ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

Self-Assembly of Synthetic Metabolons through Synthetic Protein Scaffolds: One-Step Purification, Co-immobilization, and Substrate Channeling

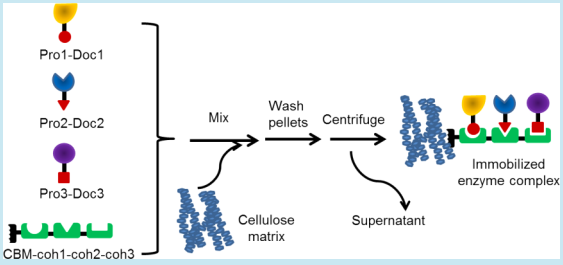
Chun You[†] and Y.-H. Percival Zhang*[,†,‡,§,∥]

[†]Biological Systems Engineering Department and [‡]Institute for Critical Technology and Applied Science (ICTAS), Virginia Tech, Blacksburg, Virginia 24061, United States

[§]DOE BioEnergy Science Center (BESC), Oak Ridge, Tennessee 37831, United States

[∥]Gate Fuels Inc., 2200 Kraft Drive, Suite 1200B, Blacksburg, Virginia 24060, United States

ABSTRACT: One-step purification of a multi-enzyme complex was developed based on a mixture of cell extracts containing three dockerin-containing enzymes and one family 3 cellulose-binding module (CBM3)-containing scaffoldin through high-affinity adsorption on low-cost solid regenerated amorphous cellulose (RAC). The three-enzyme complex, called synthetic metabolon, was self-assembled through the high-affinity interaction between the dockerin in each enzyme and three cohesins in the synthetic scaffoldin. The metabolons were either immobilized on the external surface of RAC or free when the scaffoldin contained an intein between the CBM3 and three cohesins. The immobilized and free metabolons containing triosephosphate isomerase, aldolase, and fructose 1,6-biphosphatase exhibited initial reaction rates 48 and 38 times, respectively, that of the non-complexed three-enzyme mixture at the same enzyme loading. Such reaction rate enhancements indicated strong substrate channeling among synthetic metabolons due to the close spatial organization among cascade enzymes. These results suggested that the construction of synthetic metabolons by using cohesins, dockerins, and cellulose-binding modules from cellulosomes not only decreased protein purification labor and cost for in vitro synthetic biology projects but also accelerated reaction rates by 1 order of magnitude compared to non-complexed enzymes. Synthetic metabolons would be an important biocatalytic module for in vitro and in vivo synthetic biology projects.



KEYWORDS: *biocatalytic module, cascade enzymes, in vitro synthetic biology, substrate channeling, metabolon, synthetic enzyme complex*

The primary goal of synthetic biology is the design and construction of new biological functions and systems that are better than their natural counterparts or of even non-natural ones. These biological systems are usually constructed from parts to modules to pathways to systems.[1] Synthetic biology has two distinctive directions: *in vivo* and *in vitro*.[2−4] Compared to self-duplicating living cellular systems that represent extraordinarily complex and frequently labile systems, in vitro synthetic biology based on well-defined components or cell extracts has received relatively scant attention.[2,3,5−7]

In vitro or cell-free biosystems have been used as fundamental research tools for understanding living systems for more than 100 years.[2,3] In vitro synthetic biology, however, is suggested to be a new biomanufacturing platform due to several advantages suitable in industrial applications.[3,5,6,8,9] First, in vitro biosystems can be easily controlled and accessed without cellular membranes. Cell-free protein synthesis is becoming a new alternative choice for fast protein synthesis.[5] Second, very high product yields are usually accomplished without the formation of byproduct or the synthesis of cell mass. For example, nearly12 H_2 has been produced per glucose unit of polysaccharides and water, three times the theoretical yield of anaerobic hydrogen-producing microorganisms.[10,11] Third, in vitro biosystems can implement some biological reactions that living microbes or chemical catalysts cannot implement before. For example, β-1,4-glucosidic bond linked cellulose can be converted to α-1,4-glucosidic bond linked starch by a mixture of intracellular and extracellular enzymes in one pot (in preparation for publication). Fourth, enzymatic systems without the barrier of cellular membrane usually have faster reaction rates than microbial systems. For instant, enzymatic fuel cells usually have much higher power outputs than microbial fuel cells.[12,13] Fifth, enzyme cocktails are able to tolerate toxic compounds and products better than microorganisms.[14,15] Sixth, enzyme mixtures usually work under broad reaction conditions, such as high temperature, low pH, and the presence of organic solvents or ionic liquids.[8]

Despite the above advantages, in vitro synthetic biology often requires intensive efforts in the preparation of numerous purified enzymes as building blocks, especially when most building blocks are not commercially available. Inspired by

Received: July 28, 2012
Published: August 28, 2012

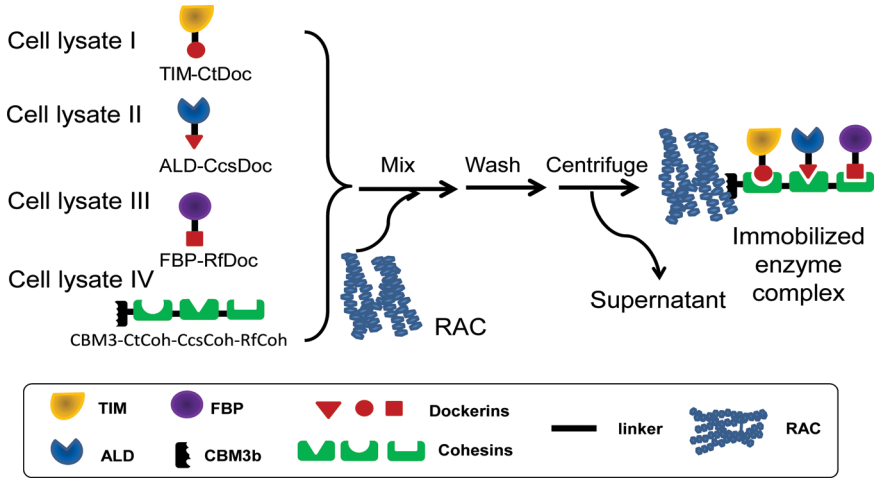


Figure 1. Schematic representation for one-step purification and co-immobilization of the synthetic three-enzyme complex, where the scaffoldin contained three different types of cohesins and one family 3 carbohydrate-binding module, and three enzymes contained respective dockerins.

natural metabolons, enzyme complexes containing cascade enzymes in metabolic pathway,[16−18] the construction of in vitro synthetic metabolons is appealing because it not only significantly decreases the protein purification cost and labor but also may increase reaction rates among metabolons containing cascade enzymes due to the substrate channeling phenomena and avoid the degradation of labile metabolites.[16,19,20]

Cellulosomes are extracellular enzyme complexes, consisting of catalytic subunits such as glycoside hydrolases (mainly), polysaccharide lyases, and carboxyl esterases bound together by non-catalytic scaffoldins.[21−23] The hydrolytic enzyme complexes are formed through the high-affinity interaction between the cohesins of scaffoldins and the dockerins of hydrolytic enzymes. Such cellulosomes can hydrolyze cellulose faster than their non-complexed cellulase mixture, especially on low-accessibility recalcitrant cellulose.[24−26] In 1994, Bayer and his co-workers[23] first proposed the construction of synthetic enzyme complexes by utilizing dockerins, cohesins, and cellulose-binding modules from cellulosomal systems. Later, numerous synthetic (mini-)cellulosomes have been constructed for facilitating enzymatic hydrolysis rates on solid polysaccharides, such as cellulose and pretreated lignocellulosic biomass.[24−27]

In this study, we developed simple approaches to purify and/or immobilize three-enzyme complexes as building modules for in vitro synthetic biology projects by using a low-cost biodegradable cellulosic material. Triosephosphate isomerase (TIM, EC 5.3.1.1), aldolase (ALD, EC 4.1.2.13), and fructose 1,6-bisphosphatase (FBP, EC3.1.3.11) in the glycolysis and gluconeogenesis pathways were chosen for demonstration purposes. These three cascade enzymes were assembled and purified easily.

RESULTS AND DISCUSSION

One-Step Metabolon Purification and Immobilization.

Thermus thermophilus TIM, *Thermotoga maritima* ALD and FBP were used for the study of in vitro assembly of the synthetic metabolons as building modules. TIM is responsible for reversibly converting glycerealdehyde-3-phosphate (G3P) to dihydroxyacetone phosphate (DHAP) (eq 1); ALD catalyzes the reversible aldol condensation of G3P and DHAP to fructose 1,6-bisphosphate (F16P) (eq 2); and FBP catalyzes the irreversible conversion of F16P to F6P (eq 3):

$$\mathrm{G3P} \overset{\mathrm{TIM}}{\longleftrightarrow} \mathrm{DHAP} \tag{1}$$

$$\mathrm{G3P} + \mathrm{DHAP} \overset{\mathrm{ALD}}{\longleftrightarrow} \mathrm{F16P} \tag{2}$$

$$\mathrm{F16P} \overset{\mathrm{FBP}}{\longrightarrow} \mathrm{F6P} + \mathrm{P_i} \tag{3}$$

These three cascade enzymes are the key enzymes in the glycolysis and gluconeogenesis pathway, which has been used to produce high-yield hydrogen from sugars and water.[10,11] Previous studies reported that a dynamic metabolon between ALD and FBP in muscle was formed,[28] and that bifunctional ALD/FBP enzymes in Archaeal microorganisms ensured a fast conversion of labile metabolite F16P to stable F6P,[29] suggesting the feasibility of substrate channeling occurring among these three enzymes. However, we found that the purified TIM, ALD, and FBP did not form the transient complex examined by affinity electrophoresis (data not shown).

To construct the static three-enzyme metabolon through high-affinity and high-specific interaction between cohesins and dockerins, the synthetic mini-scaffoldin CBM3-CtCoh-CcsCoh-RfCoh (i.e., CBM3-Scaf3) was constructed to contain a family 3 cellulose-binding module (CBM3) at the N-terminus followed by three different type cohesins from the *Clostridium thermocellum* ATCC 27405 CipA, *Clostridium cellulovorans* ATCC 35296 CbpA, and *Ruminococcus flavefaciens* ScaB.[30] The dockerin-containing TIM, ALD, and FBP were constructed by the addition of one dockerin from *C. thermocellum* CelS, *C. cellulovorans* EngE, and *R. flavefaciens* ScaA at their C-terminus, respectively. Four *E. coli* BL21 strains harboring the expression plasmids expressed soluble recombinant proteins. The four cell extracts were mixed together, and the high-specificity interaction between cohesins and dockerins allowed each pair of cohesin and dockerin to bind tightly at the molar ratio of 1:1,[31] forming the synthetic three-enzyme complex, called a synthetic metabolon (Figure 1). In the presence of solid regenerated amorphous cellulose (RAC), such synthetic three-enzyme complex was adsorbed on the surface of cellulosic material through the CBM3 in the scaffoldin. After washing and centrifugation, the synthetic metabolon containing TIM, ALD, and FBP was purified and immobilized on RAC (Figure.1).

dx.doi.org/10.1021/sb300068g | *ACS Synth. Biol.* 2013, 2, 102−110

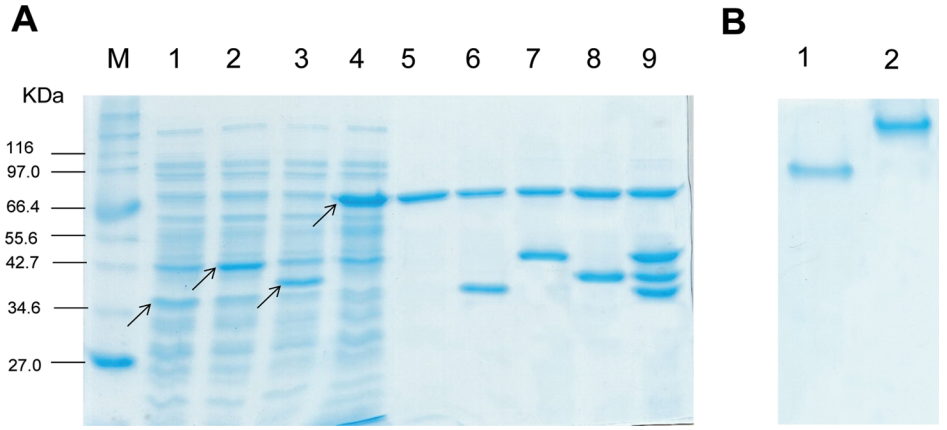


Figure 2. SDS-PAGE analysis of the *E. coli* cell extracts containing the recombinant proteins and RAC pull-down proteins (A). Lane M, protein marker; Lanes 1−4, cell extract containing TIM-CtDoc, ALD-CcsDoc, FBP-RfDoc, and CBM3-Scaf3 (i.e., CBM3-CtCoh-CcsCoh-RfCoh), respectively; Lane 5, RAC adsorbed CBM3-Scaf3, Lanes 6−8, RAC adsorbed CBM3-Scaf3 and TIM-CtDoc, ALD-CcsDoc, and FBP-RfDoc, respectively; and Lane 9, RAC adsorbed CBM3-Scaf3, TIM-CtDoc, ALD-CcsDoC, and FBP-RfDoc. Native gel (B) of the EG-eluted CBM3-Scaf3 only (Lane 1) and the EG-eluted synthetic metabolon (Lane 2).

SDS-PAGE analysis was conducted to study protein expression, purification and co-immobilization (Figure 2A). Four *E. coli* cell extracts containing TIM-CtDoc, ALD-CcsDoc, FBP-RfDoc, and CBM3-Scaf3 are shown in Lanes 1−4 of Figure 2A, respectively. The scaffoldin can be purified by RAC pull-down experiment (Lane 5, Figure 2A). When RAC was mixed with the cell extract containing the scaffoldin and another cell extract containing TIM-CtDoc, ALD-CcsDoc, or FBP-RfDoc, the high-affinity interaction between the cohesin and dockerin resulted in the formation of the unifunctional enzyme metabolon containing one scaffoldin and one dockerin-containing enzyme, exhibiting two bands in SDS-PAGE analysis (Lane 6−8). When RAC was mixed with the four cell extracts, the adsorbed three-enzyme metabolon on RAC had four bands in SDS-PAGE (Lane 9), suggesting that one scaffoldin can bind one TIM, one ALD, and one FBP. The immobilized metabolon can be eluted by ethylene glycol (EG); the eluted and dialyzed metabolon was validated in a native gel (Figure 2B, lane 2), compared to the scaffoldin only (lane 1).

The three-enzyme synthetic metabolon was suggested to be adsorbed on the surface of RAC, which has much larger external surface area than microcrystalline cellulose (Avicel).[32] The adsorption curves of the EG-eluted synthetic metabolon and scaffoldin obeyed the Langmuir isotherms well, suggesting a single layer adsorption (Figure 3A). The maximum adsorption capacity of the scaffoldin was 208 mg/g RAC (Figure 3A). Through its three cohesins, TIM, ALD, and FBP were bound with the scaffoldin, resulting in the enhanced overall protein adsorption capacity to 442 mg/g RAC. For low-accessibility Avicel, the maximum adsorption capacity of the scaffoldin was 8.3 mg/g Avicel (Figure 3B), lower than the value of green fluorescent protein-CBM3 (i.e., 14.8 mg/g Avicel),[32] possibly because the scaffoldin was a large size protein that cannot be adsorbed by small pores inside Avicel.[33] The maximum adsorption capacity of the synthetic metabolon was 16.5 mg/g Avicel, 1.99 times of that of mini-scaffoldin on Avicel. The maximum adsorption capacity ratio of the synthetic metabolon to the scaffoldin on RAC was 2.13, slightly higher than the value on Avicel (1.99), possibly because some enzymes cannot bind with the scaffoldin that was adsorbed by Avicel. Therefore, it was strongly recommended that the large-size synthetic metabolons should be immobilized on the solid

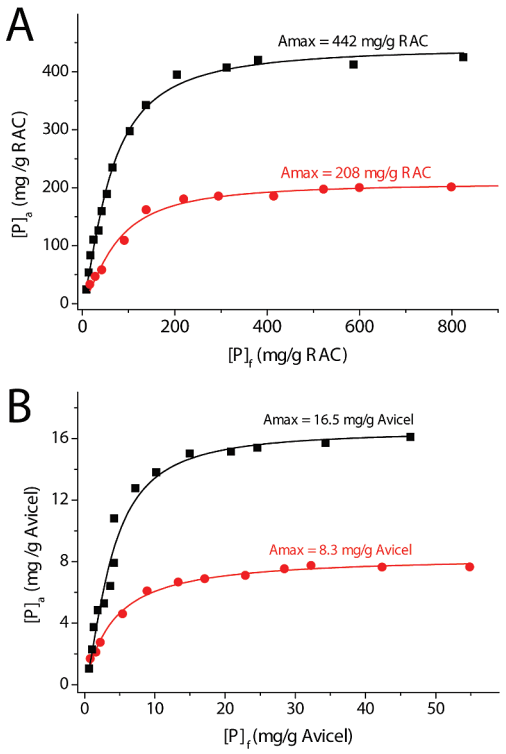


Figure 3. Isotherm adsorption of the CBM3-containing synthetic metabolon (■) and the CBM3-containing mini-scaffoldin (●) on RAC (A) and Avicel (B). The curves were fitted by the Langmuir equations.

materials with large external surface (e.g., RAC) or large-pore nanomaterials.

Large-accessibility RAC is a low-cost biodegradable cellulosic material prepared from commercial microcrystalline cellulose (Avicel). The CBM3 in scaffoldins may be replaced with other tags, such as glutathione *S*-transferase (GST) and maltose-binding protein (MBP). However, we did not recommend using the His-tag because Ni^{2+} ions on the Ni-NTA resin could affect the interaction between cohesins and dockerins (data not shown), where the presence of Ca^{2+} was vital to the formation of enzyme complexes.[23,34] One-step purification of multiple

dx.doi.org/10.1021/sb300068g | *ACS Synth. Biol.* 2013, 2, 102−110

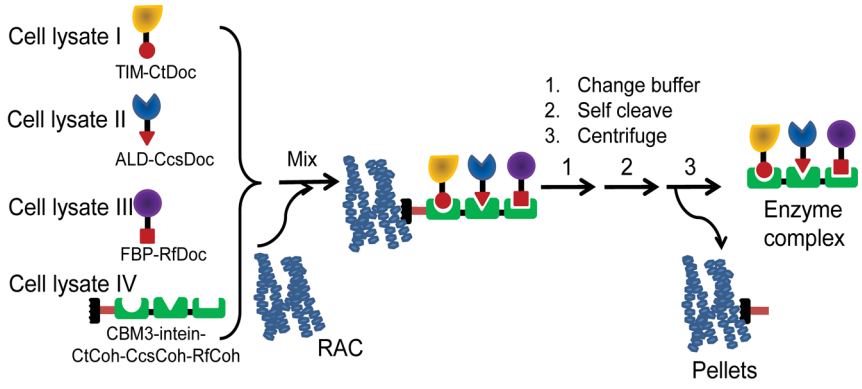


Figure 4. Schematic representation for the purification of the CBM3-free synthetic metabolon through RAC adsorption followed by intein self-cleavage.

enzymes by using such low-cost cellulosic materials would greatly decrease enzyme preparation labor for in vitro synthetic biology projects[3] and decrease enzyme purification costs in biocatalysis.[35]

Bayer and his co-workers developed a protein purification method that can specifically capture one low concentration dockerin-containing enzyme by using one-cohesin-containing protein immobilized on a cellulose-based affinity column.[31] Here we captured three enzymes as the complex and co-immobilized it on cellulose by the CBM3-tagged scaffoldin. These simple protein purification technologies have several advantages: (i) relatively high protein purity based on the SDS-PAGE analysis (Figure 2A), although target protein concentration in cell extract was pretty low; (ii) a scalable technology based on simple adsorption and centrifugation or filtration without the use of chromatographic columns; (iii) ultra-low-cost adsorbent RAC, which is biodegradable and renewable. It was worth noting that the cohesin number in synthetic scaffoldins could be increased to more than three[36] so that more metabolic enzymes could be purified through large-size scaffoldins in one step. Also, the number of cohesins and their orders in scaffoldin could be optimized, as occurred in the in vivo synthetic biology study[19] and synthetic cellulosomes.[27]

Purification of CBM3-Free Synthetic Metabolon. Although the CBM3-tagged synthetic metabolon immobilized on RAC can be used for advanced biocatalysis, solid RAC not only might influence enzymatic assays but also could decrease effective reactor volume.[37] The immobilized synthetic metabolon can be eluted by EG, but the elution efficiency was pretty low and the presence of EG may impair activities of some enzymes (e.g., potato α-glucan phosphorylase, unpublished). Therefore, we designed another method to purify a CBM3-free synthetic metabolon through RAC adsorption followed by intein self-cleavage (Figure 4). The intein module was inserted between CBM3 and the first cohesin in the scaffoldin. Four *E. coli* cell extracts contained TIM-CtDoc (Figure 5, Lane 1), ALD-CcsDoc (Lane 2), FBP-RfDoc (Lane 3), and CBM3-intein-CtCoh-CcsCoh-RfCoh (i.e., CBM3-intein-Scaf3) (Lane 4) as major products. The CBM3-intein-Scaf3 was adsorbed on RAC but not for other *E. coli* cellular proteins (lane 5). After the buffer was changed to the relatively low pH and high salt strength buffer, this scaffoldin was self-cleaved into Scaf3 (Lane 6) in the supernatant and adsorbed CBM3-intein in the RAC pellets. When four cell extracts and the RAC slurry were mixed together followed by washing and centrifugation, RAC can pull down the synthetic metabolon, having four bands in SDS-PAGE (Lane 7). After the buffer exchange, self-cleavage, and centrifugation, the CBM3-free synthetic metabolon was obtained in the supernatant (Lane 8), showing four bands including the truncated scaffoldin and the other three enzymes.

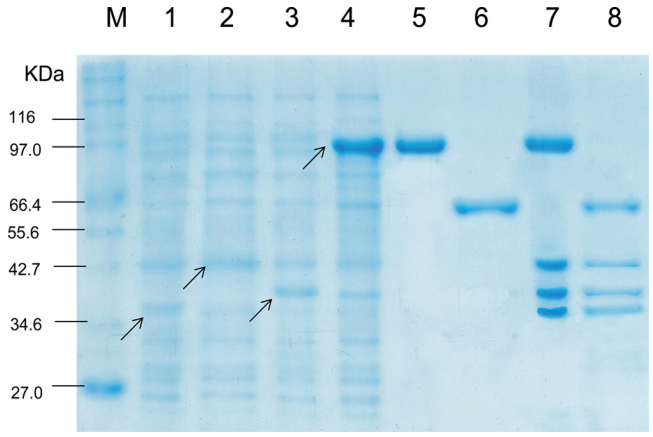


Figure 5. SDS-PAGE analysis of the *E. coli* cell extracts containing the recombinant proteins and cleaved non-immobilized enzyme complex. (A). Lane M, protein marker; Lanes 1–4, cell extract containing TIM-CtDoc, ALD-CcsDoc, FBP-RfDoc, and CBM3-intein-Scaf3 (i.e., CBM3-intein-CtCoh-CcsCoh-RfCoh), respectively; Lane 5, RAC adsorbed CBM3-intein-Scaf3; Lane 6, cleaved Scaf3; Lane 7, RAC adsorbed CBM3-intein-Scaf3, TIM-CtDoc, ALD-CcsDoC, and FBP-RfDoc; Lane 8, free synthetic metabolon containing cleaved Scaf3, TIM-CtDoc, ALD-CcsDoC, and FBP-RfDoc.

Substrate Channeling in Synthetic Metabolons. The F6P formation profiles were examined on 2.5 mM G3P mediated by a non-complexed three-enzyme mixture (2 μM each enzyme), 2 μM concentration of the RAC-immobilized synthetic metabolon, and 2 μM concentration of the CBM3-free synthetic metabolon at 37 °C (Figure 6). The non-complexed three-enzyme mixtures exhibited a very low activity, 0.28 μM F6P/min at 37 °C, because thermophilic enzymes exhibited low activities at decreased temperature. When three enzymes were assembled tightly through scaffoldins either on the surface of RAC or free in the aqueous solution, synthetic metabolons exhibited drastic increases in reaction rates. The immobilized synthetic metabolon and the CBM3-free synthetic metabolon had initial reaction rates of 13.6 and 10.7 μM/min, respectively. Such synthetic metabolons exhibited rates 38–48 times that of the non-complexed enzyme mixture at the same enzyme loading, exhibiting a high degree of substrate channeling, because the product of one enzyme can be

dx.doi.org/10.1021/sb300068g | *ACS Synth. Biol.* 2013, 2, 102–110

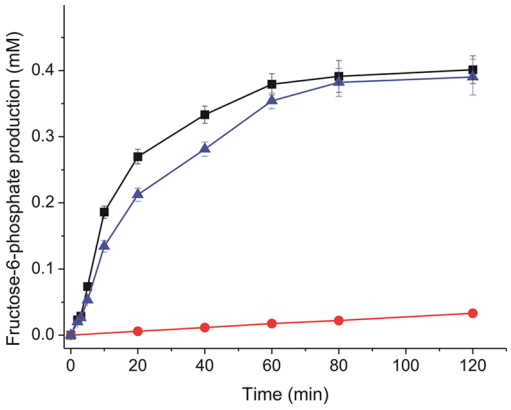


Figure 6. Degree of substrate channeling of the immobilized (■) and non-immobilized (▲) metabolons compared to the non-complexed three enzyme mixture (●) at the same enzyme loading (2 μM). The degree of substrate channeling was defined as the ratio of the initial reaction rate of the synthetic metabolon to the non-complexed enzyme mixture at the same molar enzyme loading.

transferred to an adjacent enzyme where it acts as the substrate without full equilibrium in the bulk solution. The immobilized synthetic metabolon exhibited approximately 27% higher reaction rates than the synthetic metabolon, which was possibly due to shorter enzyme−enzyme distances when the enzymes were co-immobilized on the surface of the solid adsorbent compared to those in the aqueous solution.

Only dimeric TIM is active;[38] all monomeric, dimeric and tetrameric ALDs are active;[39] and both dimeric and tetrameric FBPs are active.[40] Since the addition of dockerins on the C terminus of these three enzymes did not influence their activities, the three-enzyme complex may form a dimer with TIM, ALD, and FBP facing one another.[30] The formation of this putative dimeric three-enzyme complex was supported by the fact that the particle size of the synthetic metabolon was approximately 5-fold larger than the single mini-scaffoldin.[30] The formation of synthetic metabolons by using cohesins and dockerins could be a good choice for multimeric enzymes.

The in vitro experimental data clearly suggested that static synthetic metabolons containing TIM, ALD, and FBP exhibited tens-fold increases in reaction rates compared to that of the non-complexed enzyme mixture, implying that the association or dissociation of transient metabolons may be a very important mechanism in living organisms for the forestallment of substrate competition among different pathways, regulation of metabolic fluxes in response to environmental changes, mitigation of metabolite inhibition, and circumvention of unfavorable equilibrium and kinetics.[16,41] Since the immobilized synthetic metabolon exhibited approximately 27% higher reaction rates compared with those of the synthetic metabolon, it was speculated that the metabolons occurring on the cellular membrane might work more effectively than those in cytoplasm.

In conclusion, the three-enzyme synthetic metabolons were immobilized on the surface of RAC (Figure 1) or free (Figure 4), depending on the choice of scaffoldins and the potential applications of such synthetic metabolons. One-step purification of multiple enzymes by using low-cost cellulosic materials would greatly decrease enzyme preparation labor and cost for in vitro synthetic biology projects[3] and multi-enzyme biocatalysis.[35] Additionally, such enzyme complexes facilitated the substrate transfer among adjacent cascade enzymes, resulting in drastic increases in reaction rates (Figure 6). As a result, the use of synthetic metabolons featuring substrate

Table 1. Primers Used in This Study

primer	sequence
VF-Scaf	TACTGATGGCGCTATCATTGTTAAGCTCGAGCACCACCACCACCACCACT
VR-Scaf	CCTTCAAATTGCCTGATACCGGCATATGTATATCTCCTTCTTAAAGTTAA
IF- Scaf	TTAACTTTAAGAAGGAGATATACATATGCCGGTATCAGGCAATTTGAAGG
IR-Scaf	AGTGGTGGTGGTGGTGGTGCTCGAGCTTAACAATGATAGCGCCATCAGTA
F-intein	GCGGCCGCGAATTCCTCGAGGGTGGCAGTGTAGTACCATC
R-intein	GATGGTACTACACTGCCACCCTCGAGGAATTCGCGGCCGC
VF-TIM	AATAGATACATTGCCGTACAAGAACTAGGGATCCGAATTCGAGCTCCGTC
VR-TIM	AGTTGCCCGCCACCAAGACGCGCCGCATATGTATATCTCCTTCTTAAAGT
IF-TIM	ACTTTAAGAAGGAGATATACATATGCGGCGCGTCTTGGTGGCGGGCAACT
IR-TIM	AAGCAGTTGCCTTGTTCCAATCAACACCCGCGATCCTGAGGAGGGCGAGG
IF-CTDoc	CCTCGCCCTCCTCAGGATCGCGGGTGTTGATTGGAACAAGGCAACTGCTT
IR-CTDoc	GACGGAGCTCGAATTCGGATCCCTAGTTCTTGTACGGCAATGTATCTATT
TIM-XhoI-R	GTCCACTCGAGACCCGCGATCCTGAGGAGGGCGAGG
VF-ALD	TCTTGCATTCTTAAAAAAAGCAATATGAGTCGACAAGCTTGCGGCCGCAC
VR-ALD	TCTTTACATAAGGCATTGTCACCATATGTATATCTCCTTCTTAAAGTTAA
IF-ALD	TTAACTTTAAGAAGGAGATATACATATGGTGACAATGCCTTATGTAAAGA
IR-ALD	CTTGGATTCCTGGTAATACCTTACCCGCCTTTCCGCTGGATCCGAATATT
IF-CCDoc	AATATTCGGATCCAGCGGAAAGGCGGGTAAGGTATTACCAGGAATCCAAG
IR-CCDoc	GTGCGGCCGCAAGCTTGTCGACTCATATTGCTTTTTTTAAGAATGCAAGA
VF-FBP	CGTTGAACTCATCACACTTCCTCAGCTCGAGCACCACCACCACCACCACT
VR-FBP	TAGAAAAGTCCAGTCTGTCCAACATATGTATATCTCCTTCTTAAAGTTAA
IF-FBP	TTAACTTTAAGAAGGAGATATACATATGTTGGACAGACTGGACTTTTCTA
IR-FBP	GGAACGAGCTTTGTGCCGGGCTTTCCTCCTATTTCTTCTA
IF-RFDoc	TAGAAGAAATAGGAGGAAAGCCCGGCACAAAGCTCGTTCC
IR-RFDoc	AGTGGTGGTGGTGGTGGTGCTCGAGCTGAGGAAGTGTGATGAGTTCAACG

dx.doi.org/10.1021/sb300068g | ACS Synth. Biol. 2013, 2, 102−110

Table 2. Strains and Plasmids

stains or plasmids	characteristics	ref
	E. coli	
Top10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-*	Invitrogen
BL21 Star (DE3)	F^-*ompT hsd*SB (rB^-mB^-) *gal dcm rne131* (DE3)	Invitrogen
	Plasmids	
pET20b-cbm3-scaf3	Amp^R, mini-scaffoldin expression cassette containing a CBM3 module from *C. thermocellum* CipA and three different cohesins from *C. thermocellum*, *C. cellulovorans* and *R. flavefaciens*	this work
pET20b-cbm3-intein-scaf3	Amp^R, mini-scaffoldin expression cassette containing a CBM3 module from *C. thermocellum*, an intein module from *Synechocystis sp* DnaB and three different cohesins from *C. thermocellum*, *C. cellulovorans* and *R. flavefaciens*	this work
pET20b-tim-ctdoc	Amp^R, tim-ctdoc expression cassette containing TIM module (TTC0581) from T. thermophiles and the dockerin module from *C. thermocellum*	this work
pET20b-ald-ccsdoc	Amp^R, ald-ccdoc expression cassette containing ALD module (TM0273) from *T. maritime* and the dockerin module from *C. cellulovorans*	this work
pET20b-fbp-rfdoc	Amp^R, fbp-rfdoc expression cassette containing FBP module (TM1415) from *T. maritime* and the dockerin module from *R. flavefaciens*	this work
pET20a-tim	Amp^R, tim expression cassette containing TIM protein, which was purified based on the C-terminal 6× His tag	this work
pET28a-ald	$Kana^R$, ald expression cassette containing ALD protein, which was purified based on the C-terminal 6× His tag	gifted by J. J. Zhong
pCIF	Amp^R, cbm3-intein-fbp expression cassette containing FBP protein without any tag, which can be obtained by intein self-cleavage	20

channeling for in vitro synthetic biology projects could increase overall reaction rates or decrease enzyme use.[42]

METHODS

Chemicals. All chemicals were reagent grade or higher, purchased from Sigma (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), unless otherwise noted. Microcrystalline cellulose Avicel PH105 (20 μm) was purchased from FMC (Philadelphia, PA). Large accessibility regenerated amorphous cellulose (RAC) was prepared from Avicel through cellulose dissolution and precipitation as previously described.[32] The PCR enzyme was Phusion DNA polymerase from New England Biolabs (Ipswich, MA). The oligonucleotides were synthesized by Integrated DNA Technologies (Coraville, IA).

Strains and Medium. *Escherichia coli* Top10 was used as a host cell for DNA manipulation, and *E. coli* BL21 Star (DE3) (Invitrogen, Carlsbad, CA) was used as a host cell for recombinant protein expression. Luria−Bertani (LB) medium was used for *E. coli* cell culture and recombinant protein expression. The final concentrations of antibiotics for *E. coli* were 100 mg/L ampicillin or 50 mg/L kanamycin.

Construction of Plasmids. The sequences of all PCR primers used are listed in Table 1. The plasmids are summarized in Table 2. Restriction enzyme-free, ligase-free, and sequence-independent Simple Cloning enabled us to make chimeric plasmids rapidly.[43] All of the plasmid sequences were validated by DNA sequencing.

Plasmid pET20b-cbm3-scaf3 had an expression cassette containing one CBM3 module and one cohesin module from CipA (314−723 amino acids, GenBank Accession number: L08665) of *C. thermocellum* ATCC 27405, one cohesin module from CbpA (1078−1253 amino acids, GenBank Accession number: AAA23218.1) of *C. cellulovorans* ATCC 35296, and one cohesin module of ScaB (28−186 amino acids, GenBank Accession number: CAC34385.1) from *R. flavefaciens* in tandem. The DNA sequence encoding this chimeric mini-scaffoldin was synthesized by Genescript (Piscataway, NJ). The DNA sequence encoding this chimeric mini-scaffoldin was amplified by PCR using a pair of primer (IF-scaf and IR-scaf); the pET20b vector backbone was amplified with a pair of primers (VF-scaf and VR-scaf). The two DNA templates were *in vitro* assembled to DNA multimers by modified overlap extension PCR.[43] Five microliters of the PCR product containing 1 μg DNA multimers was transformed to *E. coli* Top 10, yielding plasmid pET20b-cbm3-scaf3.

Plasmid pET20b-cbm3-intein-scaf3 had an expression cassette containing one CBM3 module from from CipA (314−474 amino acids, GenBank Accession number: L08665) of *C. thermocellum* ATCC 27405, an intein module from *Synechocystis sp* DnaB (2−155 amino acids, GenBank Accession number: ABD24064), one cohesin module from CipA (475−723 amino acids, GenBank Accession number: L08665) of *C. thermocellum* ATCC 27405, one cohesin module from CbpA (1078−1253 amino acids, GenBank Accession number: AAA23218.1) of *C. cellulovorans* ATCC 35296, and one cohesin module of ScaB (28−186 amino acids, GenBank Accession number: CAC34385.1) from *R. flavefaciens* in tandem. The DNA sequence encoding CBM3-intein was amplified by PCR using a pair of primers (IF-scaf and R-intein) from pCIF;[20] the DNA fragment encoding three cohesins was amplified using a pair of primers (F-intein and IR-scaf) from pET20b-cbm3-scaf3; the pET20b vector backbone was amplified with a pair of primers (VF-scaf and VR-scaf). Plasmid pET20b-cbm3-intein-scaf3 was obtained based on three DNA fragments by using Simple Cloning.[43]

Plasmid pET20b-tim-ctdoc contained an expression cassette containing the *tim* gene isomerase (TTC0581) from *T. thermophilus* HB27 and a dockerin module from *C. thermocellum* CelS (673−741 amino acids, GenBank Accession number: L06942). The *tim* gene was amplified from the genomic DNA of *T. thermophilus* by a primer pair of IF-TIM and IR-TIM; the DNA fragment encoding a dockerin domain was amplified from the *C. thermocellum* genomic DNA by a primer pair of IF-CTDoc and IR-CTDoc; pET20b vector backbone was amplified with a primer pair of VF-TIM and VR-TIM. Plasmid pET20b-tim-ctdoc based on three DNA fragments was obtained by using Simple Cloning.

Plasmid pET20b-ald-ccsdoc had an expression cassette containing the *ald* gene (TM0273) from *Thermotoga maritima* and a dockerin module from endoglucanase EngE (943−1030 amino acids, GenBank Accession number: AAD39739.1) of *C. cellulovorans*. The *ald* gene was amplified with a primer pair of

IF-ALD/IR-ALD based on the genomic DNA of *T. maritima* MSB8; the DNA fragment encoding a dockerin domain was amplified by a primer pair of IF-CCDoc/IR-CCDoc based on the genomic *C. cellulovorans* DNA; the pET20b vector backbone was amplified with a primer pair of VF-ALD and VR-ALD. pET20b-ald-ccdoc was obtained based on three DNA fragments by Simple Cloning.[43]

Plasmid pET20b-fbp-rfdoc had an expression cassette containing the *fbp* gene (TM1415) from *T. maritima* and a dockerin domain from the *R. flavefaciens* ScaA (787−879 amino acids, GenBank Accession number: CAC34384.3). The DNA encoding fructose-1,6-bisphosphatase was amplified with a primer pair of IF-FBP/IR-FBP based on pCIF;[20] the DNA fragment encoding a dockerin domain was amplified with a primer pair of IF-RFDoc/IR-RFDoc based on the *R. flavefaciens* genomic DNA; pET20b vector backbone was amplified with a primer pair of VF-FBP and VR-FBP. Plasmid pET20b-fbp-rfdoc based on three DNA fragments was obtained by Simple Cloning.[43]

Plasmid pET33b-tim has an expression cassette containing only the *tim* gene. The DNA sequence of TIM was amplified from the genomic DNA of *T. thermophilus* by a primer pair of IF-TIM and IR-TIM. The PCR product was digested with NdeI/XhoI and ligated into NdeI/XhoI-digested pET20b (Novagen, Madison, WI), yielding pET20b-tim. Plasmid pET28a-ald whose expression cassette contains only *ald* gene was kindly provided by Dr. J. J. Zhong of Shanghai Jiao-Tong University (Shanghai, China).

Recombinant Protein Expression and Purification. The strains *E. coli* BL21 Star (DE3) containing the protein expression plasmids were cultivated in the LB medium supplemented with 1.2% glycerol at 37 °C. When A_{600} reached about 0.75, 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG, a final concentration) was added, and the cultivation temperature was decreased to 16 °C for ∼16 h. After centrifugation, the cell pellets were resuspended in a 50 mM HEPES buffer (pH 8.5) containing 1 mM $CaCl_2$ and 50 mM NaCl. The cells were lysed by ultrasonication. Ten microliters of the cell extracts were loaded into 12% SDS-PAGE to check the expression level of the four proteins. Protein purification of His-tag containing protein was conducted routinely as published elsewhere using Ni-NTA resin.[14,30]

One-Step Metabolon Purification and Immobilization. After rough estimation of each targeted protein expression level by SDS-PAGE, 20 mL of the cell lysate supernatant of TIM-CtDoC, 40 mL of the cell lysate supernatant of ALD-CcsDoc, and 10 mL of the cell lysate supernatant of FBP-RfDoc were mixed with 6 mL of the cell lysate supernatant of CBM3-Scaf3, making sure that TIM-CtDoc, ALD-CcsDoc, and FBP-RfDoc were mixed with the scaffoldin at a molar ratio of ca. 1:1. Then 100 mg RAC was used to adsorb CBM3-containing enzyme complex at room temperature for 5 min. After centrifugation at 3000*g* for 10 min, the RAC pellet was washed three times in 20 mL of 100 mM HEPES (pH 7.5) containing 50 mM NaCl and 1 mM $CaCl_2$. After centrifugation at 3000*g* for 10 min, the RAC pellet was obtained as immobilized enzyme complex. Two hundred micrograms of pellet was resuspended in 40 μL of 1X SDS loading buffer. After boiling for 2 min, 10 μL of the supernatant was loaded into 12% SDS-PAGE to check the interaction between mini-scaffoldin and dockerin containing enzymes. EG was used to elute the enzyme complex from RAC.[32] After dialysis in 50 mM HEPES (pH 7.5) containing 1 mM $CaCl_2$, the free enzyme complex was obtained.

Purification of CBM3-Free Metabolon through Intein Self-Cleavage. After rough estimation of each targeted protein expression level by SDS-PAGE, 20 mL of the cell lysate supernatant of TIM-CtDoC, 40 mL of the cell lysate supernatant of ALD-CcsDoc, and 10 mL of the cell lysate supernatant of FBP-RfDoc were mixed with 5 mL of the cell lysate supernatant of CBM3-intein-Scaf3, where TIM-CtDoc, ALD-CcsDoc, and FBP-RfDoc were a little in excess compared to the CBM3-intein-CtCoh-CcsCoh-RfCoh. Then 100 mg of RAC was used to adsorb the CBM3-containing metabolon at room temperature for 5 min. After centrifugation and washing two times in 10 mL of 50 mM HEPES (pH 6.5) containing 500 mM NaCl and 1 mM $CaCl_2$ at 4 °C, the RAC pellet was suspended in 4 mL of 50 mM HEPES (pH 6.5) containing 50 mM NaCl and 1 mM $CaCl_2$. The suspension solution was incubated at 37 °C for 4 h. After centrifugation at 3000*g* for 10 min, the supernatant contained the purified metabolon. Then the buffer of such metabolon was change to 50 mM HEPES (pH 7.5) containing 1 mM $CaCl_2$ by dialysis.

Enzymatic Activity Assays. The enzymatic activity was measured in a 200 mM HEPES buffer (pH 7.5) containing 10 mM $MgCl_2$, 0.5 mM $MnCl_2$, 1 mM $CaCl_2$, and 2.5 mM G3P at 37 °C. In order to determine the substrate channeling, three enzyme systems were tested to determine the enzymatic activity: RAC immobilized metabolon, non-immobilized metabolon, and the free enzyme mixture containing purified dockerin-free TIM, ALD, and FBP. Here, 2 μM of the immobilized, non-immobilized enzyme complex and free enzyme mixture were used for the activity assay. Then purified dockerin-free TIM, ALD, and FBP were equimolar with the counterparts in the enzyme complex. The reaction systems that contained only enzyme or substrate were performed as negative controls. Sixty-five microliters of the reaction sample was withdrawn at indicated time intervals. The reactions were terminated by adding 35 μL of 1.88 M perchloric acid. After centrifugation, the pH of the supernatant was neutralized with 13 μL of 5 M KOH. The production of F6P was measured by using a glucose hexokinase/glucose-6-phosphate dehydrogenase assay kit (Pointe Scientific, Canton, MI) supplemented with the recombination phosphoglucose isomerase.[20,30]

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.[24] The purity of protein samples was examined by 12% SDS-PAGE. The 12% native PAGE and SDS-PAGE were stained by Bio-Rad Bio-Safe Colloidal Coomassie Blue G-250. The intensity of the band in the gel was analyzed with Quantity One (Bio-Rad, Version 4.6.7). The protein adsorption of the metabolon and scaffoldin on cellulosic materials was determined as described previously.[32]

AUTHOR INFORMATION

Corresponding Author

*E-mail: ypzhang@vt.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported partially by the DOE BioEnergy Science Center (BESC), DOE ARPA-E Petro project, and Shell

GameChanger Program. BESC is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This work was also partially supported by the College of Agriculture and Life Sciences Bioprocessing and Biodesign Research Center at Virginia Tech.

REFERENCES

(1) Andrianantoandro, E., Basu, S., Karig, D. K., and Weiss, R. (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol. 2*, 28.
(2) Hodgman, C. E., and Jewett, M. C. (2012) Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng. 14*, 261−269.
(3) Zhang, Y.-H. P. (2010) Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: Challenges and opportunities. *Biotechnol. Bioeng. 105*, 663−677.
(4) Forster, A. C., and Church, G. M. (2007) Synthetic biology projects *in vitro*. *Genome Res. 17*, 1−6.
(5) Swartz, J. R. (2011) Transforming biochemical engineering with cell-free biology. *AIChE J. 58*, 5−13.
(6) Bujara, M., Schümperli, M., Billerbeck, S., Heinemann, M., and Panke, S. (2010) Exploiting cell-free systems: Implementation and debugging of a system of biotransformations. *Biotechnol. Bioeng. 106*, 376−389.
(7) Subsoontorn, P., Kim, J., and Winfree, E. (2012) Ensemble Bayesian Analysis of Bistability in a Synthetic Transcriptional Switch. *ACS Syn. Biol. 1*, 299−316.
(8) Panke, S., Held, M., and Wubbolts, M. (2004) Trends and innovations in industrial biocatalysis for the production of fine chemicals. *Curr. Opin. Biotechnol. 15*, 272−279.
(9) Guterl, J.-K., and Sieber, V. (2012) Biosynthesis "debugged": novel bioproduction strategies, *Eng. Life Sci.* Accepted for publication; DOI: 10.1002/elsc.201100231.
(10) Ye, X., Wang, Y., Hopkins, R. C., Adams, M. W. W., Evans, B. R., Mielenz, J. R., and Zhang, Y.-H. P. (2009) Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. *ChemSusChem 2*, 149−152.
(11) Zhang, Y.-H. P., Evans, B. R., Mielenz, J. R., Hopkins, R. C., and Adams, M. W. W. (2007) High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. *PLoS One 2*, e456.
(12) Cooney, M. J., Svoboda, V., Lau, C., Martin, G., and Minteer, S. D. (2008) Enzyme catalysed biofuel cells. *Energy Environ. Sci. 1*, 320−337.
(13) Zhu, Z. G., Sun, F., Zhang, X., and Zhang, Y.-H. P. (2012) Deep oxidation of glucose in enzymatic fuel cells through a synthetic enzymatic pathway containing a cascade of two thermostable dehydrogenases. *Biosens. Bioelectron. 36*, 110−115.
(14) Wang, Y., Huang, W., Sathitsuksanoh, N., Zhu, Z., and Zhang, Y.-H. P. (2011) Biohydrogenation from biomass sugar mediated by *in vitro* synthetic enzymatic pathways. *Chem. Biol. 18*, 372−380.
(15) Guterl, J.-K., Garbe, D., Carsten, J., Steffler, F., Sommer, B., Reiße, S., Philipp, A., Haack, M., Rühmann, B., Kettling, U., Brück, T., and Sieber, V. (2012) Cell-free metabolic engineering - Production of chemicals via minimized reaction cascades, *ChemSusChem*, DOI: 10.1002/cssc.201200365.
(16) Zhang, Y.-H. P. (2011) Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv. 29*, 715−725.
(17) Srere, P. A. (1985) The metabolon. *Trends Biochem. Sci. 10*, 109−110.
(18) Schoffelen, S., and van Hest, J. C. M. (2012) Multi-enzyme systems: bringing enzymes together in vitro. *Soft Matter 8*, 1736−1746.
(19) Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., Prather, K. L. J., and Keasling, J. D. (2009) Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol. 27*, 753−759.
(20) Myung, S., Wang, Y. R., and Zhang, Y.-H. P. (2010) Fructose-1,6-bisphosphatase from a hyper-thermophilic bacterium *Thermotoga maritima*: Characterization, metabolite stability and its implications. *Proc. Biochem. 45*, 1882−1887.
(21) Bayer, E. A., Belaich, J. P., Shoham, Y., and Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol. 58*, 521−554.
(22) Doi, R. H., and Kosugi, A. (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Microbiol. 2*, 541−551.
(23) Bayer, E. A., Morag, E., and Lamed, R. (1994) The cellulosome—a treasure-trove for biotechnology. *Trends Biotechnol. 12*, 379−386.
(24) You, C., Zhang, X.-Z., Sathitsuksanoh, N., Lynd, L. R., and Zhang, Y.-H. P. (2012) Enhanced microbial cellulose utilization of recalcitrant cellulose by an *ex vivo* cellulosome-microbe complex. *Appl. Environ. Microbiol. 78*, 1437−1444.
(25) Vazana, Y., Morais, S., Barak, Y., Lamed, R., and Bayer, E. A. (2010) Interplay between Clostridium thermocellum family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states. *Appl. Environ. Microbiol. 76*, 3236−3243.
(26) Ding, S.-Y., Xu, Q., Crowley, M., Zeng, Y., Nimlos, M., Lamed, R., Bayer, E. A., and Himmel, M. E. (2008) A biophysical perspective on the cellulosome: new opportunities for biomass conversion. *Curr. Opin. Biotechnol. 19*, 218−227.
(27) Gefen, G., Anbar, M., Morag, E., Lamed, R., and Bayer, E. A. (2012) Enhanced cellulose degradation by targeted integration of a cohesin-fused β-glucosidase into the *Clostridium thermocellum* cellulosome. *Proc. Natl. Acad. Sci. U.S.A. 109*, 10298−10303.
(28) Rakus, D., Pasek, M., Krotkiewski, H., and Dzugaj, A. (2004) Interaction between muscle aldolase and muscle fructose 1,6-bisphosphatase results in the substrate channeling. *Biochemistry 43*, 14948−14957.
(29) Say, R. F., and Fuchs, G. (2010) Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature 464*, 1077−1081.
(30) You, C., Myung, S., and Zhang, Y.-H. P. (2012) Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. *Angew. Chem., Int. Ed. 51*, 8787−8790.
(31) Demishtein, A., Karpol, A., Barak, Y., Lamed, R., and Bayer, E. A. (2010) Characterization of a dockerin-based affinity tag: application for purification of a broad variety of target proteins. *J. Mol. Recognit. 23*, 525−535.
(32) Hong, J., Ye, X, Wang, Y., and Zhang, Y.-H. P. (2008) Bioseparation of recombinant cellulose binding module-protein by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. *Anal. Chim. Acta 621*, 193−199.
(33) Wang, Q. Q., He, Z., Zhu, Z., Zhang, Y.-H. P., Ni, Y., Luo, X. L., and Zhu, J. Y. (2012) Evaluations of cellulose accessibilities of lignocelluloses by solute exclusion and protein adsorption techniques. *Biotechnol. Bioeng. 109*, 381−389.
(34) Johnson, E. A., Sakajoh, M., Halliwell, G., Madia, A., and Demain, A. L. (1982) Saccharification of complex cellulosic substrates by the cellulase system from *Clostridium thermocellum*. *Appl. Environ. Microbiol. 43*, 1125−1132.
(35) Tufvesson, P. r., Lima-Ramos, J., Nordblad, M., and Woodley, J. M. (2011) Guidelines and cost analysis for catalyst production in biocatalytic processes. *Org. Proc. Res. Dev. 15*, 266−274.
(36) Krauss, J., Zverlov, V. V., Schwarz, W. H. (2012) In vitro reconstitution of the complete *Clostridium thermocellum* cellulosome and synergistic activity on crystalline cellulose, *Appl. Environ. Microbiol.* Published ahead of print; DOI: 10.1128/aem.07959-07911.
(37) Cao, L., Langen, L. v., and Sheldon, R. A. (2003) Immobilised enzymes: carrier-bound or carrier-free? *Curr. Opin. Biotechnol. 14*, 387−394.
(38) Maes, D., Zeelen, J. P., Thanki, N., Beaucamp, N., Alvarez, M., Thi, M. H., Backmann, J., Martial, J. A., Wyns, L., Jaenicke, R., and Wierenga, R. K. (1999) The crystal structure of triosephosphate isomerase (TIM) from *Thermotoga maritima*: a comparative thermostability structural analysis of ten different TIM structures. *Proteins 37*, 441−453.

(39) Sherawat, M., Tolan, D. R., and Allen, K. N. (2008) Structure of a rabbit muscle fructose-1,6-bisphosphate aldolase A dimer variant. *Acta Crystallogr., Sect. D 64*, 543−550.
(40) Gerbling, K.-P., Steup, M., and Latzko, E. (1985) Fructose-1,6-bisphosphatase from *Synechococcus leopoliensis*. *Eur. J. Biochem. 147*, 207−215.
(41) Srere, P. A. (1987) Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem. 56*, 89−124.
(42) Zhang, Y.-H. P. (2011) Simpler is better: high-yield and potential low-cost biofuels production through cell-free synthetic pathway biotransformation (SyPaB). *ACS Catal. 1*, 998−1009.
(43) You, C., Zhang, X.-Z., and Zhang, Y.-H. P. (2012) Simple Cloning via direct transformation of PCR product (DNA multimer) to *Escherichia coli* and *Bacillus subtilis*. *Appl. Environ. Microbiol. 78*, 1593−1595.

■ NOTE ADDED AFTER ASAP PUBLICATION

Figure 6 was incorrect in the version published ASAP September 12, 2012, the corect version reposted November 9, 2012.