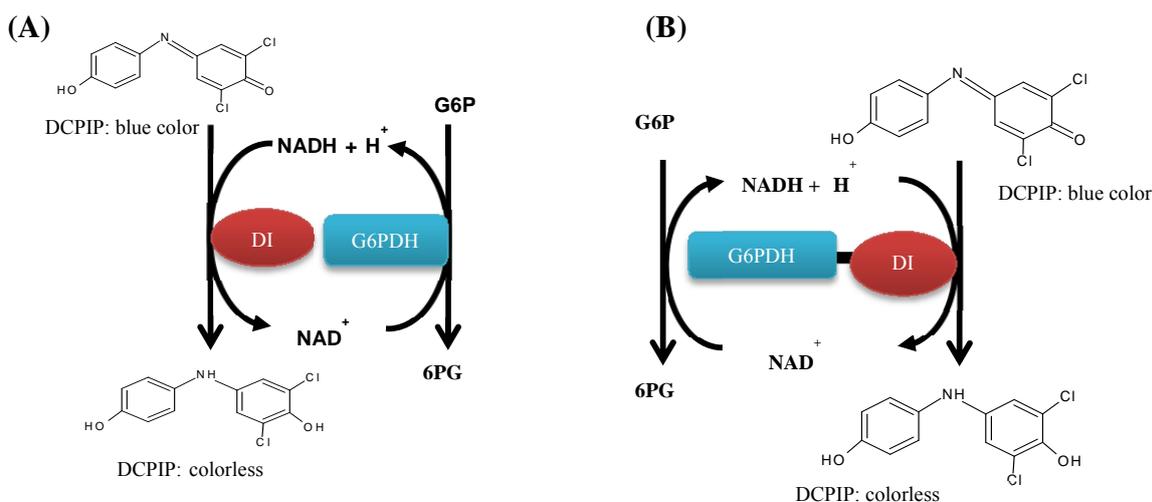
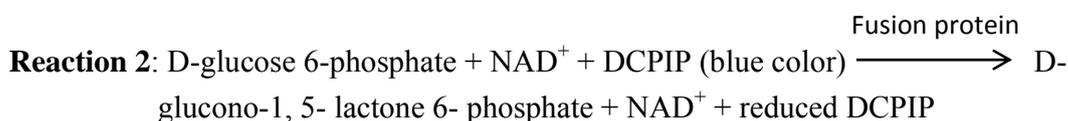
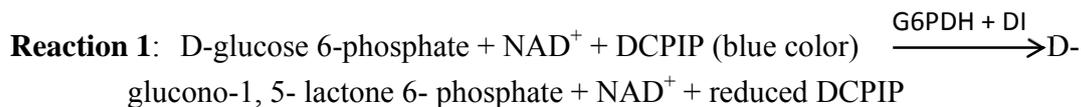
**Table 3-2.** Specific activities of G6PDH, DI, as well as two fusion proteins G6PDH-DI and DI-G6PDH.

Enzyme	Substrate	Specific activity* (unit / $\mu\text{mol}$ of enzyme)
G6PDH	G6P + NAD <sup>+</sup>	710 $\pm$ 15
DI-G6PDH	G6P + NAD <sup>+</sup>	76 $\pm$ 3
G6PDH-DI	G6P + NAD <sup>+</sup>	108 $\pm$ 5
DI	NADH + DCPIP	5066 $\pm$ 16
DI-G6PDH	NADH + DCPIP	5120 $\pm$ 20
G6PDH-DI	NADH + DCPIP	5933 $\pm$ 6

\*, specific activities were measured at 23°C and the concentration was 0.03 $\mu\text{M}$  and 0.06  $\mu\text{M}$  for DI and G6PDH, respectively.

### 3.3.2 Substrate channeling

To study the proximity or substrate channeling effect, two sets of reactions were run in parallel (Fig. 3-3). In the first set, the multi-enzyme system formed by a mixture of G6PDH and DI was used to catalyze the reaction 1 (Fig. 3-3A). In the second set of reactions the new fusion enzymes were employed to catalyze the reaction 2 (Fig. 3-3B).



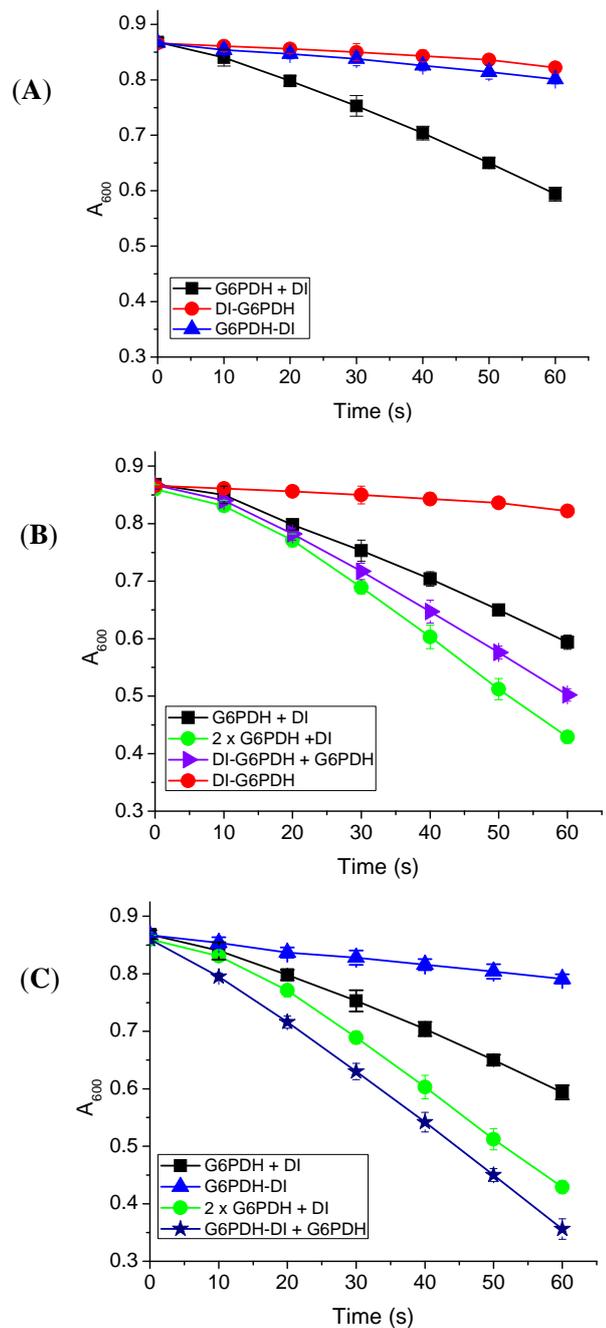
**Fig. 3-3.** Schematic representation of the catalytic reactions in the mixture of free G6PDH and DI system (A) and the possible substrate channeling in the mixture of free G6PDH and fusion enzyme G6PDH/DI system (B).

Catalytic performances of a mixture of G6PDH and DI, G6PDH-DI and DI-G6PDH at the same enzyme concentration of 0.06  $\mu\text{M}$  were compared in a 50 mM HEPES buffer (pH 7.5) containing 2.5 mM G6P and 0.1 mM DCPIP at room temperature. Fig. 3-4A clearly shows that much slower reaction rates were observed for two fusion proteins than the G6PDH/DI mixture, suggesting that drastic decreases in G6PDH activity in the fusion protein severely impaired the fusion protein performance and there was no substrate channeling between two catalytic enzymes in the two fusion proteins.

When G6PDH concentration was increased by two fold (Fig. 3-4B), the overall reaction rate of the mixture of 0.12  $\mu\text{M}$  G6DPH and 0.06  $\mu\text{M}$  DI approximately doubled that of the mixture of 0.06  $\mu\text{M}$  G6DPH and 0.06  $\mu\text{M}$  DI, indicating that reaction mediated by G6PDH was the rate-limiting step in this two-step cascade reaction system. The mixture of G6DPH and DI-G6PDH exhibited a drastic increase in reaction rate compared to DI-G6PDH by nine fold (Fig. 3-4B); approximately 30% higher than the reaction rate sum of the G6PDH/DI mixture and DI-G6PDH, implying the formation of heteromultimeric G6PDH and DI-G6PDH could restore greatly G6PDH activity. The mixture of 0.06  $\mu\text{M}$  G6DPH and 0.06  $\mu\text{M}$  DI-G6PDH exhibited a lower reaction rate compared to the mixture of 0.12  $\mu\text{M}$  G6PDH and 0.06  $\mu\text{M}$  DI (Fig. 3-4B), suggesting that there was no substrate channeling between G6DPH and DI-G6PDH.

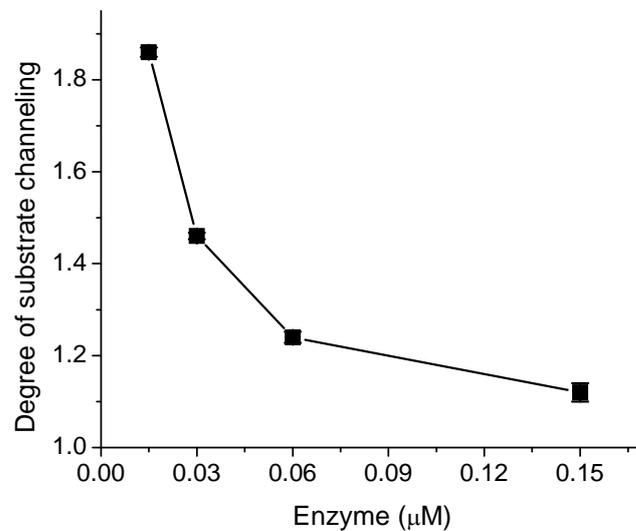
Surprisingly, the mixture of 0.06  $\mu\text{M}$  G6PDH and 0.06  $\mu\text{M}$  the fusion protein G6PDH-DI exhibited higher reaction rates than the mixture of 0.12  $\mu\text{M}$  G6PDH and 0.06  $\mu\text{M}$  DI, suggesting that there was substrate channeling among G6PDH and G6PDH-DI (Fig. 3-4C). The mixture of G6PDH and G6PDH-DI also enabled to restore G6PDH activity in the fusion protein G6PDH-DI. Results of Fig. 3-4B and c clearly suggested that whether substrate channeling was observed in fusion proteins containing cascade enzymes or not depended on their orientation.

Furthermore, we investigated the effects of enzyme concentration on the degree of substrate channeling of the mixture of G6PDH/G6PDH-DI relative to the mixture of  $2 \times$  G6PDH/DI at the



**Fig. 3-4.** Enzyme activity assays of free enzymes, fusion enzymes and the mixture of G6PDH and fusion enzymes G6PDH-DI and DI-G6PDH. (A) Enzyme activity assays among free enzyme mixture G6PDH/DI, fusion protein G6PDH-DI and DI-G6PDH, (B) enzyme activity assays among free enzyme mixtures G6PDH/DI, 2xG6PDH/DI, fusion enzymes DI-G6PDH, the mixture of G6PDH/DI-G6PDH, (C) enzyme activity assays among free enzyme mixture G6PDH/DI, 2xG6PDH/DI, fusion enzymes G6PDH-DI, the mixture of G6PDH/G6PDH-DI.

same enzyme loading (Fig. 3-5). The degree of substrate channeling decreased from 1.86 to 1.12 when the enzyme concentration increased from 0.015 to 0.15  $\mu\text{M}$ . This trend was in good agreement with the previous report pertaining to the activity of complexed enzyme relative to that of non-complexed enzymes (You et al. 2012c). At low enzyme concentrations, the channeling substrate took longer time to transfer G6PDH to DI in the free enzyme mixture than in the G6PDH/G6PDH-DI system because the distance between cascade enzymes the latters were fixed, independent of enzyme concentration. These data suggested that the degree of substrate channeling depends on enzyme concentration and implied that substrate channeling effects could be more important for low concentration cellular proteins in cytoplasm.



**Fig. 3-5.** The effects of enzyme concentration on the degree of substrate channeling of the mixture of G6PDH/G6PDH-DI relative to the mixture of  $2 \times$  G6PDH/DI at the same enzyme loading.

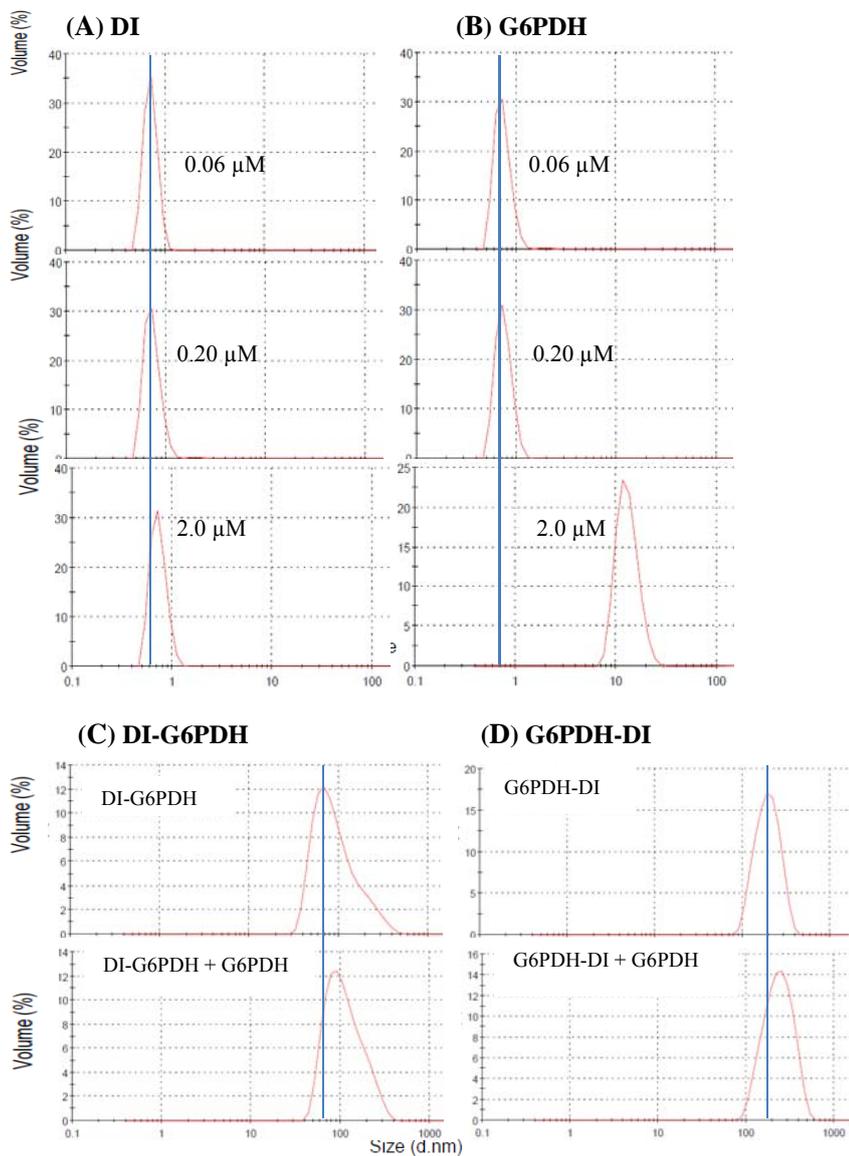
### 3.3.3 Particle size in aqueous solution

Most intracellular enzymes are believed to form active multimers in cytoplasm, where macromolecules, such as other enzymes, nucleic acids (i.e., RNA and DNA), structural proteins, and membranes occupy 30-40% of cellular interior. Equilibrium constants for protein associations can be increased by one to three orders of magnitude, depending on sizes and shapes of macromolecular reactants and products, and of background macromolecules. Therefore, it was interesting to determine whether G6PDH, DI, and their fusion proteins form multimers at low protein concentrations (e.g., 0.06  $\mu\text{M}$ ) or not.

The particle size of the individual enzymes and fusion enzymes were examined by dynamic light scattering through a Malvern Zetasizer Nano ZS system. The particle size of DI was nearly constant. The diameters of DI were 0.71 nm, 0.71 nm and 0.71 nm at the concentration of 0.06  $\mu\text{M}$ , 0.2  $\mu\text{M}$ , 2  $\mu\text{M}$ , respectively (Fig. 3-6A). It suggested that DI could be a monomer regardless its concentration. In contrast, the particle size of G6PDH depended on its concentration. The diameters of G6PDH were approximately 0.73 nm, 0.74 nm and 13.16 nm at the concentration of 0.06  $\mu\text{M}$ , 0.2  $\mu\text{M}$ , 2  $\mu\text{M}$ , respectively (3-6B). Such data suggested that G6PDH could form a very large size multimeric network when protein concentration was high enough. To our limited knowledge, this phenomenon was first reported in the aqueous solution. The formation of such macromolecular structure in G6DPH at a relatively high protein concentration was so loose that they cannot be detected by conventional assays, such as native gel electrophoresis (Figure 3-2B). In partially support to it, the *Neurospora crassa* G6PDH was found to form more active tetramers from dimers when NADPH concentration was increased to 2 mM (Scott 1971). Also, in vivo study suggested that the product of an gene *opcA*, adjacent to *zwf* encoding G6DPH was

involved in its oligomerization and activation of G6PDH (Sundaram et al. 1998). Similarly, fibrinogen was found to form a homodimer when bovine serum albumin (BSA) exceeded 40 g/L and to form multimerization with an estimated activity coefficient of 10 when BSA concentration was 80 g/L. The activity of fibrinogen in blood plasma was an order of magnitude larger than that exhibited in dilute solutions. Previous study pertaining to the addition of macromolecular crowding agents -- polyethylene glycol (PEG) 6000 and PEG 35000 also lead to a sharp decrease in dissociation constants of phosphoenol pyruvate/carbohydrate phosphotransferase system (Francke et al. 2003; Rohwer et al. 1998).

Furthermore, the diameters of fusion proteins DI-G6PDH, G6PDH-DI were 68.06 nm, 181.1 nm at enzyme concentration of 0.06  $\mu$ M, respectively. However, when equal-molar G6PDH was mixed with fusion enzyme DI-G6PDH or G6PDH-DI, the diameter of the particle size increased to 91.28 nm and 223.8 nm at the enzyme concentration of 0.06  $\mu$ M, respectively (Fig. 3-6C and D). These results suggested that DI-G6PDH or G6PDH-DI and G6PDH formed heterogeneous multimers. Such loosely-associated heteromeric enzyme complexes at low protein concentrations in the aqueous solution were not detected by conventional assays, such as native gel or size-exclusion chromatography.



**Fig. 3-6.** Particle sizes of the individual enzyme components and fusion enzymes determined by the Malvern Zetasizer Nano ZS system. (A) DI at a concentration of 0.06  $\mu\text{M}$ , 0.20  $\mu\text{M}$ , and 2.0  $\mu\text{M}$ ; (B) G6PDH at a concentration of 0.06  $\mu\text{M}$ , 0.20  $\mu\text{M}$ , and 2.0  $\mu\text{M}$ ; (C) Fusion protein DI-G6PDH and the mixture of DI-G6PDH and G6PDH at the concentration of 0.06 $\mu\text{M}$ ; (D) Fusion protein G6PDH-DI and the mixture of G6PDH-DI and G6PDH at the concentration of 0.06 $\mu\text{M}$ .

### 3.4 Discussion

Fusion protein, one of the most popular approach to generate artificial bifunctional and multifunctional enzymes, is the creation of chimeric proteins in which two or more distant genetic elements are combined by a short linker sequence to yield a single polypeptide that exhibits more than one activity. The expression of a fusion protein is an attractive alternative for pathway engineering in bacteria and plants. Fusion proteins implement the coordinated expression of component enzymes. Additionally, the close proximity of active sites within fusion proteins has been suggested in some cases to increase reaction rates by allowing the transfer of metabolites between sites before re-equilibration with the surrounding environment can occur (Bulow and Mosbach 1991). For example, Mosbach and co-workers created an in-frame fusion protein between LacZ and the dimeric galactose dehydrogenase from *Pseudomonas fluorescens* for the sequential hydrolysis of lactose followed by the oxidation of the galactose, forming the corresponding lactone (Ljungcrantz et al. 1989). The resulted fusion enzyme displayed kinetic advantages over the identical native system in conversion of lactose to galactonolactone which suggests that the proximity conferred by the fusion protein enabled substrate channeling. However, not all the fusion proteins can cause synergistic reaction. For example, Snell and co-workers explored enzyme fusions for enhancing bacterial and plant-based biosynthesis of polyhydroxybutyrate (PHB) granules by expressing a fusion between the *Ralstonia eutropha* thiolase (PhaA) and reductase (PhaB) enzymes in *E. coli* (Kourtz et al. 2005). The resulting fusion protein exhibited thiolase and reductase enzyme activity that were three-fold and nine-fold less than those of the individual expressed thiolase and reductase enzymes, respectively.

In this study, we have developed the fusion proteins which consist of G6PDH and DI from *G. stearothermophilus* with a ten-amino acid linker in two orientations. The fusion protein retained good DI activity and lower G6PDH activity. The activity of DI in the fusion protein was not affected with the presence of G6PDH; however, G6PDH activity in the fusion protein was severely impaired by the DI addition in fusion proteins. According to the above results, we demonstrated that there are some key challenges associated with enzyme fusions. First, the fact that many metabolic enzymes are multimeric presents an assembly conundrum. That is, while a fusion of two normally monomeric enzymes is straightforward, the fusion of two normally multimeric enzymes requires subunit assembly for activity. These higher ordered structures can interfere with the non-covalent interactions that are required for the activity of individual enzyme subunits and can lead to a partial protein network that may or may not exhibit enzymatic activity (Conrado et al. 2008). Second, folding of large multidomain proteins is typically inefficient in bacteria (Chang et al. 2005; Netzer and Hartl 1997), thus numerous enzyme fusions may misfold or aggregate when expressed in bacteria.

Surprisingly, when the same amount of free G6PDH enzyme was added to the fusion protein, an obviously increased enzyme activity was observed, especially higher enzyme activity was detected in the mixture of G6PDH and G6PDH-DI compared to the free enzyme mixture of DI and double amount of G6PDH. According to the results observed, we hypothesized that after adding free G6PDH to the fusion protein, the spatial organization changed dramatically and the active multimeric enzyme complex formed, thus substrate channeling occurred. This hypothesis was further proved by dynamic light scattering.

## **4. Conclusions and perspectives**

### **4.1 Conclusions**

A putative RpiB was cloned from the thermophilic bacterium *thermotoga maritima*, heterologously expressed in *Escherichia coli*, purified by one step heat treatment, and biochemically characterized. High-purity *T. maritima* RpiB was simply purified by optimized heat treatment and it exhibited high TTN values under its suboptimal conditions.

Two fusion proteins which consist of G6PDH and DI from *G. stearothermophilus* with a ten amino acids linker in two orientations were developed. The fusion proteins were purified with His-tag. The DI activity in the fusion protein was not affected with the presence of G6PDH; however, G6PDH activity in the fusion proteins was severely impaired by the DI presence in fusion proteins. However, we observed that the mixture of G6PDH and a fusion protein G6PDH-DI not only restored G6PDH activity through the formation of heteromultimers but also facilitated substrate channeling between DI and G6PDH, especially at low enzyme concentrations.

### **4.2 Perspectives**

The production of biofuels is still too costly by SyPaB. It is mainly attributed to costly unstable enzymes. Enzyme costs can be decreased by use of (hyper)-thermostable enzymes, enzyme

immobilization, simple enzyme purification, large-scale production of recombinant proteins. A number of (hyper- ) thermostable enzymes have been cloned and expressed in *E. coli*, e.g. *C. thermocellum* PGM (Wang and Zhang 2010), *T. maritima* 6PGDH (Wang and Zhang 2009) and FBP (Myung et al. 2010). Both rational and random approaches have been applied to construct highly active and highly stable enzymes. One step heat treatment purification of RpiB decreased enzyme purification costs significantly, which may offer new strategies for the purification of other thermostable enzymes.

We developed two fusion proteins which consist of glucose-6-phosphate dehydrogenase (G6PDH) and diaphorase (DI) from *G. stearothermophilus* and we found their enzyme activities decreased dramatically compared to the free enzyme reaction system. After mixing free G6PDH enzyme with fusion proteins, the activities of fusion proteins can be restored partially or completely. We will further demonstrate the spatial structure change of the G6PDH through some biochemical and biological methods.

## References

- Adams MW, Stiefel EI. 1998. Biological hydrogen production: not so elementary. *Science* 282(5395):1842-3.
- Ansell RJ, Small DA, Lowe CR. 1999. Synthesis and properties of new coenzyme mimics based on the artificial coenzyme CL4. *J Mol Recognit* 12(1):45-56.
- Atkinson MR, Deutscher MP, Kornberg A, Russell AF, Moffatt JG. 1969. Enzymatic synthesis of deoxyribonucleic acid. XXXIV. Termination of chain growth by a 2',3'-dideoxyribonucleotide. *Biochemistry* 8(12):4897-904.
- Atsumi S, Hanai T, Liao JC. 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451(7174):86-9.
- Banki MR, Feng L, Wood DW. 2005. Simple bioseparations using self-cleaving elastin-like polypeptide tags. *Nat Methods* 2(9):659-61.
- Barnard GC, Henderson GE, Srinivasan S, Gerngross TU. 2004. High level recombinant protein expression in *Ralstonia eutropha* using T7 RNA polymerase based amplification. *Protein Expr Purif* 38(2):264-71.
- Bauer MW, Bylina EJ, Swanson RV, Kelly RM. 1996. Comparison of a beta-glucosidase and a beta-mannosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Purification, characterization, gene cloning, and sequence analysis. *J Biol Chem* 271(39):23749-55.
- Bok JD, Yernool DA, Eveleigh DE. 1998. Purification, characterization, and molecular analysis of thermostable cellulases CelA and CelB from *Thermotoga neapolitana*. *Appl Environ Microbiol* 64(12):4774-81.
- Borges A, Hawkins CF, Packman LC, Perham RN. 1990. Cloning and sequence analysis of the genes encoding the dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase components of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Eur J Biochem* 194(1):95-102.
- Bulow L, Ljungcrantz P, Mosbach K. 1985. Preparation of a soluble bifunctional enzyme by gene fusion. *Nat. Biotechnol.* 3(9):821-823.
- Bulow L, Mosbach K. 1991. Multienzyme systems obtained by gene fusion. *Trends Biotechnol* 9(7):226-31.

- Chakraborty S, Sakka M, Kimura T, Sakka K. 2008a. Cloning and expression of a *Clostridium kluyveri* gene responsible for diaphorase activity. *Biosci Biotechnol Biochem* 72(3):735-41.
- Chakraborty S, Sakka M, Kimura T, Sakka K. 2008b. Two proteins with diaphorase activity from *Clostridium thermocellum* and *Moorella thermoacetica*. *Biosci Biotechnol Biochem* 72(3):877-9.
- Chang HC, Kaiser CM, Hartl FU, Barral JM. 2005. De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *J Mol Biol* 353(2):397-409.
- Conrado RJ, Varner JD, DeLisa MP. 2008. Engineering the spatial organization of metabolic enzymes: Mimicking nature's synergy. *Curr. Opin. Biotechnol.* 19(5):492-499.
- Cosgrove MS, Naylor C, Paludan S, Adams MJ, Levy HR. 1998. On the mechanism of the reaction catalyzed by glucose 6-phosphate dehydrogenase. *Biochemistry* 37(9):2759-67.
- Demain AL, Newcomb M, Wu JHD. 2005. Cellulase, clostridia, and ethanol. *Microbiol Mol Biol Rev* 69(1):124 -154.
- Demirbas MF. 2009. Biorefineries for biofuel upgrading: A critical review. *Applied Energy* 86:S151-S161.
- Dewar MJ, Storch DM. 1985. Alternative view of enzyme reactions. *Proc Natl Acad Sci U S A* 82(8):2225-9.
- Dische Z, Borenfreund E. 1951. A new spectrophotometric method for the detection and determination of keto sugars and trioses. *J Biol Chem* 192(2):583-7.
- Domagk GF, Doering KM, Chilla R. 1973. Purification and properties of ribose-phosphate isomerase from *Candida utilis*. *Eur J Biochem* 38(2):259-64.
- Easterby JS. 1981. A generalized theory of the transition time for sequential enzyme reactions. *Biochem J* 199(1):155-61.
- Eric Hodgman C, Jewett MC. 2012. Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng.*(0):DOI: 10.1016/j.ymben.2011.09.002.
- Essenberg MK, Cooper RA. 1975. Two ribose-5-phosphate isomerases from *Escherichia coli* K12: partial characterisation of the enzymes and consideration of their possible physiological roles. *Eur J Biochem* 55(2):323-32.
- Farmer TB, Caprioli RM. 1991. Assessing the multimeric states of proteins: studies using laser desorption mass spectrometry. *Biol Mass Spectrom* 20(12):796-800.

- Francke C, Postma PW, Westerhoff HV, Blom JG, Peletier MA. 2003. Why the phosphotransferase system of *Escherichia coli* escapes diffusion limitation. *Biophys J* 85(1):612-22.
- Geck MK, Kirsch JF. 1999. A novel, definitive test for substrate channeling illustrated with the aspartate aminotransferase/malate dehydrogenase system. *Biochemistry* 38(25):8032-7.
- Georis J, de Lemos Esteves F, Lamotte-Brasseur J, Bougniet V, Devreese B, Giannotta F, Granier B, Frere JM. 2000. An additional aromatic interaction improves the thermostability and thermophilicity of a mesophilic family 11 xylanase: structural basis and molecular study. *Protein Sci* 9(3):466-75.
- Giver L, Gershenson A, Freskgard PO, Arnold FH. 1998. Directed evolution of a thermostable esterase. *Proc Natl Acad Sci U S A* 95(22):12809-13.
- Hartley JL. 2006. Cloning technologies for protein expression and purification. *Curr Opin Biotechnol* 17(4):359-66.
- Hill J, Nelson E, Tilman D, Polasky S, Tiffany D. 2006. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc Natl Acad Sci U S A* 103(30):11206-11210.
- Hong J, Wang Y, Ye X, Zhang YH. 2008. Simple protein purification through affinity adsorption on regenerated amorphous cellulose followed by intein self-cleavage. *J Chromatogr A* 1194(2):150-4.
- Horecker BL, Smyrniotis PZ, Seegmiller JE. 1951. The enzymatic conversion of 6-phosphogluconate to ribulose-5-phosphate and ribose-5-phosphate. *J Biol Chem* 193(1):383-96.
- Hove-Jensen B, Maigaard M. 1993. *Escherichia coli* rpiA gene encoding ribose phosphate isomerase A. *J Bacteriol* 175(17):5628-35.
- Huang WD, Zhang Y-HP. 2011. Analysis of biofuels production from sugar based on three criteria: Thermodynamics, bioenergetics, and product separation *Energy Environ. Sci.* 4:784-792.
- Huck JH, Verhoeven NM, Struys EA, Salomons GS, Jakobs C, van der Knaap MS. 2004. Ribose-5-phosphate isomerase deficiency: new inborn error in the pentose phosphate pathway associated with a slowly progressive leukoencephalopathy. *Am J Hum Genet* 74(4):745-51.

- Ishikawa K, Matsui I, Payan F, Cambillau C, Ishida H, Kawarabayasi Y, Kikuchi H, Roussel A. 2002. A hyperthermostable D-ribose-5-phosphate isomerase from *Pyrococcus horikoshii* characterization and three-dimensional structure. *Structure* 10(6):877-86.
- Iturrate L, Sanchez-Moreno I, Doyaguez EG, Garcia-Junceda E. 2009. Substrate channelling in an engineered bifunctional aldolase/kinase enzyme confers catalytic advantage for C-C bond formation. *Chem. Commun.* 2009(13):1721-1723.
- Kalscheuer R, Stolting T, Steinbuchel A. 2006. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* 152(Pt 9):2529-36.
- Kholodenko BN, Demin OV, Westerhoff HV. 1993. 'Channelled' pathways can be more sensitive to specific regulatory signals. *FEBS Lett* 320(1):75-8.
- Klein-Marcuschamer D, Ajikumar PK, Stephanopoulos G. 2007. Engineering microbial cell factories for biosynthesis of isoprenoid molecules: beyond lycopene. *Trends Biotechnol.* 25(9):417-424.
- Kourtz L, Dillon K, Daughtry S, Madison LL, Peoples O, Snell KD. 2005. A novel thiolase-reductase gene fusion promotes the production of polyhydroxybutyrate in *Arabidopsis*. *Plant Biotechnol J* 3(4):435-47.
- Levy HR. 1979. Glucose-6-phosphate dehydrogenases. *Adv Enzymol Relat Areas Mol Biol* 48:97-192.
- Liu T, Vora H, Khosla C. 2010. Quantitative analysis and engineering of fatty acid biosynthesis in *E. coli*. *Metab. Eng.* 12(4):378-386.
- Liu W, Hong J, Bevan DR, Zhang YH. 2009. Fast identification of thermostable beta-glucosidase mutants on cellobiose by a novel combinatorial selection/screening approach. *Biotechnol Bioeng* 103(6):1087-94.
- Ljungcrantz P, Carlsson H, Mansson MO, Buckel P, Mosbach K, Buelow L. 1989. Construction of an artificial bifunctional enzyme, .beta.-galactosidase/galactose dehydrogenase, exhibiting efficient galactose channeling. *Biochemistry* 28(22):8786-8792.
- Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. 2002. Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol Mol Biol Rev* 66(4):739-739.
- Mariano S, Roos AK, Mowbray SL, Salmon L. 2009. Competitive inhibitors of type B ribose 5-phosphate isomerases: design, synthesis and kinetic evaluation of new D-allose and D-allulose 6-phosphate derivatives. *Carbohydr Res* 344(7):869-80.

- Meyer FM, Gerwig J, Hammer E, Herzberg C, Commichau FM, Völker U, Stülke J. 2011. Physical interactions between tricarboxylic acid cycle enzymes in *Bacillus subtilis*: Evidence for a metabolon. *Metab. Eng.* 13(1):18-27.
- Middaugh CR, MacElroy RD. 1976. The effect of temperature on ribose-5-phosphate isomerase from a mesophile, *Thiobacillus thioparus*, and a thermophile, *Bacillus caldolyticus*. *J Biochem* 79(6):1331-44.
- Myung S, Wang YR, Zhang YHP. 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 X-Z, 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. Prog.* 27:969–975.
- Netzer WJ, Hartl FU. 1997. Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388(6640):343-9.
- Okuno H, Nagata K, Nakajima H. 1985. Purification and properties of glucose-6-phosphate dehydrogenase from *Bacillus stearothermophilus*. *J Appl Biochem* 7(3):192-201.
- Olive C, Geroch ME, Levy HR. 1971. Glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides*. Kinetic studies. *J Biol Chem* 246(7):2047-57.
- Orita I, Sakamoto N, Kato N, Yurimoto H, Sakai Y. 2007. Bifunctional enzyme fusion of 3-hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase. *Appl. Microbiol. Biotechnol.* 76(2):439-445.
- Ovadi J, Huang Y, Spivey HO. 1994. Binding of malate dehydrogenase and NADH channelling to complex I. *J Mol Recognit* 7(4):265-72.
- Pettersson H, Pettersson G. 2001. Kinetics of the coupled reaction catalysed by a fusion protein of [beta]-galactosidase and galactose dehydrogenase. *Biochim. Biophys. Acta: Protein Struct. Mol. Enzymol.* 1549(2):155-160.
- Podar M, Reysenbach AL. 2006. New opportunities revealed by biotechnological explorations of extremophiles. *Curr Opin Biotechnol* 17(3):250-255.
- Rogers TA, Bommarius AS. 2010. Utilizing Simple Biochemical Measurements to Predict Lifetime Output of Biocatalysts in Continuous Isothermal Processes. *Chem Eng Sci* 65(6):2118-2124.

- Rohwer JM, Postma PW, Kholodenko BN, Westerhoff HV. 1998. Implications of macromolecular crowding for signal transduction and metabolite channeling. *Proc Natl Acad Sci U S A* 95(18):10547-52.
- Roos AK, Burgos E, Ericsson DJ, Salmon L, Mowbray SL. 2005. Competitive inhibitors of *Mycobacterium tuberculosis* ribose-5-phosphate isomerase B reveal new information about the reaction mechanism. *J Biol Chem* 280(8):6416-22.
- Rudolph J, Stubbe J. 1995. Investigation of the mechanism of phosphoribosylamine transfer from glutamine phosphoribosylpyrophosphate amidotransferase to glycinamide ribonucleotide synthetase. *Biochemistry* 34(7):2241-50.
- Ryan JD, Fish RH, Clark DS. 2008. Engineering cytochrome P450 enzymes for improved activity towards biomimetic 1,4-NADH cofactors. *Chembiochem* 9(16):2579-82.
- Salge JR, Dreyer BJ, Dauenhauer PJ, Schmidt LD. 2006. Renewable hydrogen from nonvolatile fuels by reactive flash volatilization. *Science* 314(5800):801-4.
- Satyanarayana T, Raghukumar C, Shivaji S. 2005. Extremophilic microbes: Diversity and perspectives. *Current Science* 89(1):78-90.
- Scott WA. 1971. Physical Properties of Glucose 6-Phosphate Dehydrogenase from *Neurospora crassa*. *J. Biol. Chem.* 246(20):6353-6359.
- Shaw AJ, Podkaminer KK, Desai SG, Bardsley JS, Rogers SR, Thorne PG, Hogsett DA, Lynd LR. 2008. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proc Natl Acad Sci U S A* 105(37):13769-74.
- Sorensen KI, Hove-Jensen B. 1996. Ribose catabolism of *Escherichia coli*: characterization of the rpiB gene encoding ribose phosphate isomerase B and of the rpiR gene, which is involved in regulation of rpiB expression. *J Bacteriol* 178(4):1003-11.
- Stern AL, Burgos E, Salmon L, Cazzulo JJ. 2007. Ribose 5-phosphate isomerase type B from *Trypanosoma cruzi*: kinetic properties and site-directed mutagenesis reveal information about the reaction mechanism. *Biochem J* 401(1):279-85.
- Sundaram S, Karakaya H, Scanlan DJ, Mann NH. 1998. Multiple oligomeric forms of glucoses-6-phosphate dehydrogenase in cyanobacteria and the role of OpeA in the assembly process. *Microbiology* 144(6):1549-1556.
- Swartz JR. 2011. Transforming biochemical engineering with cell-free biology. *AIChE J.* 58:5-13.

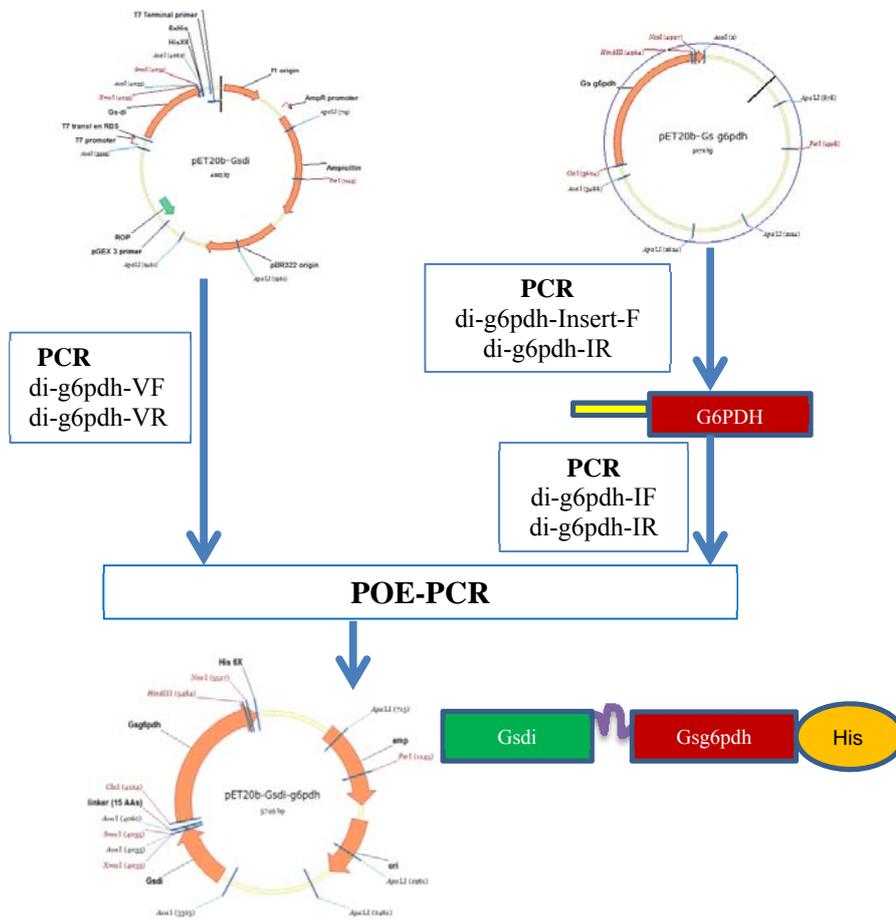
- Turner P, Mamo G, Karlsson EN. 2007. Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microbial Cell Factories* 6.
- Ushiroyama T, Fukushima T, Styre JD, Spivey HO. 1992. Substrate channeling of NADH in mitochondrial redox processes. *Curr Top Cell Regul* 33:291-307.
- Wang Y, Huang W, Sathitsuksanoh N, Zhu Z, Zhang Y-HP. 2011. Biohydrogenation from biomass sugar mediated by *in vitro* synthetic enzymatic pathways. *Chem. Biol.* 18:372-380.
- Wang Y, Zhang YH. 2010. A highly active phosphoglucomutase from *Clostridium thermocellum*: cloning, purification, characterization and enhanced thermostability. *J Appl Microbiol* 108(1):39-46.
- Wang Y, Zhang YHP. 2009. Overexpression and simple purification of the *Thermotoga maritima* 6-phosphogluconate dehydrogenase in *Escherichia coli* and its application for NADPH regeneration. *Microbial Cell Factories* 8:30.
- Westerhoff HV, Welch GR. 1992. Enzyme organization and the direction of metabolic flow: physicochemical considerations. *Curr Top Cell Regul* 33:361-90.
- Winter RT, van den Berg TE, Colpa DI, van Bloois E, Fraaije MW. 2012. Functionalization of Oxidases with Peroxidase Activity Creates Oxiperoxidases: A New Breed of Hybrid Enzyme Capable of Cascade Chemistry. *ChemBioChem* 13(2):252-258.
- Xiong Y, DeFraia C, Williams D, Zhang X, Mou Z. 2009. Deficiency in a cytosolic ribose-5-phosphate isomerase causes chloroplast dysfunction, late flowering and premature cell death in Arabidopsis. *Physiol Plant* 137(3):249-63.
- Xu Q, Schwarzenbacher R, McMullan D, von Delft F, Brinen LS, Canaves JM, Dai X, Deacon AM, Elsliger MA, Eshagi S and others. 2004. Crystal structure of a ribose-5-phosphate isomerase RpiB (TM1080) from *Thermotoga maritima* at 1.90 Å resolution. *Proteins* 56(1):171-175.
- Ye X, Wang Y, Hopkins RC, Adams MW, Evans BR, Mielenz JR, Zhang YH. 2009. Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. *ChemSusChem* 2(2):149-52.
- Yeoman CJ, Han Y, Dodd D, Schroeder CM, Mackie RI, Cann IK. 2010. Thermostable enzymes as biocatalysts in the biofuel industry. *Adv Appl Microbiol* 70:1-55.

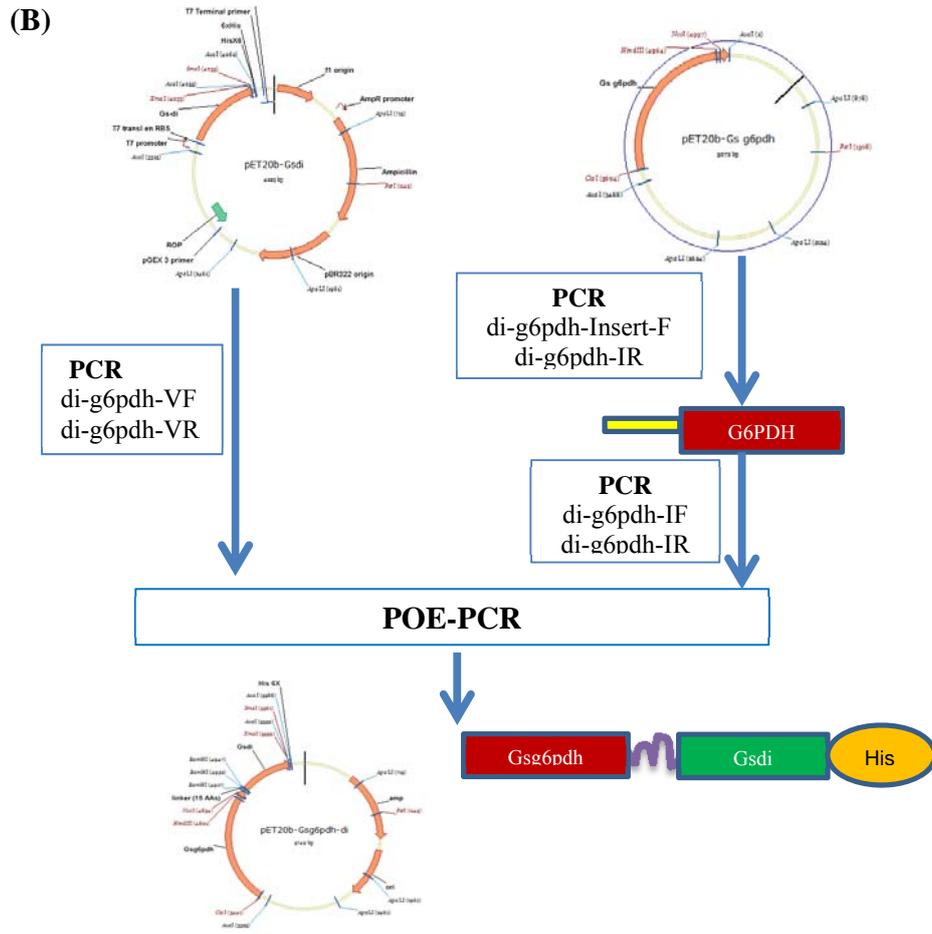
- Yoon RY, Yeom SJ, Kim HJ, Oh DK. 2009. Novel substrates of a ribose-5-phosphate isomerase from *Clostridium thermocellum*. J Biotechnol 139(1):26-32.
- You C, Zhang X-Z, Sathitsuksanoh N, Lynd LR, Zhang Y-HP. 2012a. Enhanced microbial cellulose utilization of recalcitrant cellulose by an *ex vivo* cellulosome-microbe complex. Appl. Environ. Microbiol. 78(5):1437-1444.
- You C, Zhang X-Z, Zhang Y-HP. 2012b. Simple Cloning: direct transformation of PCR product (DNA multimer) to *Escherichia coli* and *Bacillus subtilis*. Appl. Environ. Microbiol. 78:1593-1595.
- You C, Zhang XZ, Zhang YH. 2012c. Simple cloning via direct transformation of PCR product (DNA Multimer) to *Escherichia coli* and *Bacillus subtilis*. Appl Environ Microbiol 78(5):1593-5.
- Zeng YC, Zhang SZ. 1989. Purification and properties of beta-glucosidase from *Aspergillus phoenicis*. Wei Sheng Wu Xue Bao 29(3):195-9.
- Zhang RG, Andersson CE, Skarina T, Evdokimova E, Edwards AM, Joachimiak A, Savchenko A, Mowbray SL. 2003. The 2.2 Å resolution structure of RpiB/AlsB from *Escherichia coli* illustrates a new approach to the ribose-5-phosphate isomerase reaction. J Mol Biol 332(5):1083-94.
- Zhang Y-HP. 2011a. Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol. Adv. 29:715-725.
- Zhang Y-HP. 2011b. What is vital (and not vital) to advance economically-competitive biofuels production. Proc. Biochem. 46:2091-2110.
- Zhang YH, Lynd LR. 2004. Kinetics and relative importance of phosphorolytic and hydrolytic cleavage of cellodextrins and cellobiose in cell extracts of *Clostridium thermocellum*. Appl Environ Microbiol 70(3):1563-9.
- Zhang YH, Lynd LR. 2005. Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci U S A 102(20):7321-5.
- Zhang YH, Sun J, Zhong JJ. 2010. Biofuel production by in vitro synthetic enzymatic pathway biotransformation. Curr Opin Biotechnol 21(5):663-9.
- Zhang YHP. 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 YHP. 2010. Production of Biocommodities and Bioelectricity by Cell-Free Synthetic Enzymatic Pathway Biotransformations: Challenges and Opportunities. *Biotechnol Bioeng* 105(4):663-677.
- Zhang YHP. 2011c. Simpler Is Better: High-Yield and Potential Low-Cost Biofuels Production through Cell-Free Synthetic Pathway Biotransformation (SyPaB). *Acs Catalysis* 1(9):998-1009.
- Zverlov VV, Velikodvorskaya GA, Schwarz WH. 2002. A newly described cellulosomal cellobiohydrolase, CelO, from *Clostridium thermocellum*: investigation of the exo-mode of hydrolysis, and binding capacity to crystalline cellulose. *Microbiology* 148(Pt 1):247-55.

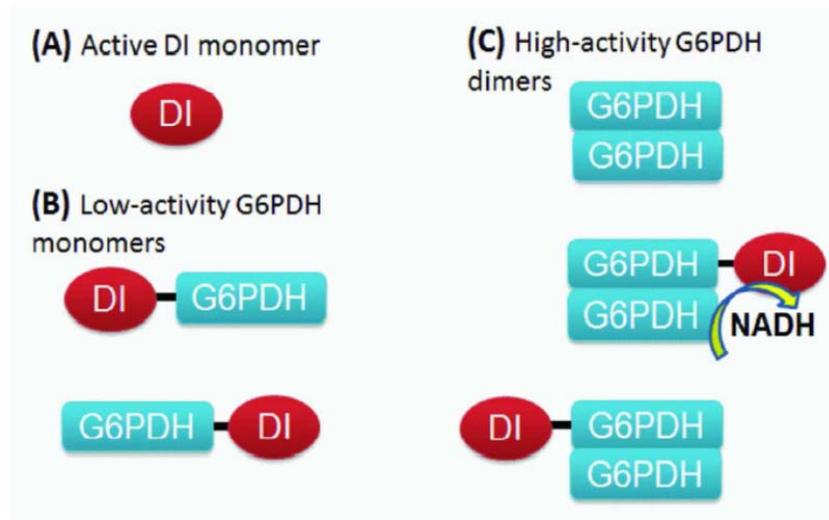
# Appendix

(A)





**Fig A1.** The scheme of construction fusion plasmids encoding two fusion proteins DI-G6PDH (A) and G6PDH-DI (B)



**Fig A2.** Active DI monomer (A), low-activity G6PDH monomers in fusion proteins (B), high-activity G6PDH multimers (C)-G6PDH only, G6PDH-DI/G6PDH featuring substrate channeling, and DI-G6PDH/G6PDH.