Metabolic Engineering 49 (2018) 13–20

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

# Artificial Protein Scaffold System (*AProSS*): An efficient method to optimize exogenous metabolic pathways in *Saccharomyces cerevisiae*

Tianyi Li[a,1], Xiuqi Chen[a,d,1], Yizhi Cai[b,c,*], Junbiao Dai[a,b,*]

[a] Key Laboratory of Industrial Biocatalysis (Ministry of Education) and Center for Synthetic and Systems Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China
[b] Shenzhen Key Laboratory of Synthetic Genomics and Center for Synthetic Genomics, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[c] Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, M1 7DN, Manchester, UK
[d] Department of Biology, Johns Hopkins University, 3400N. Charles Street, Baltimore, MD, USA

ARTICLE INFO

*Keywords:*
Synthetic biology
Metabolic engineering
Artificial scaffold protein
Protein-protein interaction domain
Golden gate assembly

ABSTRACT

Scaffold proteins influence cellular signaling by orchestrating multiple enzymes, receptors or ion channels, and could be tailored to enhance the efficiency of biochemical reactions by positioning related enzymes physically together. However, the number of applicable domains remains small, and the construction of scaffold proteins with optimal domain ratio could be tedious and time-consuming. In this study, we outlined a modular design to quickly assemble scaffold proteins using protein interaction domains, which have been constructed into a standardized vector. We generated multiple protein interaction domains and ligands for making artificial scaffold proteins. At the same time, we developed a robust Golden-Gate-based molecular toolkit for the construction of artificial scaffold proteins, allowing a variance of domain types, number, and positions. The synthesized domain-ligand interaction was verified by yeast two-hybrid and split-GFP assays. Using synthetic scaffolds, we demonstrated an increase in the yield of two target products by 29% and 63% respectively. Moreover, we demonstrated that the synthetic scaffold could be applied to rewire the metabolic flux. Our system could be a useful tool for metabolic engineering and beyond.

## 1. Introduction

Microbial production of natural products has been achieved by introducing product-specific enzymes or the entire metabolic pathway into readily engineered organisms such as *E. coli* and budding yeast, a process now we call metabolic engineering (Bailey, 1991). The range of chemicals that can be produced from metabolic engineering has expanded substantially (Alonso-gutierrez et al., 2013; Jin and Alper, 2005; Chemler et al., 2010; Li, Tan, 2017; Cone et al., 2003; Meng et al., 2017; Awan et al., 2017) since the term was coined around 20 years ago, in part due to notable advances in fields related to metabolic engineering. However, introducing a metabolic pathway into a heterologous host can cause problems. The synthetic metabolic pathways lack endogenous regulatory components. Therefore, the introduction of such pathways often leads to growth retardation and causes metabolic imbalance due to the accumulation of over-expressed proteins, end products, and intermediates which are often toxic (Lee et al., 2012 ; Nielsen, Nielsen, 2017; Tan et al., 2016). Moreover, imbalances in a metabolic pathway often elicit a stress response in central metabolism (Keasling, 2010). Currently, tools to alter the regulatory landscape within the cells are neither available nor well-understood (Yadav et al., 2012). This leaves pathway balancing as one of the most challenging issues in metabolic engineering that require further investigation.

The principal objective of balancing a metabolic pathway is to produce more of a target product by reducing potential flux imbalances and cellular burden in the host organism. This is mainly accomplished by eliminating the production of excessive intermediate metabolites and precursors which results in efficient conversion of intermediates, substrates, and co-factors to desired products (Jones et al., 2015). Among the several successful pathway-balancing approaches, post-translational modulation takes advantage of synthetic scaffolds to recruit metabolic enzymes of interest pathways to form synthetic complexes and increase spatial orientation of substrates. This is achieved

* Corresponding authors at: Shenzhen Key Laboratory of Synthetic Genomics and Center for Synthetic Genomics, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China.
*E-mail addresses:* yizhi.cai@manchester.ac.uk (Y. Cai), junbiao.dai@siat.ac.cn (J. Dai).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ymben.2018.07.006
Received 1 May 2018; Received in revised form 11 July 2018; Accepted 12 July 2018

either by protein fusions for enzyme cascades or synthetic scaffolds to dock the enzymes in close proximity using DNA, RNA, or proteins (Sachdeva et al., 2014; Delebecque et al., 2011; Dueber et al., 2009). Dueber et al. (2009) demonstrated that by introducing synthetic scaffold proteins, the production of an end product, mevalonate could be improved significantly. They assembled a synthetic scaffold with three protein domains, GBD from rat N-WASP, SH3 from mouse Crk, and PDZ from mouse α-syntrophin. Meanwhile, they fused the corresponding ligands on the three enzymes in the mevalonate pathway, namely AtoB, HMGS, and HMGR. By varying the number of each domain on the scaffold, they successfully found an optimal combination and improve the mevalonate production by 77 folds.

Although introducing the synthetic scaffold proteins can be a very tempting approach for metabolic engineers to optimize their pathways, two major problems lie ahead. The construction of a synthetic scaffold protein with optimal configuration can be very time-consuming and laborious. As indicated in the previous study, constructing the fusion proteins with multiple domains requires a significant amount of repetitive lab work. One has to configure the scaffold protein specifically for each metabolic pathway, which would be essential for the performance of the scaffolding strategy. Having an only limited number of available protein domains is another obstacle for broad application of this strategy. The most commonly used scaffold protein domains are limited to the three candidates from Dueber's study, whereas few tried to adapt new protein domains into the scaffold (Kim and Hahn, 2014) and none have built a scaffold protein with 4 or more protein domains. However, there is need to co-localize more than three enzymes since most of the pathways we are optimizing today involve three or more enzymes, such as astaxanthin synthesis in *Saccharomyces cerevisiae* and violacein biosynthetic pathway (Ukibe et al., 2009; Lee et al., 2013). Meanwhile, the scaffold systems engineered to recruit pathway enzymes are predominately in *E. coli*, while there is an urgent demand for them in yeast and higher eukaryotes (Horn and Sticht, 2015).

In this study, we designed a standardized assembly protocol and constructed a set of Golden-Gate-based molecular parts, which we named Artificial Protein Scaffold System (*AProSS*). We designed and constructed a hierarchical three-step assembly system to integrate any protein domain with variable repeats into a scaffold protein. On the other hand, we searched through structural studies and found three strong interaction protein domains with their corresponding ligands. These three domains were named as FE, YAP, and BIR. FE, and YAP are two WW domains of human FE65 (Meiyappan et al., 2007) and YAP65 protein (Macias et al., 1996), and BIR is a protein domain from inhibitor-of-apoptosis protein(IAP) (Liu et al., 2000). We chose the interactional center of these proteins and synthesized the gene *de novo*, with subsequently modularized the protein domains into a standardized library together with the three previously verified domains, namely, GBD, SH3 and PDZ. The domain-ligand interactions were verified by yeast two-hybrid and split GFP assays. Constructing scaffold proteins with *AProSS* increased the production of violacein and deoxyviolacein by 29% and 63% respectively, while violacein/deoxyviolacein ratio increased by 18% with rewired flux.

## 2. Materials and methods

### 2.1. Strains and cultivation

*E. coli* DH5α was used for plasmid construction and amplification. Yeast strain JDY26 (*MATa ade2–101 trp1–901 leu2–3.112 his3Δ200 ura3–52 gal4Δ gal80Δ SPAL::URA3 LYS2::GAL1-HIS GAL2-ADE2 met2:GAL7-LacZ can1R*) was used to host the constructs and subjected to yeast two-hybrid analysis. Yeast strain BY4741 (*MATa leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*) was used to host the split GFP constructs and subjected to FACS analysis. Yeast strain JDY52-URR-His3 (*MATa his3Δ200 leu2Δ0 lys2Δ0 trp1Δ63 ura3Δ0 met15Δ0*), a derivative of S288C with I-*Sce*I recognition sites flanking the URR1-His3-URR2 fragment at HO locus (Guo et al., 2015) was used as a host for violacein biosynthetic pathway. Yeast cells were cultured in YPD (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose) and SC-Ura [6.7 g/L YNB, 0.01μmol/L $Fe(NH_4)_2(SO_4)_2$, 20 g/L glucose; The mixture of amino acids and other nutrients] medium.

### 2.2. Plasmids construction

For our designed plasmids, we took advantage of the pSMART®HCKan and pSMART®LCAmp plasmids combined with the cloned RFP gene. We added two *Ear*I restriction endonuclease sites to both terminals of the linear DNA by PCR. The RFP gene was also constructed by PCR with additional *Ear*I on the primers, followed by *Ear*I digestion and ligation. With the KanR and AmpR genes already on the vectors, we inserted the *Bsa*I and *Bsm*BI onto the upstream and downstream of the RFP site by mutagenic PCR. For pLV vectors, we separated the two digestion sites with *Bsm*BI inside and *Bsa*I flanking outside around the RFP gene and inserted a 9-amino acid of glycine-serine repeat between the *Bsa*I site and *Bsm*BI site upstream, in frame with the RFP starting from $pLV_2$.

### 2.3. Selection and synthesis of new protein interaction domains

The new protein interaction domains were cherry-picked by reading through structural studies about protein-ligand interaction (Dueber et al., 2009). After we acquired the DNA sequence of the protein domain, we removed the stop codon, optimized the codon composition using *GeneDesign* (Richardson et al., 2006), mutated the *Bsa*I, *Bsm*BI and *Ear*I digestion sites, and added the additional sequences to the 5′ and 3′ ends:

Upstream:

caggaaacagctatgacc**GGTCTC**aCTCGnnnnnnnnnnnnnnnnnnnn

Downstream:

tgtaaaacgacggccagt**GGTCTC**gTTGAnnnnnnnnnnnnnnnnnnnn

"nnn" indicates the DNA region of protein domain, the bold and underline parts are the *Bsa*I recognition and digestion overhangs. The whole sequence could then be synthesized through primer alignment (Richardson et al., 2006) or direct chemical synthesis.

### 2.4. One-pot AProSS assembly

One-pot Golden Gate assembly is described in many studies (Engler et al., 2008, 2009; Yuan et al., 2017). In the *AProSS* method, constructing a scaffold ORF composes of three steps. The first step, clone PCR product into a kanamycin resistant vector pMV, mixed with reaction mix [1.5 μL 10 × T4 DNA ligase reaction buffer (New England BioLabs, M0202), 0.15 μL 100 × bovine serum albumin (BSA, New England BioLabs), 2.5 U T4 DNA ligase (Enzymatics, Beverly, MA, L6030-HC-L), and 10 U of *Bsa*I (New England BioLabs, Beverly, MA, R0535 or R0580, respectively)] to a final volume of 15 μL. One-pot digestion-ligation assembly was performed in a thermocycler as follows: 37 °C for 60 min, 50 °C for 15 min and 80 °C for 15 min. Five microliters of each assembly reaction were transformed into 100 μL of competent DH5α *E. coli* cells and plated on the appropriate selection media. The second step, subclone the fragment into a ampicillin resistant vector pLV with location order, mixed with reaction mix [1.5 μL 10 × T4 DNA ligase reaction buffer (New England BioLabs, M0202), 0.15 μL 100 × bovine serum albumin (BSA, New England BioLabs), 2.5 U T4 DNA ligase (Enzymatics, Beverly, MA, L6030-HC-L), and 10 U of *Bsm*BI (New England BioLabs, Beverly, MA, R0535 or R0580, respectively)] to a final volume of 15 μL. One-pot digestion-ligation assembly was performed in a thermocycler as follows: 55 °C for 60 min and 25 °C for 60 min with T4 DNA ligase was added as soon as the thermocycler cool to 25 °C, followed 55 °C for 15 min and 80 °C for 15 min Product was transformed into competent DH5α *E. coli* cells as the described in the first step. The third step, similar to the first step with a series of