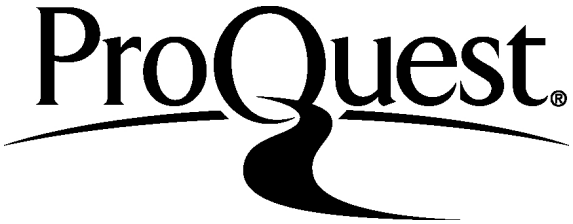
MIMICKING NATURE'S SYNERGY: ENGINEERING FN3-BASED PROTEIN SCAFFOLDS FOR MULTI-ENZYME ASSEMBLY

by

Long Chen

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master in Chemical Engineering

Spring 2015



MIMICKING NATURE'S SYNERGY: ENGINEERING FN3-BASED PROTEIN SCAFFOLDS FOR MULTI-ENZYME ASSEMBLY

by

Long Chen

Approved: _____
Wilfred Chen, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Abraham M. Lenhoff, Ph.D.
Chair of the Department of Chemical and Biomolecular Engineering

Approved: _____
Babatunde Ogunnaike, Ph.D.
Dean of the College of Engineering

Approved: _____
James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

At the end of this important and precious experience of my life, I would like to thank lots of people for their support and caring during my study and research in the University of Delaware.

First of all, I would like to express my deepest gratitude to my thesis advisor Prof. Wilfred Chen for his support and guidance on my research. His passion for research, his high scientific standards and his diligence set a good example for me. I learned a lot from him about doing research, critical thinking and presentations, *et al*. Also, I want to express my thankfulness to Prof. Beris, Prof. Colby, Prof. Dhurjati, Prof. Furst, Prof. Lobo, Prof. Ogunnaike and Prof. Papoutsakis for their enjoyable and unforgettable classes.

In the last couple of years, it's really my great pleasure to know a lot of friends and colleagues at UD. My classmates make my life more colorful. Special thanks to all Chen lab group members for their help, encouragement and advices. I would also like to thank ARPA-E for the financial support.

Last but not least, I would like to thank my family and my girlfriend for their caring, understanding and support along I finish the thesis. They are always with me through the good and bad times and support all my decisions!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Multi-enzyme cascade reactions are common in nature. Compared to free-floating enzyme catalysis, enzyme co-localization, existing in some natural systems, can improve the efficiency of the reactions thanks to kinetic benefits. In engineering multi-enzyme reactions, multi-functional enzyme complexes can be constructed to mimic nature's synergy and "substrate-channeling" behavior.

This thesis focuses on engineering the spatial organization of enzymes using an antibody mimic, the 10^{th} human fibronectin type III (FN3) domain, as the building block. Due to the small size, good solubility and stability of FN3, along with the diversity of FN3-binders, a complex protein scaffold platform can be built. By employing the specific and tight binding interactions between FN3-binders and their corresponding monobodies (mutated FN3), FN3-binder fused enzymes can be docked onto the protein scaffold and result in the formation of synthetic "substrate-channels". Two particular pathways have been studied and executed. First, the scaffold was designed to be displayed on the cell surface of *S. cerevisiae* for designer cellulosome construction which aims at efficient hydrolysis of biomass. Experimental results showed that three engineered FN3 domains have orthogonal and high-affinity binding with their targets. Second, the scaffold was designed to be applied in *E.coli* cells to assemble the metabolic enzymes involved in the conversion of methanol for the production of n-butanol in order to achieve a metabolic flux enhancement.

Chapter 1

INTRODUCTION

1.1 Multi-functional Enzyme Complex

With the rapid development of systems biology, synthetic biology and metabolic engineering, there is an ever-increasing interest in developing and optimizing biological routes for the production of valuable compounds. Despite all these efforts, a large number of engineered pathways are characterized by low productivity (Atsumi et al. 2008) and imbalanced flux (Pitera et al. 2007), as well as by undesirable side reactions (Jiang et al. 2005). To solve these problems, it is important to learn how natural biological systems employ spatial organization of metabolic enzymes as a strategy for efficient metabolite production (Conrado et al. 2008).

Based on studies of systems ranging from the pyruvate dehydrogenase complex and the enzymes of glycolysis to the machinery of fatty acid oxidation and amino acid biosynthesis (Ovadit & Srere 1996), the co-localization of multiple-enzymes appears to be a common feature of cellular metabolism in all organisms. By using this co-localization strategy, the metabolism can successfully proceed and even be facilitated by metabolite channeling, which directly transfers the intermediate from one enzyme to an adjacent enzyme without the need of free aqueous-phase diffusion. This metabolite channeling decreases the transit times of intermediates and prevents intermediate accumulation (Ovádi & Saks 2004). Furthermore, the enzymes are brought into close proximity by co-localization so that they can act synergistically on

the substrate. From all these kinetic benefits, the substrate conversion will be enhanced. Inspired by nature, we can engineer the spatial organization of multi-enzyme systems so that synergy can be achieved among the targeted enzymes.

Basically, there are three strategies to build artificial multi-enzyme systems. Enzyme encapsulation on polymer materials is the traditional technique, but it is seriously limited by the cost associated with the scale-up of cell-free systems. Thus, it would be preferable to organize enzymes inside the living cells because cell-based systems are easier to manipulate and scale-up. One strategy is the creation of chimeric proteins by combining two or more distant genetic elements with a short linker sequence to generate a single multifunctional polypeptide. However, it has been shown that large fusions of enzymes are problematic because of misfolding and proteolysis of the fusions in heterologous hosts. Thus, an alternative to protein fusions is to bring the enzymes into close proximity using post-translational assembly (Conrado et al. 2008). This is the strategy we employ in this thesis. The synthetic scaffolds built from modular parts can assemble the pathway enzymes into the system in a programmable manner using engineered protein-protein interactions.

The best example for post-translational multi-enzyme assembly is the natural cellulosome system in some cellulolytic anaerobes. Typical bacterial cellulosomes are composed of a scaffoldin that contains at least one cellulose-binding domain (CBD) and a number of cohesin modules, which tightly bind to the complementary dockerin modules borne by the catalytic cellulases. In this way, cellulases are brought close to each other with optimized ratios allowing them to act synergistically on crystalline cellulose, resulting in a more efficient architecture than the corresponding free enzyme systems (Mingardon et al. 2007).

1.2 The 10th Human Fibronectin Type III (FN3) Domain

Antibodies are the most common platforms for protein-protein interaction studies, but as large, multichain, disulfide-stabilized proteins, they tend to misfold in the reducing environment of the cytoplasm (Wojcik et al. 2010). Alternatively, the 10th human fibronectin type III domain (referred to simply as FN3 hereinafter) is a good candidate for protein scaffolds as an antibody mimic. FN3 is a monomeric, highly stable, 94-residue domain, with an immunoglobulin-like fold formed by seven antiparallel β-strands (Figure 1). Unlike most of the immunoglobulin superfamily members, FN3 is disulfide-bond free and overexpression of correctly folded FN3 in *E. coli* is straightforward. It has three loop regions (BC, DE and FG) that are analogous to antibody complementary determining regions (CDRs). Residues in these loops can tolerate mutations without impairing the overall stability enabling us to generate monobodies (mutated FN3) that can bind to virtually any given target (FN3-binder) (Koide et al. 1998). Some of the FN3-binders generated in literature are listed in Table 1.

Because of the small size, good solubility of FN3 and homogeneity (Just several amino acids are diversified in the three loops between repeating FN3 modules.) of FN3-based scaffolds, we expect to build larger tandem protein scaffolds without any concern about misfolding and proteolytic problems. This is supported by a previous report indicating that a protein scaffold composed of eight repeating FN3 domains can be solubly expressed in *E. coli* without any degradation (Baca et al. 2011). By employing the specific interaction between a monobody and its corresponding binder, we can construct FN3-based scaffolds to assemble multi-enzyme complexes. For constructing multi-enzyme assemblies, FN3-binders must have the following

characteristics: monomeric, small size, good expression and solubility, high and specific binding affinity to the corresponding monobody. The FN3-binders generated in literature with the desired characteristics are highlighted in Table 1. Additional suitable binding pairs can be screened for by diversifying the amino acids in BC, DE and FG loops. The modular nature of this system makes it highly flexible because we can simply construct different FN3-binder tagged enzymes to engineer different multi-enzyme cascades.

Chapter 2

CREATION OF DESIGNER CELLULOSOMES USING FN3-BASED PROTEIN SCAFFOLDS FOR EFFICIENT CELLULOSE HYDROLYSIS

2.1 Introduction

To lessen the nation's dependence on petroleum oil, ethanol is a promising candidate for transportation fuels. Cellulosic biomass is the most abundant and sustainable raw material for biofuel production because of its high sugar content. It's estimated that about 1.4 billion tons of cellulosic biomass can be fed into bioenergy and bioproduct industries throughout the United States per year (Office et al. 2005). To make bio-ethanol more competitive with fossil fuels in the market, technologies have to be employed to reduce its production cost.

Because of the recalcitrant nature of cellulose, the synergistic actions of endoglucanase, exoglucanase and β-glucosidase are required for the complete hydrolysis of cellulose to glucose. Large quantity supplements of expensive cellulase enzymes impede the cost reduction of bio-ethanol. A method called consolidated bioprocessing (CBP) which combines enzyme production, cellulose saccharification and ethanol fermentation was proposed several years ago for the low-cost production (Lynd et al. 2008). *Saccharomyces cerevisiae* is an attractive candidate for CBP because of its high ethanol productivity and its inherent high tolerance to ethanol (Lynd et al. 1999). Previously in our group, an artificial cellulosome based on the species specific interaction between dockerins and cohesins was constructed and

displayed on the yeast cell surface inspired by the natural cellulosome systems in some anaerobic microbes (Tsai et al. 2009). A dockerin/cohesin-based trifunctional minicellulosome which specifically accrues exoglucanase CelE from *C. cellulolyticum*, endoglucanase CelA from *C. thermocellum* and β-glucosidase BglA from *R. flavefaciens*, was effectively displayed on the yeast cell surface. This work aimed at realizing simultaneous saccharification and fermentation of cellulose for ethanol production by endowing yeast cells with the ability of efficient cellulose hydrolysis (Tsai et al. 2010). However, the creation and display of complex artificial cellulosomes with more than three cohesins on the yeast cell surface have not been realized probably because of the complexity of folding the larger chimeric scaffoldins (Tsai et al. 2013). To circumvent this problem, we proposed to use FN3 interaction pairs as alternative building blocks for complex cellulosome assembly.

2.2 Materials and Methods

2.2.1 Strains and plasmids

All procedures for DNA manipulation were performed according to standard methods (Sambrook and Russel, 2003). The high fidelity Phusion DNA polymerase (Thermo Scientific, Odessa, TX) and Taq DNA polymerase (Promega, Madison, WI) were used for PCR amplification with a S1000^{TM} Thermal Cycler (Bio-Rad, Hercules, CA). *Escherichia coli* strain DH5α (F– Φ80*lacZ*ΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1) was used as the host for genetic manipulations.

To construct the expression vector for BglA-ySUMO-his6, the gene coding for ySUMO was obtained by PCR from plasmid pET14b-SUMO-FynSH3 (Huang et al.

2012), which was a gracious gift from Dr. Brian K. Kay from University of Illinois at Chicago, with the forward primer 5'- AAT TCC CCG CGG ATG AGC GAC TCG GAA GTG -3' and the reverse primer 5'- ATA CCG CTC GAG ACC GCC AAT TTG TTC ACG -3'. The amplified gene fragment was cloned into *Sac*II and *Xho*I linearized plasmid pET24a-BglA-PE1A-his6 (Sun et al. 2014) to generate plasmid pET24a-BglAySUMO. For the construction of pET24a-CelEGFP, GFP fragment was PCRed from pET24a-GFP with the forward primer 5'- ATA CAT GAA TTC GGA TCC GGC GGC AGC AGC CCG AGC ACC CCG CCG ACC CCG AGC CCG AGC ACC CCG CCG GGC GGC AGC CCG CGG ATG GTG AGC AAG GGC -3' and reverse primer 5'- ATACAT CTC GAG CTT GTA CAG CTC GTC CAT -3'. The amplified gene fragment was inserted into *Sac*II and *Xho*I linearized plasmid pET24a-CelE-Dc-his6 (Tsai et al. 2009) to generate plasmid pET24a-CelEGFP. To construct pET24a-CelAMBP, MBP PCR fragment was inserted into *Sac*II and *Xho*I linearized plasmid pET24a-CelA-Dt-his6 (Tsai et al. 2009). The forward primers were 5'- CCG ACC CCG AGC CCG AGC ACC CCG CCG GGC GGC AGC CCG CGG ATG AAA ATC GAA GAA GGT AAA CTG -3' and 5'- GTC GAC GAA TTC GGA TCC GGC GGC AGC AGC CCG AGC ACC CCG CCG ACC CCG AGC CCG -3'. The reverse primer was 5'- GTG GTG CTC GAG AGT CTG CGC GTC TTT C -3'.

For the construction of the FN3 domain yeast cell surface displaying vector, the genes for three monobodies were synthesized using IDT gBlocks. Then the gene fragment was amplified by PCR. Because these monobodies have the same framework, the forward primer 5'- AAT ACCC CCG GGT CAT ATG GCC GTT TCT GAT GTT-3' and the reverse primer 5'- AAT CCG CTC GAG TTA GCT GGT ACG GTA GTT

AAT-3' were the same. Amplified gene fragment was cloned into *Nhe*I and *BamH*I linearized plasmid pCTCON2-scaf3 (Tsai et al. 2009).

2.2.2 Expression of BglA-ySUMO-his6, CelE-GFP-his6 and CelA-MBP-his6 enzyme fusions

The three enzyme fusions were expressed in *E. coli* BLR [F- *omp*T *hsd*SB (rB-mB-) *gal dcm*(DE3) Δ(*srl-rec*A)306::Tn10(*Tet*R); Novagen, Madison,WI]. Overnight cultures were inoculated into 25 ml Lysogeny broth (LB) media (10.0 g/liter tryptone, 5.0 g/liter yeast extract, 10.0 g/liter NaCl) with an initial OD_{600} of 0.05 supplemented with 50 μg/ml kanamycin and incubated at 37˚C until OD_{600} reached 0.5. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.25 mM and cells were grown at 20˚C overnight. Cells were then harvested by centrifugation, resuspended in phosphate-buffered saline (PBS; 8.0 g/liter NaCl, 0.2 g/liter KCl, 1.44 g/liter Na_2HPO_4, 0.24 g/liter KH_2PO_4), and lysed by ultrasonic disruption using a sonicator. The cell lysate was centrifuged to remove insoluble cell debris.

2.2.3 Display of monobodies on the yeast cell surface

S. cerevisiae strain EBY100 [*MATa AGA1::GAL1-AGA1::URA3 ura3-52 trp1 leu2-1 his3-200 pep4::HIS3 prb1-1.6R can1 GAL*] was used for surface display of monobodies. Yeast cells harboring pCTCON2-FN3(ySUMO)-myc, pCTCON2-FN3(GFP)-myc, pCTCON2-FN3(MBP)-myc were precultured in SDC medium (20.0 g/liter dextrose, 6.7 g/liter yeast nitrogen base without amino acids, 5.0 g/liter Casamino Acids) for 18.0 hrs at 30˚C. These precultures were subinoculated into 25 ml SGC medium (20.0 g/liter galactose, 6.7 g/liter yeast nitrogen base without amino

acids, 5.0 g/liter Casamino Acids) with an initial OD600 of 0.1 and grown for 48 hrs at 20°C.

2.2.4 Docking enzyme fusion onto yeast cell surface displayed monobody

E. coli cell lysates containing BglA-ySUMO-his6, CelE-GFP-his6 or CelA-MBP-his6 were mixed and incubated with yeast cells displaying monobody for 1hr at 4°C in PBS. After incubation, yeast cells were washed and harvested by centrifugation (3,000 g, 10 min) at 4°C and resuspended in the same buffer for further use.

2.2.5 Immunofluorescence microscopy

Yeast cells displaying monobodies were harvested by centrifugation, washed with PBS, and resuspended in 250 μL of PBS containing 1 mg/ml BSA and 0.5 μg of anti-c-Myc or anti-His immunoglobulin G (IgG; Invitrogen) for 4 hrs at room temperature with occasional mixing. Cells were then pelleted and washed with PBS before resuspension in PBS plus 1 mg/mL BSA and 0.5 μg anti-mouse IgG conjugated with Alexa 488 (Molecular Probes). After incubation for 2 hrs, cells were pelleted, washed twice with PBS, and resuspended in PBS to an OD600 of 1. Whole-cell fluorescence was measured with a fluorescence microplate reader (Synergy4; BioTek, VT) at an excitation wavelength of 485 nm and an emission wavelength of 535 nm. The sensitivity was adjusted to be 100. For fluorescence microscopy (Olympus BX51), 10 μL cell suspensions were spotted onto slides and a coverslip was added. Images from Alexa 488 were captured with the QCapture Pro6 software.

2.3 Results and Discussion

To synergistically enhance cellulose hydrolysis using exoglucanase, endoglucanase and β-glucosidase, we envision to construct and display a trifunctional FN3-based scaffoldin with an internal CBD on the yeast cell surface to specifically recruit the three cellulases (Figure 2). The aim of this work is to demonstrate that the FN3-based scaffold can be used for designer cellulosome constructions. The engineered yeast strains displaying FN3-based trivalent mini-cellulosome will be compared to the analogous trivalent cohesin/dockerin cellulosome in terms of cellulose hydrolysis and ethanol production, with the expectation that we can achieve comparable synergistic effect among the three cellulases (Tsai et al. 2009).

According to literature, monobodies FN3(GFP), FN3(MBP) and FN3(SUMO) can bind onto GFP, MBP and ySUMO, respectively, with high affinity (Gilbreth et al. 2011; Koide et al. 2012; Gilbreth et al. 2008). Before constructing the full trifunctional FN3-based mini-cellulosome using these three binding interaction pairs, the binding capacities and specificities of monobodies displayed on the yeast cell surface must be checked. Therefore, we displayed three monobodies on the yeast cell surface and checked their binding with their corresponding binders, which were visualized by immunofluorescence microscopy (Figure 3). As shown in Figure 4, the fluorescence intensities resulting from the his6-tag and c-Myc tag demonstrated that those three monobodies remain functional when displayed on yeast cell surface and the bindings between three monobodies are orthogonal. The specificity between different monobody pairs is very important for designer cellulolosome assembly, especially when we want to control the sequence and stoichiometry of different cellulases on the scaffoldin.

2.4 Future Work

The next step in this work is to use these three monobodies to construct a trifunctional mini-scaffoldin for cellulose hydrolysis as a proof of concept. To further improve the design of the synthetic scaffoldin, it is also important to consider the contribution of other components in raw lignocellulosic biomass. Natural cellulose degradation is severely impeded by cellulose encapsulation within a hemicellulose (mainly xylan) matrix; thus, xylan degradation is desirable. Research has shown that enhanced xylanase-mediated xylan solubilization increased the accessibility of cellulases to cellulose by exposing new cellulose chains, thereby amplifying cellulose hydrolysis (Moraïs et al. 2010). Furthermore, xylan is the second most abundant renewable polysaccharide in nature after cellulose. Simultaneous saccharification and fermentation of cellulose and xylan can strengthen the overall economics of processing lignocellulosic biomass. To further improve the functionality of the FN3-based designer cellulosome and its efficiency in hydrolyzing complex polysaccharides (e.g. wheat straw), we will introduce xylanases' synergistic contribution into the system. Using cohesin-dockerin interactions, Morais et al demonstrated that cooperation among xylanases and cellulases can be promoted by their integration into a single designer cellulosome. Up to 1.6-fold improvement in untreated wheat straw degradation was achieved by the designer cellulosome compared to those of the wild-type, free enzymes (Moraïs et al. 2010). In the future, new enzymatic activities can be evaluated by introducing desirable synergistic contributions for a more complex FN3-based designer cellulosome construct. The ultimate goal is to approach the rates of cellulose degradation exhibited by natural cellulosomes.

Chapter 3

APPLICATION OF FN3-BASED PROTEIN SCAFFOLD IN E. COLI CELLS FOR THE METABOLIC FLUX ENHANCEMENT

3.1 Introduction

Recent increase in the supply of natural gas and a growing price spread between natural gas and petroleum spurred research and technology developments on natural gas-to-liquid (GTL) fuels, because the inherent low volumetric energy density of natural gas impedes its usage in transportation (Conrado & Gonzalez 2014). Bioconversion is promising in addressing this problem because of its high specificity and high process energy efficiency under mild conditions and also for its low capital expenses (Haynes & Gonzalez 2014). The economic analysis of natural gas bioconversion to liquid fuels revealed that its economic viability depends on three major breakthroughs: high efficiency activation of methane, high efficiency biosynthesis of liquid fuel from methanol derived from methane activation and downstream design of efficient bioreactor (Haynes & Gonzalez 2014). This chapter focuses on high-efficiency conversion of methanol to liquid fuel n-butanol using engineered methylotrophic *E. coli* cells.

The first key step of methanol utilization is the oxidation of methanol to formaldehyde. Some Gram-negative bacteria use pyrroloquinoline quinone (PQQ) and periplasmic methanol dehydrogenase (MDH) to oxidize methanol (Keltjens et al. 2014) whie the MDHs in some Gram-positive methylotrophs, such as Bacillus strains, are

NAD^+-dependent and function in the cytoplasm (Krog et al. 2013). Since *E. coli* cells natively lack PQQ, it is a better choice to heterogeneously express NAD^+-dependent MDHs to engineer methylotrophic *E. coli* cells (Matsushita et al. 1997; Stites et al. 2000). To assimilate cytotoxic formaldehyde into the central metabolism, a synthetic ribulose monophosphate (RuMP) pathway can be introduced into *E. coli* cells. In the RuMP pathway, formaldehyde in the presence of ribulose-5-phosphate will be converted into D-arabino-3-hexulose-6-phosphate, which will be isomerized into fructose-6-phosphate. The two enzymes catalyzing these steps, 3-hexulose-6-phosphate synthase (HPS) and 6-phosphate-3-hexuloisomerase (PHI), have to be expressed heterogeneously (Ato et al. 2006). Then fructose-6-phosphate will go into the central metabolism via glycolysis and fed into the engineered n-butanol production pathway (Shen et al. 2011) or be used to regenerate the formaldehyde acceptor ribulose-5-phosphate via the pentose phosphate pathway (Ato et al. 2006). The main pathways are shown in Figure 5.

There are several technical challenges involved in the bioconversion of methanol into butanol. First, both energy demands and biomass requirements of host cells must be met by a C1 carbon source. Second, the product of methanol activation, formaldehyde, is cytotoxic to *E. coli* cells. Low cytoplasmic concentration has to be maintained. Third, the NAD^+-dependent MDHs naturally favor the reduction of formaldehyde to methanol instead of the reverse reaction (Krog et al. 2013). Therefore, a fast rate of formaldehyde condensation into hexulose-6-phosphate (formaldehyde "kinetic trap") is desired to achieve fast rate of methanol conversion. All these challenges can potentially be solved by using a protein scaffold to accrue the pathway enzymes in the hope of achieving higher metabolic flux (Figure 6).

The feasibility of protein scaffolds used for enhancing metabolic fluxes has been demonstrated previously (Dueber et al. 2009; Moon et al. 2010). Dueber et al (2009) utilized SH3, PDZ and GBD domains as building blocks to construct synthetic protein scaffolds and applied them in mevalonate producing *E. coli* strains. The native mevalonate biosynthetic pathway suffers from flux imbalance between HMG-CoA synthase (HMGS) and HMG-CoA reductase (HMGR) which results in the accumulation of cytotoxic intermediate HMG-CoA (Pitera et al. 2007). By optimizing synthetic scaffold expression levels (balancing the scaffold to enzyme ratio, Good et al. 2011; Whitaker & Dueber 2011), the stoichiometry of three mevalonate biosynthetic enzymes recruited to the synthetic complex and the composition of protein scaffolds, they engineered a mevalonate producing *E. coli* strain with 77-fold improvement in product titer, reduced metabolic burden and faster growth because of lowered levels of the cytoplasmic HMG-CoA concentration. The same set of scaffolds was applied inside *E. coli* cells to further improve glucaric acid productivity by Moon et al (2010) despite of its high production level (1g/L). After the same optimization steps, the glucaric acid productivity was enhanced by about 5-fold compared with the non-scaffolded control.

For the assembly of a multifunctional enzyme complex, interaction affinity of binding pairs used as building blocks is a crucial factor because low-affinity interaction may result in incomplete complex assembly. As indicated in Dueber et al (2009), only about 5% of the scaffold was docked with all three enzymes because of low interaction affinity (Kd ~ 1 μM) of the three binding pairs. One notable benefit of FN3 interaction pairs is their high binding affinity (in the range of low nM, see Table 1), which can be further improved by directed evolution if necessary.

3.2 Materials and Methods

3.2.1 Expression and purification of *B. methanolicus* MGA3 MDH3 and *M. gastri* HPS

The gene fragments for *B. methanolicus* MGA3 MDH3 and *M. gastri* HPS were cloned into pETM6 plasmid vector (Xu et al. 2012) using *Nde*I and *Spe*I restriction sites. During cloning, a His6-tag was added onto the N-terminal of both enzymes. They were expressed in *E. coli* BLR [F- *omp*T *hsd*SB (rB-mB-) *gal dcm* (DE3) Δ(*srl-recA*)306::Tn10(TetR); Novagen, Madison,WI]. Overnight cultures were inoculated into 25 ml Lysogeny broth (LB) medium (10.0 g/liter tryptone, 5.0 g/liter yeast extract, 10.0 g/liter NaCl) with an initial OD_{600} of 0.05 supplemented with 100 μg/ml ampcillin and incubated at 37˚C until OD_{600} reached 0.5. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.20 mM and cells were grown at 37˚C for another 3 hrs. Cells were then harvested by centrifugation, resuspended in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, PH=7.4) and lysed by ultrasonic disruption using a sonicator. The cell lysate was centrifuged to remove insoluble cell debris. The two enzymes were purified using Ni-NTA column and dialyzed overnight at 4˚C in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, PH=7.4). Enzyme purity was checked by running an SDS-PAGE and enzyme concentration was determined by the method of Braford using BSA as a standard.

3.2.2 Conditions for MDH and HPS activity assays

To check the activity of MDH on formaldehyde, 0.1 mM HCHO, 0.5 mM NADH and 1 μM his6-MDH were mixed in ice-cold HEPES buffer solution followed

by incubation at 37˚C. At the time points of 0 min, 5 min and 10 min, aliquots were taken out and put on ice until HCHO concentration measurements were taken. For HPS assay on formaldehyde, it was assayed by measuring the ribulose-5-phosphate dependent disappearance of formaldehyde. Because of the high price and inhibitory effect of commercial ribulose-5-phosphate (Britain & Kemp 1974), ribulose-5-phosphate was generated *in situ* for each assay from ribose-5-phosphate by the action of phosphoriboisomerase. The assay mixture was in 50mM HEPES buffer solution containing 5 mM $MgCl_2$, 0.1 mM formaldehyde, 6 U/ml of phosphoriboisomerase, 5 mM D-ribose 5-phosphate (sodium salt) and was incubated for 15 min at 30°C. After the preincubation, 5 μM or 20 μM purified his6-HPS was added and formaldehyde consumption was followed at 37°C by taking timed samples from the assay mixture. Formaldehyde disappearance not dependent on 3-hexulose phosphate synthase activity was corrected for by running parallel assays lacking ribose 5-phosphate. After each timed sample, $HClO_4$ was added to a final concentration of 0.5 M to stop further formaldehyde consumption. The PH value of the sampled solution was brought to neutral with KOH. Formaldehyde concentration was determined colorimetrically by the method of Nash (1953).

To check MDH activity on methanol and assess the effect of HPS to MDH ratio on methanol conversion rate, ribulose-5-phosphate was generated *in situ* as stated before. In the reaction mixture were 2 M methanol, 5 mM NAD^+, 5 mM ribulose-5-phosphate generated, 4 μM MDH and corresponding concentration of HPS. The reaction was carried out in a Corning 96-well microtiter plate and incubated in 37°C shaker. Since methanol is highly volatile, the plate was covered with plastic film. At

certain time points, absorbance at 340nm was measured to track the production of NADH.

3.3 Preliminary Results and Discussion

Although the biological relevance of formaldehyde reduction by methanol dehydrogenases remains elusive, almost all NAD^+-dependent MDHs favor formaldehyde reduction instead of the reverse direction (Krog et al. 2013). Also, formaldehyde is highly cytotoxic to *E. coli* host cells. Given these facts, an extremely low concentration of formaldehyde has to be maintained, and it is therefore desired to construct a formaldehyde "kinetic trap" surrounding MDH to "drain" away formaldehyde produced by MDH. To proceed, several questions have to be answered first. What concentration of formaldehyde are we targeting? To maintain the targeted formaldehyde concentration, what HPS to MDH ratio must be used? Or rather, what are the numbers for x and y needed in the scaffold configuration (Figure 6)? As a proof of this formaldehyde "kinetic trap" concept, we simply mixed HPS with MDH in different ratios, expecting to see that higher ratios would result in faster NADH production. The result we obtained confirmed our hypothesis (Figure 7). When 2M methanol is used, the equilibrium concentration of NADH or formaldehyde is only about 250 μM, which is consistent with a previous study and conclusion that MDH3 favors formaldehyde reduction (Krog et al. 2013). Another extracted piece of information is that when formaldehyde concentration accumulates to about 100μM, the reverse reaction starts to dominate. Thus the targeted formaldehyde concentration should be around 100 μM. To further determine the minimum HPS to MDH ratio, their competence or specific activities on formaldehyde need to be compared (Figure

8). We can see that the specific activity of HPS on formaldehyde is 1/4 of that of MDH.

$$CH_3OH \underset{k_2}{\overset{k_1}{\rightleftarrows}} HCHO \xrightarrow{k_3} H6P$$

When there is no HPS: $\frac{d[\mathrm{HCHO}]}{dt} = \frac{k_{cat}[\mathrm{MDH}][\mathrm{CH_3}OH]}{K_m + [\mathrm{CH_3}OH]} - k_2[\mathrm{MDH}][HCHO]$

At equilibrium, $[\mathrm{HCHO}] = \frac{1}{k_2[\mathrm{MDH}]} \times \frac{k_{cat}[\mathrm{MDH}][\mathrm{CH_3}OH]}{K_m + [\mathrm{CH_3}OH]} = 250\mu M$.

After we introduce HPS into the system: $\frac{d[\mathrm{HCHO}]}{dt} = \frac{k_{cat}[\mathrm{MDH}][\mathrm{CH_3}OH]}{K_m + [\mathrm{CH_3}OH]} - k_2[\mathrm{MDH}][HCHO] - \mathrm{k_3}[\mathrm{HPS}][\mathrm{HCHO}]$. At steady state, $[\mathrm{HCHO}] = \frac{1}{k_2[\mathrm{MDH}] + \mathrm{k_3}[\mathrm{HPS}]} \times \frac{k_{cat}[\mathrm{MDH}][\mathrm{CH_3}OH]}{K_m + [\mathrm{CH_3}OH]} = 100\mu M$. From the calculation, we can get $k_3[\mathrm{HPS}] = 1.5k_2[\mathrm{MDH}]$. Because the specific activity of HPS is 1/4 of that of MDH, or $k_3 = 0.25k_2$, [HPS]=6[MDH].

From the preliminary results and calculation, we can conclude that six times excess of HPS has to be used compared with MDH to maintain a low formaldehyde concentration. This information is critical to break the thermodynamic equilibrium of MDH converting formaldehyde back into methanol and to maintain normal growth of *E. coli* host cells.

3.4 Future Work

In response to the problem of diminishing oil, people started to shift to more-abundant natural gas and proposed the Methanol Economy concept (Olah 2013). There has been a lot of progress in the research of natural gas-to-liquid (GTL) bioconversion technologies. Bogorad et al (2014) demonstrated a biocatalytic pathway,

termed the methanol condensation cycle (MCC), converting methanol to ethanol or n-butanol using a cell-free system. Although the carbon yields from methanol to ethanol or n-butanol are satisfying, the conversion rate and percentage of methanol are far away from being enough. A recent study by (Müller et al. 2015) showed that engineered methylotrophic *E. coli* cells were endowed with the ability of utilizing methanol and integrating it into the central metabolism by using ^{13}C-labling experiments. However, the synthetic methylotrophic *E. coli* cells fail to grow on methanol. The underlying reasons might be the depletion of intermediate ribulose-5-phosphate, which helps to assimilate formaldehyde produced from methanol, or the accumulation of formaldehyde, resulting from pathway imbalances or poor metabolic pathway flux from C1 methanol.

To solve the problem of ribulose-5-phosphate depletion, we can mimic the natural methylotrophic bacteria. Jakobsen et al (2006) studied the gene transcription of methylotrophic bacterium *Bacillus methanolicus* growing on methanol, found that those genes, *glpX*, *fba*, *tkt*, *pfk* and *rpe* involved in the pentose phosphate pathway were upregulated. This might be resulting from requirement of ribulose-5-phosphate regeneration from fructose-6-phosphate. Thus, those five genes will be cloned into a plasmid and expressed in *E. coli* host cells in order to supply enough ribulose-5-phosphate.

After we get to the milestone of growing synthetic methylotrophic *E. coli* cells on methanol, the problems of formaldehyde accumulation and poor pathway flux can potentially be overcome by using the protein scaffolding strategy (Figure 7). The metabolic enzymes will be fused to different FN3-binders. To ensure their expression levels and specific activity not being compromised, a combinatory approach can be

employed to get the best match between metabolic enzymes and FN3-binders. Since Dueber et al (2009) and Moon et al (2010) demonstrated that the scaffold architecture and composition are crucial for metabolic pathway enhancement, varieties of FN3-based protein scaffolds will be co-expressed with metabolic enzyme fusions. The consumption rate of methanol and growth of *E. coli* host cells will be compared. Then the genes involved in the n-butanol production pathway (Shen et al. 2009) will be introduced into the winning *E. coli* strain followed by study of methanol conversion to n-butanol.

Chapter 4

SUMMARY AND CONCLUSION

In this thesis, we explored the feasibility of using the engineered 10^{th} human fibronectin type III domain (FN3) to construct protein scaffolds for the assembly of multi-functional enzyme complexes. Both *in vitro* and *in vivo* applications were discussed.

Because the 10^{th} human fibronectin type III domain is extremely stable, single-domain protein with an immunoglobulin-like fold but lacking disulfide bonds, it has been engineered to bind a variety of targets (S. Koide et al. 2012). By using the technique of yeast cell surface display, the author confirmed that the three FN3 domains targeting ySUMO, MBP and GFP can bind with their corresponding target with high-affinity and orthogonality. Many other orthogonal FN3 interaction pairs can be discovered by using the same technique or directed evolution from scratch. Chapter two discussed the potential of using FN3 domains to construct complex designer cellulosomes displayed on the yeast cell surfaces for the efficient hydrolysis of plant cell wall. Unlike cohesin-dockerin requiring free calcium ion for high-affinity interaction, the high-affinity binding of FN3 pairs is independent from ions which makes engineered FN3 domains as perfect candidate for the construction of protein scaffolds to assemble metabolic pathway enzymes. Chapter three discussed the potential of FN3-based protein scaffolds in enhancing metabolic pathway flux from

methanol to n-butanol in synthetic methylotrophic *E. coli* cells. For the future work, the benefit of these protein scaffolds will be assessed.

TABLE

Table 1. Examples of FN3-binders reported in literatures. Those suitable for constructing post-translational multi-enzyme complexes are highlighted in the red box. SUMO: small ubiquitin-like modifier; SH2: Src homonogy 2; SH3: Src homonogy 3; TNF: tumor necrosis factor; LBD: ligand binding domain

FN3-binder	Monomer size (kDa)	Active state	Binding affinity (Kd)	Reference
Yeast SUMO	11.3	monomer	45 nM	Gilbreth et al. 2011
Abelson SH2	12	monomer	7 nM	Wojcik et al. 2010
Fyn SH3	7	monomer	166 nM	Huang et al. 2012
Maltose binding domain (MBP)	42.5	monomer	5.7 nM	Gilbreth et al. 2008
GFP	27	monomer	31 nM	Koide et al. 2012
Human SUMO1	11.5	monomer	10 nM	Gilbreth et al. 2011
Human SUMO4	10.7	monomer	7 nM	Gilbreth et al. 2011
Lysozyme	15	monomer	2.6 pM	Hackel et al. 2008
Ubiquitin	8.5	monomer	5 uM	Koide et al. 1998
TNF-α	17.4	trimer	20 pM	Xu et al. 2002
Estrogen receptor α-LBD	62.8	Dimer	Not reported	Koide et al. 2002
Human αvβ3 integrin	198	heterodimer	330 pM	Richards et al. 2003

FIGURES

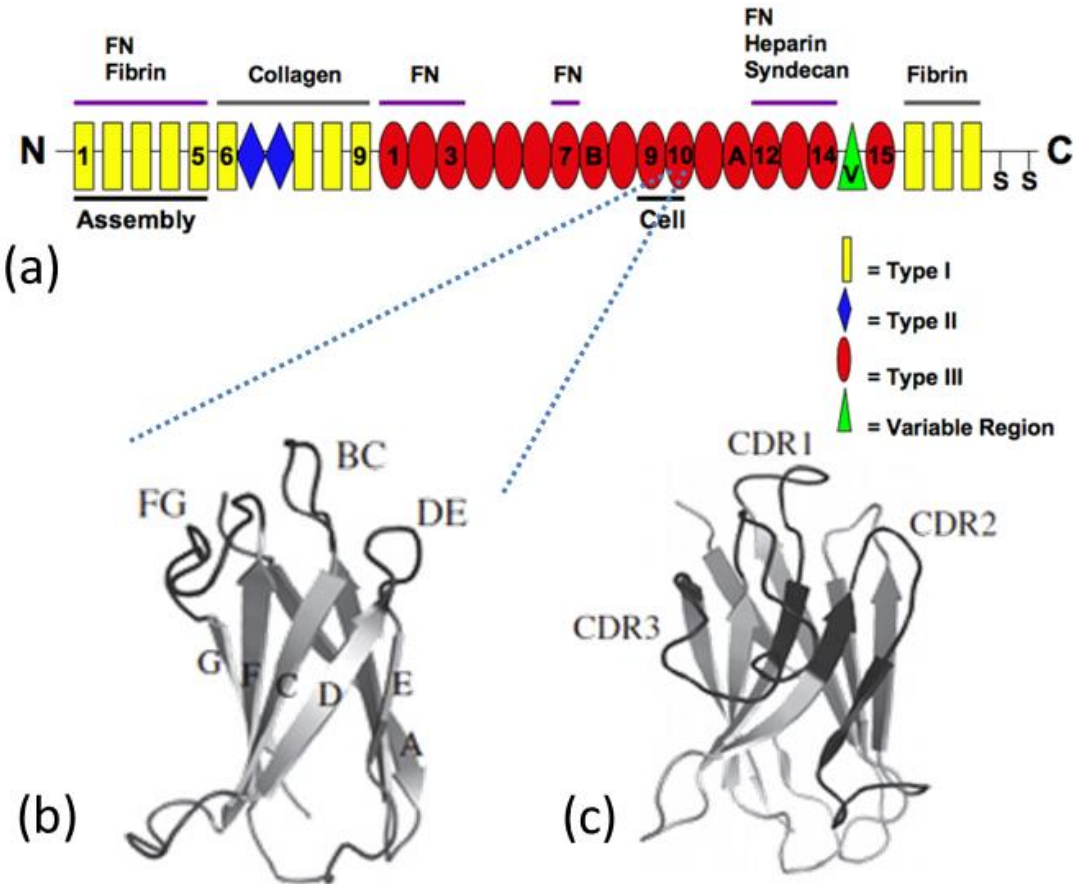


Figure 1. Makeup of human fibronectin and structure of the FN3 domain. (a) Schematic of human fibronectin protein. It consists of many repeats of three types of small fibronectin domains (adapted from http://en.wikipedia.org/wiki/Fibronectin). (b) 3-D structure of human tenth fibronectin type 3 domain (adapted from Koide et al. 1998). (c) 3-D structure of immunoglobulin VH domain. The locations of complementary determining regions (CDRs) are indicated.

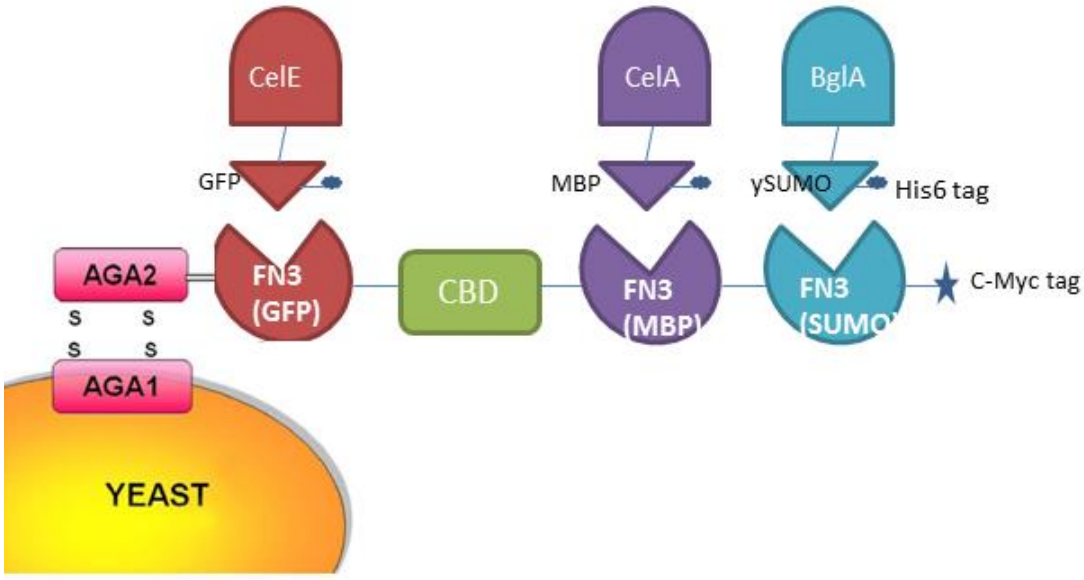


Figure 2. Construction of trifunctional mini-cellulosome using engineered FN3 domains and display it on the yeast cell surface for efficient biomass hydrolysis.

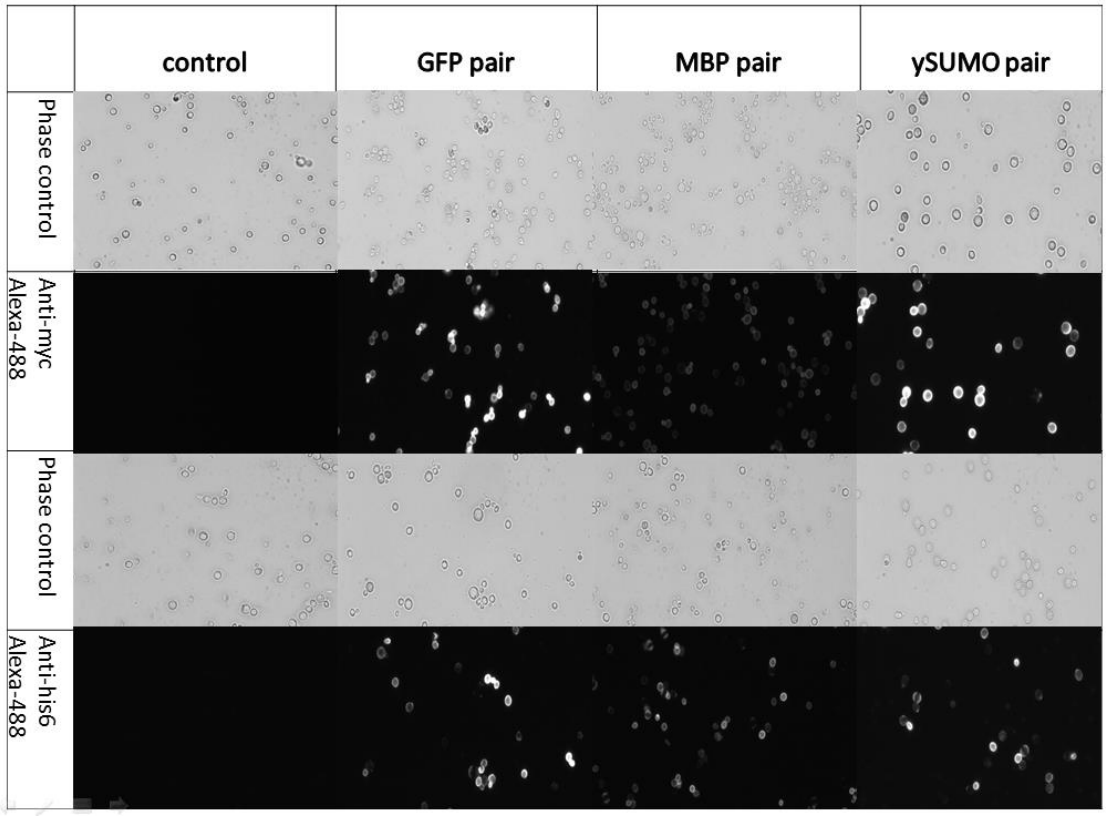


Figure 3. Phase control and immunofluorescence micrographs of yeast cells displaying monobodies. After binding experiment, cells were probed with anti-C-myc or anti-His serum and fluorescently stained with a goat anti-mouse IgG conjugated with Alexa Fluor 488. Yeast cells EBY100 displaying nothing were used as control.

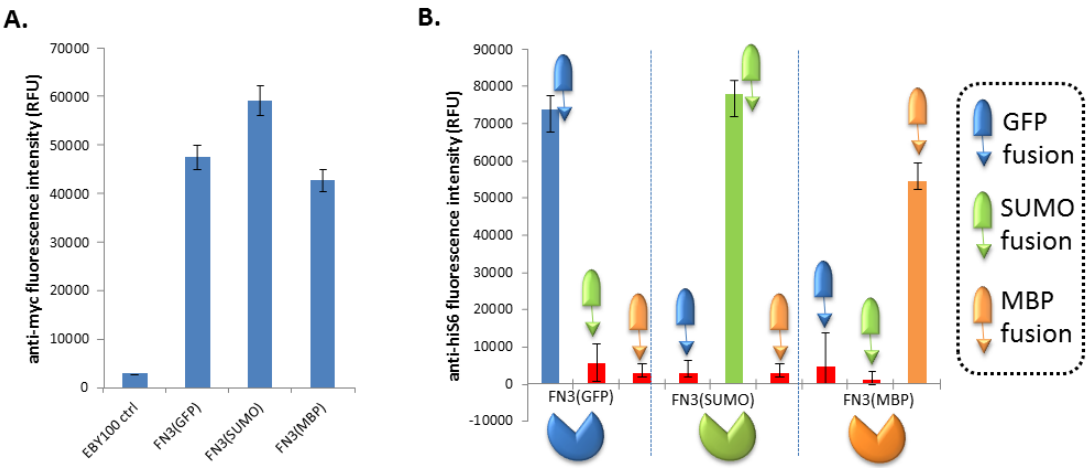


Figure 4. Whole cell fluorescence intensity of yeast cells displaying monoblodies or with different cellulase enzyme fusions docked on the displayed monobodies. (A) Cells were probed with anti-C-myc serum and fluorescently stained with a goat anti-mouse IgG conjugated with Alexa Fluor 488. Yeast cells EBY100 displaying nothing were used as control. (B) After binding experiment, cells were probed with anti-His serum and fluorescently stained with a goat anti-mouse IgG conjugated with Alexa Fluor 488. For all three pairs, anti-His fluorescence intensities are slightly higher than those of anti-C-myc. The reason might be the unspecific interaction of cellulase fusion with yeast cells.

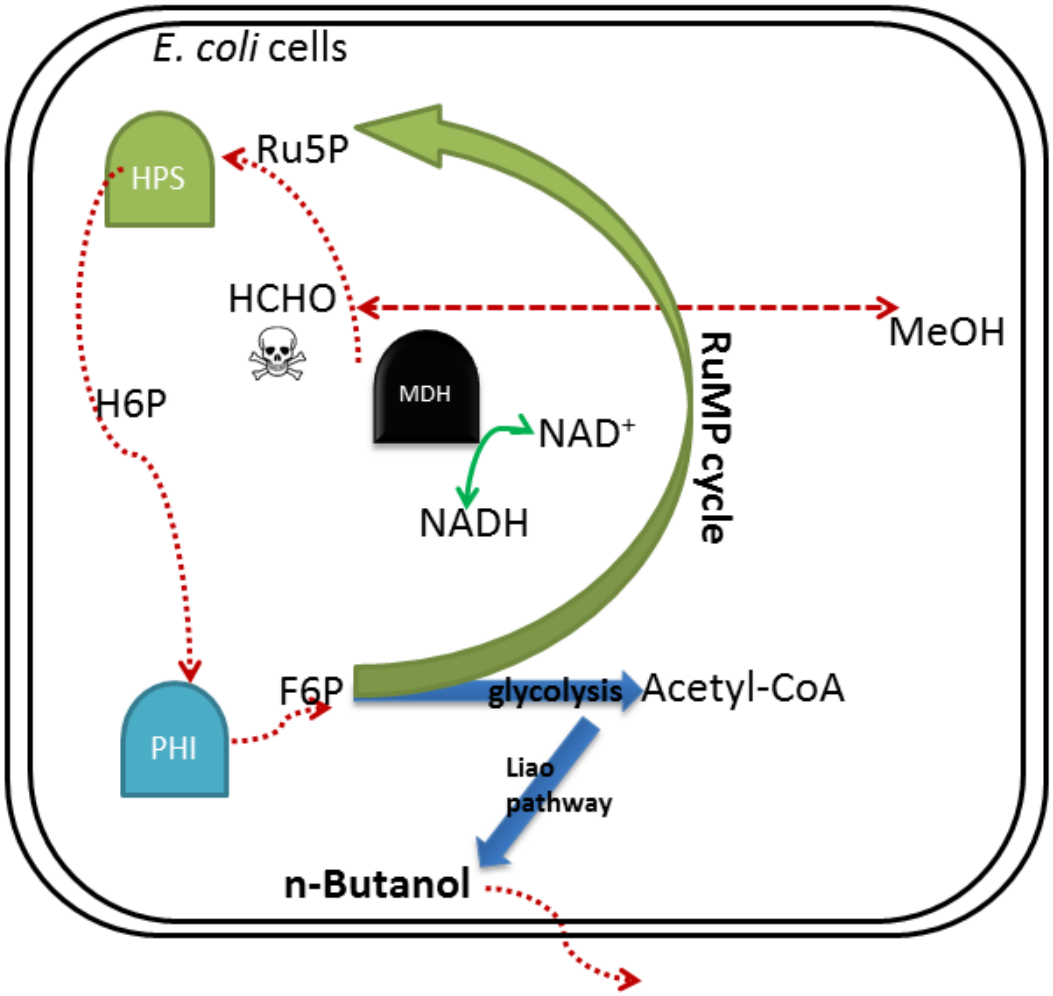


Figure 5. Metabolic pathways of methanol conversion to n-Butanol. A). Free enzyme system. Three enzymes, MDH, HPS and PHI are expressed heterogeneously inside *E. coli* cells to convert methanol into F6P. F6P will be converted for the production of butanol, as reported in Shen et al (2011). Ru5P: Ribulose-5-phosphate; H6P: Hexulose-6-phosphate; F6P: Fructose-6-phosphate; MDH: methanol dehydrogenase; HPS: Hexulose-P-synthase; PHI: Hexulose-P-isomerase

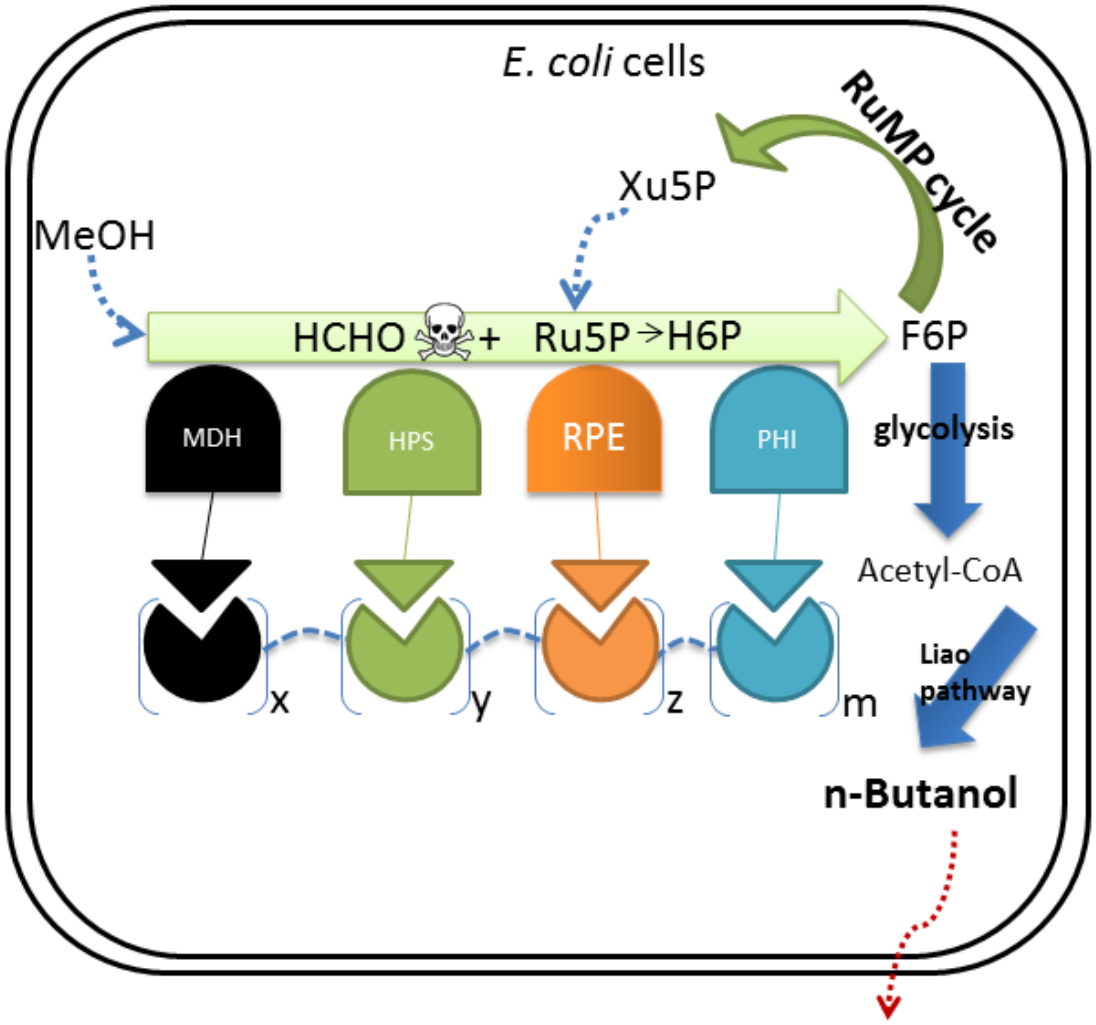


Figure 6. Scaffold-mediated enzyme complex construction. The upstream enzymes involved in the conversion of methanol are colocalized onto a FN3-based protein scaffold. To ensure enough amount of Ru5P to condense formaldehyde into H6P, RPE which converts xylulose-5-phosphate into ribulose-5-phosphate is also recruited. By varying x, y, z and m, we can control the stoichiometry of the four enzymes to achieve maximum metabolic flux. Xu5P: Xylulose-5-phosphate; Ru5P: Ribulose-5-phosphate; H6P: Hexulose-6-phosphate; F6P: Fructose-6-phosphate; MDH: methanol dehydrogenase; HPS: Hexulose-P-synthase; PHI: Hexulose-P-isomerase; RPE: Ribulose-5-Phosphate-3-Epimerase.

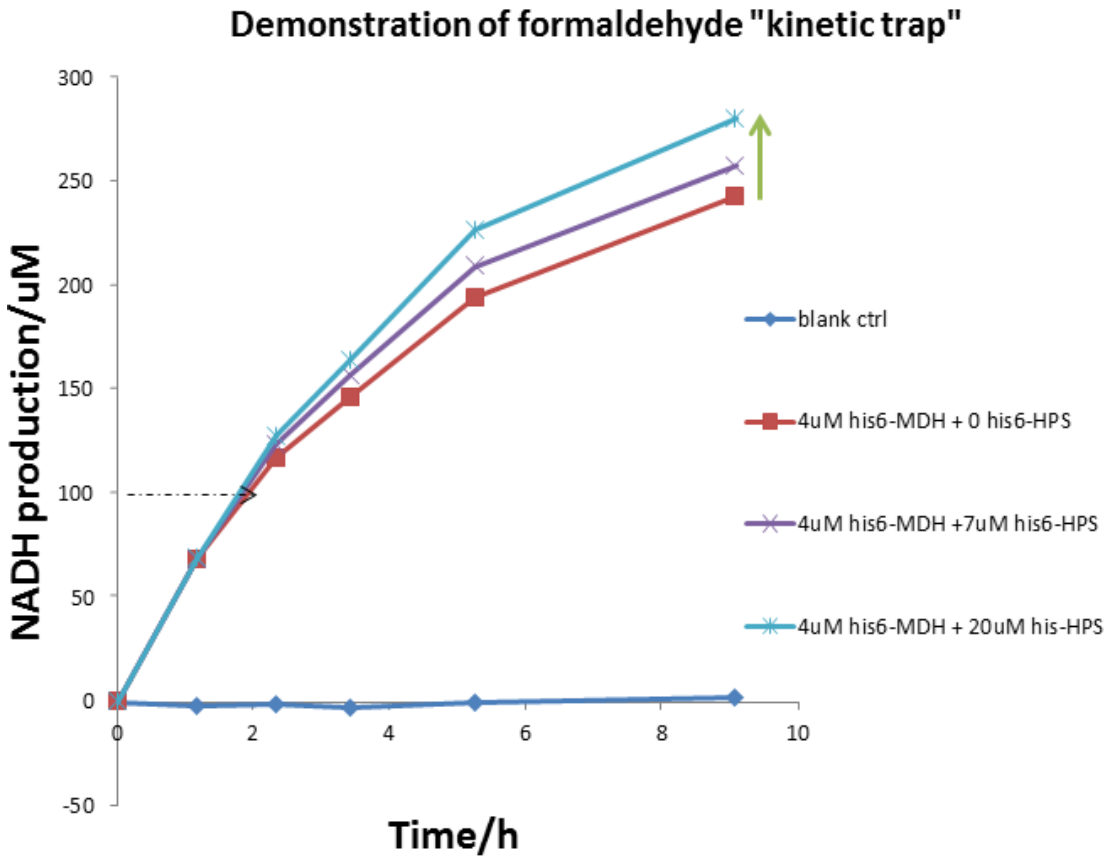


Figure 7. Use varied HPS to MDH concentration ratios to demonstrate the feasibility of formaldehyde "kinetic trap". Assay condition: T=37°C, PH=7.4, 2M methanol, 5mM NAD^+, 5mM ribose-5-phosphate, 6U/ml phosphoriboisomerase.

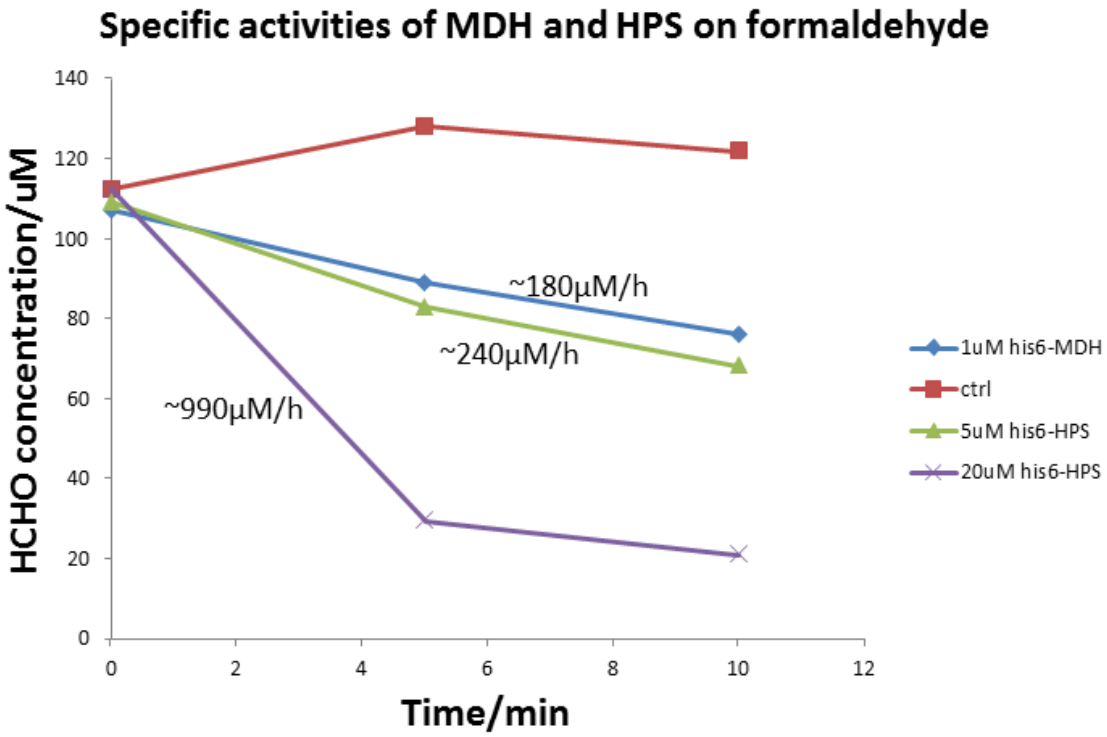


Figure 8. Specific activities of MDH and HPS on formaldehyde. Assay condition: T=37°C, PH=7.4, 0.1mM HCHO, 0.5mM NADH or 5mM ribose-5-phosphate with 6U/ml phosphoriboisomeras.

REFERENCES

Ato, N.K., Urimoto, H.Y. & Hauer, R.K.T., 2006. The Physiological Role of the Ribulose Monophosphate Pathway in Bacteria and Archaea. *Biosci. Biotechnol. Biochem.* 70(1), pp.10–21.

Atsumi, S., Hanai, T. & Liao, J.C., 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, 451(7174), pp.86–9.

Baca, M., Thisted, T., Swers, J. U.S. Patent 130, 324, 2011.

Bogorad, I.W. et al., 2014. Building carbon-carbon bonds using a biocatalytic methanol condensation cycle. *PNAS*, 111(45), pp.15928–33.

Britain, G. & Kemp, M.B., 1974. Hexose Phosphate Synthase from *Methylococcus capsulatus* Makes D-arabino-3-Hexulose Phosphate. *Biochem. J.*, 139, pp.129–134.

Conrado, R.J. & Gonzalez, R., 2014. Envisioning the Bioconversion of Methane to Liquid Fuels. *Science*, 343, pp.621–623.

Conrado, R.J., Varner, J.D. & DeLisa, M.P., 2008. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr. Opin. Biotechnol.*, 19(5), pp.492–9.

Dueber, J.E. et al., 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.*, 27(8), pp.753–759.

Gilbreth, R.N. et al., 2008. A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces. *J. Mol. Biol.*, 381(2), pp.407–18.

Gilbreth, R.N. et al., 2011. Isoform-specific monobody inhibitors of small ubiquitin-related modifiers engineered using structure-guided library design.*PNAS*, 108(19), pp.7751–6.

Good, M.C., Zalatan, J.G. & Lim, W., 2011. Scaffold proteins: hubs for controlling the flow of cellular information. *Science*, 332(6030), pp.680–686.

Haynes, C. a & Gonzalez, R., 2014. Rethinking biological activation of methane and conversion to liquid fuels. *Nat. Chem. Bio.*, 10(5), pp.331–9.

Huang, R., Fang, P. & Kay, B.K., 2012. Isolation of monobodies that bind specifically to the SH3 domain of the Fyn tyrosine protein kinase. *New Biotechnol.*, 29(5), pp.526–33.

Huang, R., Fang, P. & Kay, B.K., 2012. Isolation of monobodies that bind specifically to the SH3 domain of the Fyn tyrosine protein kinase. *New Biotechnol.*, 29(5), pp.526–33.

Nash, T., 1953. The Colorimetric Estimation of Formaldehyde by Means of the Hantzsch Reaction. *Biochem J.*, 55(3), pp.416–421.

Jakobsen, Ø.M. et al., 2006. Upregulated transcription of plasmid and chromosomal ribulose monophosphate pathway genes is critical for methanol assimilation rate and methanol tolerance in the methylotrophic bacterium Bacillus methanolicus. *J. Bacteriol.*, 188(8), pp.3063–3072.

Jiang, H., Wood, K. V & Morgan, J.A., 2005. Metabolic Engineering of the Phenylpropanoid Pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.*, 71(6), pp.2962–2969.

Keltjens, J.T. et al., 2014. PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Appl. Microbiol. Biotechnol.*, 98(14), pp.6163–83.

Koide, A. et al., 2012. Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold. *J. Mol. Biol.*, 415(2), pp.393–405.

Koide, A. et al., 1998. The Fibronectin Type III Domain as a Scaffold for Novel Binding Proteins. *J. Mol. Biol.*, 284, pp.1141–51.

Koide, S., Koide, A. & Lipovšek, D., 2012. Target-binding proteins based on the 10th human fibronectin type III domain (^{10}Fn3). *Methods Enzymol.*, 503, pp.135–56.

Krog, A. et al., 2013. Methylotrophic *Bacillus methanolicus* encodes two chromosomal and one plasmid born NAD+ dependent methanol dehydrogenase paralogs with different catalytic and biochemical properties. *PloS one*, 8(3), p.e59188.

Lynd, L.R. et al., 2008. How biotech can transform biofuels. *Nat. Biotechnol.*, 26(2), pp.169–172.

Lynd, L.R., Wyman, C.E. & Gerngross, T.U., 1999. Biocommodity engineering. *Biotechnol. Progr.*, 15(5), pp.777–793.

Matsushita, K. et al., 1997. *Escherichia coli* is unable to produce pyrroloquinoline quinone (PQQ). *Microbiology*, 143, pp.3149-56.

Mingardon, F. et al., 2007. Exploration of new geometries in cellulosome-like chimeras. *Appl. Environ. Microbiol.*, 73(22), pp.7138–49.

Moon, T.S. et al., 2010. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metab. Eng.*, 12(3), pp.298–305.

Moraïs, S. et al., 2010. Cellulase-xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic substrate. *MBio*, 1(5), pp.3–10.

Müller, J.E.N. et al., 2015. Engineering *Escherichia coli* for methanol conversion. *Metab. Eng.*, 12, pp.1–12.

Office, R.E., Information, T. & Ridge, O.,2005. Biomass as Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of Billion-Ton Annual Supply.

Olah, G., 2013. Towards oil independence through renewable methanol chemistry. *Angew. Chem. Int. Ed.*, 52(1), pp.104–7.

Ovádi, J. & Saks, V., 2004. On the origin of intracellular compartmentation and organized metabolic systems. *Mol. Cell. Biochem.*, 256, pp.5–12.

Ovadit, J. & Srere, P.A., 1996. Metabolic Consequences of Enzyme Interactions. *Cell Biochem. Funct.*, 14, pp.249–258.

Pitera, D.J. et al., 2007. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab. Eng.*, 9(2), pp.193–207.

Shen, C.R. et al., 2011. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.*, 77(9), pp.2905–15.

Stites, T.E., Sih, T.R. & Rucker, R.B., 2000. Synthesis of [^{14}C] pyrroloquinoline quinone (PQQ) in *E . coli* using genes for PQQ synthesis from *K. pneumoniae*. *Biochim. Biophys. Acta.*, 1524, pp.247–252.

Sun, Q. et al., 2014. Creation of artificial cellulosomes on DNA scaffolds by zinc finger protein-guided assembly for efficient cellulose hydrolysis. *Chem. Commun.*, 50(12), pp.1423–5.

Tsai, S., Dasilva, N.A. & Chen, W., 2013. Functional Display of Complex Cellulosomes on the Yeast Surface via Adaptive Assembly. *ACS Synth Biol*., 2, pp.2–9.

Tsai, S.-L. et al., 2009. Functional assembly of minicellulosomes on the *Saccharomyces cerevisiae* cell surface for cellulose hydrolysis and ethanol production. *Appl. Environ. Microbiol.*, 75(19), pp.6087–93.

Tsai, S.-L., Goyal, G. & Chen, W., 2010. Surface display of a functional minicellulosome by intracellular complementation using a synthetic yeast consortium and its application to cellulose hydrolysis and ethanol production. *Appl. Environ. Microbiol.*, 76(22), pp.7514–20.

Whitaker, W.R. & Dueber, J.E., 2011. Metabolic pathway flux enhancement by synthetic protein scaffolding. *Methods Enzymol.*, 497, pp.447-468.

Wojcik, J. et al., 2010. A Potent and Highly Specific FN3 Monobody Inhibitor of the Abl SH2 Domain. *Nat. Struct. Mol. Biol.*, 17(4), pp.519–527.

Xu, P. et al., 2012. ePathBrick: A Synthetic Biology Platform for Engineering Metabolic Pathways in *E. coli*. *ACS Synth Biol*. 1, pp.256-266.