BIOTECHNOLOGICAL PRODUCTS AND PROCESS ENGINEERING

Experimental design-aided systematic pathway optimization of glucose uptake and deoxyxylulose phosphate pathway for improved amorphadiene production

Congqiang Zhang · Ruiyang Zou · Xixian Chen · Gregory Stephanopoulos · Heng-Phon Too

Abstract Artemisinin is a potent antimalarial drug; however, it suffers from unstable and insufficient supply from plant source. Here, we established a novel multivariate-modular approach based on experimental design for systematic pathway optimization that succeeded in improving the production of amorphadiene (AD), the precursor of artemisinin, in *Escherichia coli*. It was initially found that the AD production was limited by the imbalance of glyceraldehyde 3-phosphate (GAP) and pyruvate (PYR), the two precursors of the 1-deoxy-D-xylulose-5-phosphate (DXP) pathway. Furthermore, it was identified that GAP and PYR could be balanced by replacing the phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) with the ATP-dependent galactose permease and glucose kinase system (GGS) and this resulted in fivefold increase in AD titer (11 to 60 mg/L). Subsequently, the experimental design-aided systematic pathway optimization (EDASPO) method was applied to systematically optimize the transcriptional expressions of eight critical genes in the glucose uptake and the DXP and AD synthesis pathways. These genes were classified into four modules and simultaneously controlled by T7 promoter or its variants. A regression model was generated using the four-module experimental data and predicted the optimal expression ratios among these modules, resulting in another threefold increase in AD titer (60 to 201 mg/L). This EDASPO method may be useful for the optimization of other pathways and products beyond the scope of this study.

Keywords Amorphadiene · Experimental design-aided systematic pathway optimization · Multivariate-modular approach · Deoxyxylulose phosphate pathway · The phosphotransferase system

C. Zhang · R. Zou · X. Chen · G. Stephanopoulos · H.-P. Too
Chemical and Pharmaceutical Engineering, Singapore-MIT Alliance, 4 Engineering Drive 3, Singapore, Singapore

C. Zhang · X. Chen · H.-P. Too (✉)
Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, 8 Medical Drive, Blk MD7, #05-04, Singapore 117597, Singapore
e-mail: heng-phon_too@nuhs.edu.sg

G. Stephanopoulos
Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA

Introduction

In *Escherichia coli* and other bacteria, isoprenoids are synthesized through the 1-deoxy-D-xylulose-5-phosphate (DXP) pathway originating from two precursors (pyruvate (PYR) and glyceraldehyde 3-phosphate (GAP)). As equal amounts of two precursors are optimal for the DXP pathway, the imbalance of the two precursors could reduce the biosynthesis of isoprenoids. The conversion from G3P to pyruvate involves three reversible enzymatic steps and one essentially irreversible step (between phosphoenolpyruvate (PEP) and PYR, Fig. 1) (Farmer and Liao 2001). And it was reported that redirection of metabolic flux from PYR to PEP by the overexpression of the gluconeogenic enzyme, PEP synthetase (*ppsA*), enhanced lycopene yield (Farmer and Liao 2000; Farmer and Liao 2001). The results indicate that it is possible to balance the G3P and PYR through the manipulation of the ratio between PEP and PYR, and redirection of metabolic flux
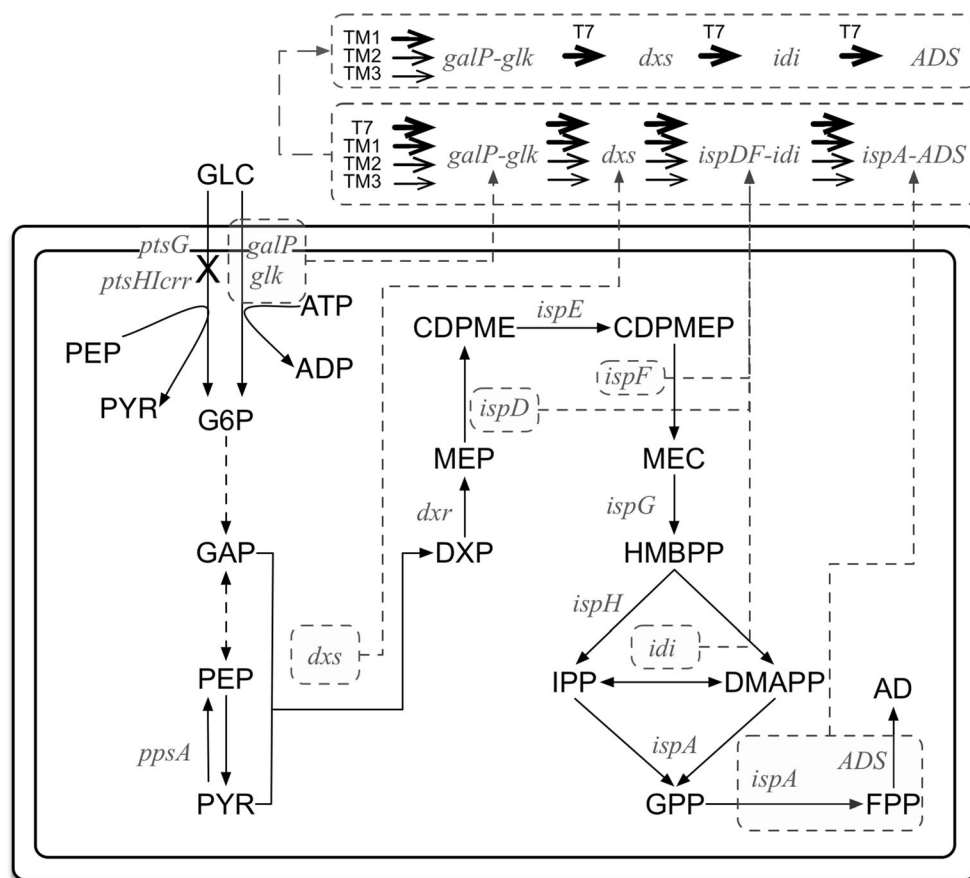
Fig. 1 Global control and balance of metabolic pathways for AD production. T7, TM1, TM2, and TM3 promoters were used to fine control the expression of the four different modules (*galP-glk*, *dxs*, *ispDF-idi* or *idi*, and *ispA-ADS*). The metabolites in the figure were as follows: glucose (GLC), glucose-6-phosphate (G6P), glyceraldehyde 3-phosphate (GAP), phosphoenolpyruvate (PEP), pyruvate (PYR), 1-deoxy-D-xylulose 5-phosphate (DXP), 2C-methyl-D-erythritol 4-phosphate (MEP), 4-diphosphocytidyl-2C-methyl-D-erythritol (CDPME), 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate (CDPMEP), 2C-methyl-D-erythritol 2,4-diphosphate (MEC),

hydroxylmethylbutenyl diphosphate (HMBPP), isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), geranyl diphosphate (GPP), farnesyl pyrophosphate (FPP), and amorpha-4,11-diene (AD). The enzyme-coding genes in the figure were as follows: PTS enzyme IIBC (*ptsG*), histidine protein (*ptsH*), PTS enzyme I (*ptsI*), PTS enzyme IIA(*crr*), galactose permease (*galP*), glucose kinase (*glk*), phosphoenolpyruvate synthetase (*ppsA*), DXP synthase (*dxs*), DXP reductase (*dxr*), CDPME synthase (*ispD*), CDPME kinase (*ispE*), CDPMEP synthase (*ispF*), HMBPP synthase (*ispG*), HMBPP reductase (*ispH*), IPP isomerase (*idi*), *FPP* synthase (*ispA*), and *AD* synthase (*ADS*)

from PYR to GAP is beneficial for isoprenoid production. Similarly, inactivation of the carbohydrate phosphotransferase system (PTS) can also increase the ratio between PEP and PYR (Chatterjee et al. 2001; Flores et al. 1996; Zhang et al. 2009b), thus is potentially applicable to enhance isoprenoid biosynthesis in *E. coli*.

The phosphotransferase system (PTS) is composed of enzyme I (EI), histidine protein (HPr), sugar-specific enzyme IIA (EIIA), and enzyme IICB (EIICB), encoded by gene *ptsHIcrr* operon and *ptsG*, respectively. The PTS is the major glucose uptake pathway in *E. coli*, and half of the PEP produced from glucose is metabolized into PYR during the glucose uptake by the PTS (Gorke and Stulke 2008). Thus, the deletion of the PTS can increase the intracellular PEP/PYR ratio and thus increase the intracellular GAP/PYR ratio. However, the disruption of the PTS resulted in the decreased glucose uptake rate and significantly impeded cell growth

(Martínez et al. 2008). To address this issue, adaptive evolution was used to select fast-growing strains (Flores et al. 1996). Transcriptome analysis in these fast-growing strains indicated that the expressions of gene *galP* (coding for galactose permease), *glk* (coding for glucose kinase), and *pgi* (coding for phosphoglucose isomerase) were increased as compared to the wild-type strain, suggesting that these genes were important in enhancing glucose uptake (Flores et al. 2007; Flores et al. 2002). Hence, it was reported that the overexpression of *galP* and *glk* enhanced the growth rate of microbes in glucose medium (Flores et al. 2007; Lu et al. 2012) and improved the yields of valuable products such as aromatics (Yi et al. 2003), recombinant protein (De Anda et al. 2006), and ethanol (Hernandez-Montalvo et al. 2003). Combinatorial modulation of *galP* and *glk* gene expressions has been used to improve cell growth rate (Lu et al. 2012), but it is unclear if such an approach will increase the production of isoprenoids.

Pathway optimization by modulating the activities and expressions of enzymes in the pathway is critical for improving product yields. Optimized pathways can minimize the accumulation of intermediate metabolites with potential cytotoxicity and thus decrease the metabolic burdens and increase the final product titers (Anthony et al. 2009; Pfleger et al. 2006). Multivariate-modular approach to balance metabolic pathway has been successfully used to systematically search for the optimal conditions for taxadiene production, resulting in 15, 000-fold increase in taxadiene titer over the control (Ajikumar et al. 2010). Modulating plasmid copy numbers or engineering different types of promoters (Jones et al. 2000) has been used to control and balance the expression levels of multiple genes for more than 10 years. However, these approaches are time consuming and resource demanding and may not cover a sufficiently broad search space to allow the identification of conditions for optimal productivity.

In this study, the hypothesis that amorphadiene (AD) production from glucose can be improved by enhancing the availability of GAP through PEP was tested. By deleting the PTS and overexpressing the ATP-dependent galactose permease and glucose kinase system (GGS), the AD production was significantly increased by fivefold. Optimization of the production of AD was then carried out by experimental design-aided systematic pathway optimization (EDASPO) to globally regulate four distinct modules (including eight critical genes in the glucose uptake and the DXP and AD synthesis pathways) with small number of experiments. Using T7 promoter variants with different transcriptional efficiencies to control multiple modules simultaneously (Too et al. 2014), EDASPO method enables the rapid discovery of the optimal conditions for high productivity with a small number of experiments. Based on the relative transcriptional efficiencies of T7 promoter and its variants, regression analysis of the four-module data predicted the optimal expression ranges for each module to further increase AD production. And the prediction was validated experimentally, resulting in another threefold enhancement in AD titer.

Materials and methods

Bacterium strains and plasmids

E. coli K12 MG1655 ΔrecAΔendA DE3 was used as the parental strain (Ajikumar et al. 2010). MG1655 ΔPTS::FRT (PTS- for short) strain was obtained by replacing operon *ptsHIcrr* with a kanamycin-resistant gene using the λ Red recombination system (Zhang et al. 2013). And the kanamycin-resistant gene was subsequently excised by the Flp/FRT site-specific recombination system (Datsenko

2000). Chromosomal deletions were verified by testing for antibiotic markers and colony PCR analysis.

Plasmid pET-*galP-glk* was constructed from three parts, pETaa vector, gene *galP*, and gene *glk*, by cross-lapping in vitro assembly (CLIVA) method using I-pETR(galPF), I-galPF(pETR), I-galPR(glkF), I-glkF(galPR), I-pETF(glkR), and I-glkR(pETF) (Table 1) (Zou et al. 2013). Plasmids pETA-TM1/2/3-*galP-glk* were constructed from pET-*galP-glk* by mutating T7 promoter to TM1/TM2/TM3 promoter using the oligo TM1, TM2, and TM3 (Table 1). Plasmid p15A-spec-T7-*ppsA* was constructed from two parts, 15A-spec vector and *ppsA* gene, by CLIVA method using I-*ppsA*(t7pr)-F, I-*ppsA*(t7tf)-R, I-T7P(-)-R, and I-T7T(-)-F (Table 1).

All the information about the strains, plasmids, and primers is in Tables 1 and 2.

Construction of the four-module expression system

The four modules were controlled simultaneously under T7 and T7 variant promoters with different transcriptional efficiencies (TM1, TM2, and TM3; the transcriptional efficiencies of TM1, TM2, and TM3 were 92, 37, and 16 % of that of T7 promoter, respectively). TM1, TM2, and TM3 promoters were generated as described in the patent (Table 1) (Too et al. 2014). A portion of combinations was chosen based on the principles of experimental design and a priori knowledge about three-module (*dxs*, *ispDF-idi*, *ispA-ADS*) optimization results. In the three-module experiments, the construct 231 (where the three modules *dxs*, *ispDF-idi*, and *ispA-ADS* were controlled by TM2, TM3, and TM1 promoters, respectively) and the construct 332 (where the three modules *dxs*, *ispDF-idi*, and *ispA-ADS* were controlled by TM3, TM3, and TM2 promoters, respectively) produced the highest titers of AD and thus were included in this experiment. To cover a broad range, each module was controlled by three different promoters (TM1, TM2, and TM3). Each promoter for each module was repeated at least three times (for example, TM1 for the first module was used in four constructs, P1231, P1332, P1121, and P1313). As a result, 14 construct were chosen and tested with four different concentrations of isopropyl β-D-1-thiogalactopyranoside (IPTG) (Table 3).

Media and culture conditions

Defined media composition

The defined media without carbon contained 4 g/L (NH$_4$)$_2$HPO$_4$, 3.50 g/L KH$_2$PO$_4$ and 9.28 g/L K$_2$HPO$_4$,

Table 1 Strains, plasmids, and primers used in this study

Strains/plasmids/primers	Description/sequence	References
Strains		
MG1655 DE3	*E. coli* K12 MG1655 $\Delta recA \Delta endA$ DE3	(Ajikumar et al. 2010)
MG1655 DE3 ΔPTS::FRT	MG1655 DE3 derived, $\Delta pstHIcrr$::FRT	This study
Plasmids		
p15A-spec-T7-*ppsA*	pAC-LYC derived (Cunningham et al. 1994) by replacing *crtEBI* with *ppsA* under the control of T7 promoter and replacing chloramphenicol-resistant gene with spectinomycin-resistant gene	This study
pETA-*galP-glk*	pET-11a (Novagen) derived, *galP-glk* operon under the control of T7 promoter	This study
pETA-TM1-*galP-glk*	pETA-*galP-glk* derived by mutating T7 promoter to TM1 promoter	This study
pETA-TM2-*galP-glk*	pETA-*galP-glk* derived by mutating T7 promoter to TM2 promoter	This study
pETA-TM3-*galP-glk*	pETA-*galP-glk* derived by mutating T7 promoter to TM3 promoter	This study
pBAD-galP-glk	pBAD-B (Invitrogen) derived, *galP-glk* operon was placed under the control araBAD promoter	This study
pACM-T7-*dxs*-T7-*idi*-T7-*ADS*-*ispA*	pAC-LYC derived, *dxs-idi-ADS-ispA* operon was placed under the control of T7 promoter	This study
pACM-T7-*dxs-idi-ispDF*-T7-*ADS-ispA*	pAC-LYC derived, *dxs-idi-ispDF-ADS-ispA* operon was placed under the control of T7 promoter	This study
pETK-lacI	pET-11a derived, ampicillin-resistant gene was replaced with kanamycin-resistant gene using CLIVA	This study
pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF*, and *ADS* were placed under the control of TM2, TM3, and TM1 promoters, respectively	(Too et al. 2014)
pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF*, and *ADS* were placed under the control of TM3, TM3, and TM2 promoters, respectively	(Too et al. 2014)
pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF*, and *ADS* were placed under the control of TM1, TM2, and TM1 promoters, respectively	(Too et al. 2014)
pACM-TM1-*dxs*-TM1-*idi-ispDF*-TM1-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF*, and *ADS* were placed under the control of TM1, TM1, and TM1 promoters, respectively	(Too et al. 2014)
pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	pAC-LYC derived, *dxs*, *idi-ispDF*, and *ADS* were placed under the control of TM3, TM1, and TM3 promoters, respectively	(Too et al. 2014)
Primers		
TM1 promoter	TAATACGACTCACTA<u>AT</u>GGGGA	This study
TM2 promoter	TAATACGACTCACT<u>CG</u>AGGGGA	This study
TM3 promoter	TAATACGACTCACTATAG<u>AA</u>GA	This study
I-pETR(*galP*F)	AGCGTCAGGCAT<u>A*TGTATATCTCCTTC*TTAAAGTT</u>	This study
I-*galP*F(pETR)	AAGGAGATATA*CAT<u>ATGCCTGACGCT*AAAAAAC</u>	This study
I-*galP*R(*glk*F)	GATCTGCCTCCT*<u>TTAATCGTGAGCG*CCTA</u>	This study
I-*glk*F(*galP*R)	CGCTCACGATTA*A<u>AGGAGGCAGATC*AAATGACAAAG</u> TATGCATTAGTC	This study
I-pETF(*glk*R)	GGTCACATTCT*GTA<u>AAAGCTGAGTTGGC*TGC</u>	This study
I-*glk*R(pETF)	GCCAACTCAGCTT*<u>TTACAGAATGTGACC*TAAGG</u>	This study
I-gg-F(pBADR)	ATTAACCA*TGCCTG*ACGCTAAAAAAC	This study
I-gg-R(pBADF)	AACAGCCT*TACAGA*ATGTGACCTAAGG	This study
I-pBAD-F(ggR)	TCTGTAAG*GCTGTT*TTGGCGGATGAG	This study
I-pBAD-R(ggF)	CAGGCATG*GTTAAT*TCCTCCTGTTAGCCC	This study
I-*ppsA*(t7pr)-F	CCCTCTAGAA*ATAATTTTGTT*TTTCTCAAACCGTTCATTTA	This study
I-*ppsA*(t7tf)-R	CCAACTCA*GCTTCCTTT*AACGCAGGATGTCTGTGAA	This study
I-T7P(-)-R	AACAAAATTATT*TCTAGAGGG*GAA	This study
I-T7T(-)-F	AAAGGAAG*CTGAGTTGG*CTG	This study

The asterisked oligos were phosphorothioate-modified oligos used for CLIVA cloning (Zou et al. 2013)

1.7 g/L citric acid, 0.5 g/L $MgSO_4$, 2.5 mg/L $CoCl_2 \cdot 6H_2O$, 15.0 mg/L $MnSO_4 \cdot 4H_2O$, 1.5 mg/L $CuSO_4 \cdot 2H_2O$, 3 mg/L H_3BO_3, 2.5 mg/L $Na_2MoO_4 \cdot 2H_2O$, 13 mg/L $Zn(CH_3COO)_2$, 60 mg/L Fe(III) citrate, and 8.4 mg/L EDTA. Glucose (20 g/L) was added into the defined media before use.

Table 2 Name list of strains carrying different plasmids in this study

Strains	Plasmids	Short name	Reference
MG1655 DE3 ΔPTS::FRT	pETA-*galP-glk* and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	PGG	This study
MG1655 DE3	pETA-*galP-glk* and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	MGG	This study
sMG1655 DE3	pETK-lacI and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	MGG01	This study
MG1655 DE3	p15A-spec-T7-*ppsA*, pETK-lacI and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	MGG02	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk* and pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	P1231	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	P2231	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-TM2-*dxs*-TM3-*idi-ispDF*-TM1-*ADS*	P3231	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk* and pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	P1332	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	P2332	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-TM3-*dxs*-TM3-*idi-ispDF*-TM2-*ADS*	P3332	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk* and pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	P1121	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	P2121	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-TM1-*dxs*-TM2-*idi-ispDF*-TM1-*ADS*	P3121	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-TM1-*dxs*-TM1-*idi-ispDF*-TM1-*ADS*	P2111	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-TM1-*dxs*-TM1-*idi-ispDF*-TM1-*ADS*	P3111	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk* and pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	P1313	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	P2313	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-TM3-*dxs*-TM1-*idi-ispDF*-TM3-*ADS*	P3313	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM1-*galP-glk* and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	P1000	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM2-*galP-glk* and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	P2000	This study
MG1655 DE3 ΔPTS::FRT	pETA-TM3-*galP-glk* and pACM-T7-*dxs*-T7-*idi-ADS-ispA*	P3000	This study

All the amorpha-4,11-diene titers reported here were based on the strains grown in the defined media supplemented with 20 g/L glucose.

2xpy medium composition

The 2xpy medium contained 20 g/L peptone, 10 g/L yeast extract, and 10 g/L NaCl. The medium was adjusted to pH 7.0 and autoclaved at 121 °C for 20 min.

Culture conditions

New transformed strains were rescued and grown in 2xpy medium overnight, and 1 % (*v/v*) overnight grown cell culture was inoculated into 1 mL defined media in a 14-mL BD Falcon™ tube. Dodecane with 100 mg/L beta-caryophyllene was added to 20 % (*v/v*) of the medium volume. Cells were grown at 28 °C with 300 rpm shaking speed and induced with different dosages of IPTG when OD_{600} reached the range of 0.5~1.0. The media were supplemented with 100 mg/L ampicillin, 34 mg/L chloramphenicol, 50 mg/L kanamycin, and 100 mg/L spectinomycin to maintain the pET-11a-derived plasmids, pAC-LYC-derived plasmids, pETK-lacI, and p15A-spec-T7-*ppsA*, respectively.

AD quantification

AD was extracted by diluting 5 μL dodecane phase into 495 μL ethyl acetate and analyzed on an Agilent 7980A gas chromatograph equipped with an Agilent 5975C mass spectrometer (GC/MS). Beta-caryophyllene (Sigma-Aldrich) was used as an equivalent, and the GC-MS condition was set as the method reported by scanning for only two ions (204 and 189 *m/z* ion) (Tsuruta et al. 2009).

Calculation of relative expression of each module

Relative expression level of each module was calculated based on the promoter transcriptional efficiencies and plasmid copy numbers. It is calculated using the following equation.

$$E_i(C) = \frac{M_i \times N_i}{\sum\limits_{j=1}^{4} M_j \times N_j}$$

where E_i is the relative expression level of the module i in the construct C, M_i is the promoter transcriptional efficiency of module i, and N_i is the plasmid copy number of module i.

For example, the transcriptional efficiencies of TM1, TM2, and TM3 were 92, 37, and 16 % of that of T7 promoter,

respectively; the first module was in the pAC-derived vector (p15a origin) with around ten copies, and the other three modules were in the pET-derived vector (pBR322 origin) with about 30 copies. Therefore, the relative expression level of the first module (*galP-glk*) in strain P1231 was

$$E_1(P1231) = \frac{92\ \%\ \times\ 10}{92\ \%\ \times\ 10\ +\ (37\ \%\ +\ 16\ \%\ +\ 92\ \%)\ \times\ 30} = 0.17$$

Results

Overexpression of *ppsA* increased AD production

In this study, the native DXP pathway in *E. coli* was used to produce AD (Fig. 1). The DXP pathway starts from PYR to GAP. PYR is abundant inside the cells and known to be exported out of the cells (Kreth et al. 2013), and the redirection of metabolic flux from PYR to GAP by the overexpression of PEP synthetase (*ppsA*) has been shown to enhance lycopene yield (Farmer and Liao 2000). To test the hypothesis that the overexpression of *ppsA* can similarly enhance the production of another isoprenoid, AD, cells were transformed with a plasmid encoding the endogenous *ppsA* gene under the control of the T7 promoter. As high expression of *ppsA* is toxic and negatively affects the growth of *E. coli* (Farmer and Liao 2000; Patnaik et al. 1992), the expression of *ppsA* was moderately induced at levels that did not significantly affect cell growth (Fig. 2b). As shown in Fig. 2, the AD titer (mg/L) and specific yield (mg/L/OD) increased to about twofold when *ppsA* was overexpressed in all the conditions examined (different concentrations of IPTG). The result was consistent with the hypothesis that increasing the availability of GAP through increasing the intracellular PEP (Fig. 1) would increase AD production, albeit modestly.

Overexpression of *galP* and *glk* increased AD production

Similar to the *ppsA* overexpression, the PTS knockout can also increase the intracellular PEP. As glucose uptake via the PTS is known to consume half of PEP produced from glucose and produces PYR (Hernandez-Montalvo et al. 2003), an attempt was made to use an alternative pathway for the cells to import glucose. Unlike the PTS which uses PEP as phosphate donor, the GGS transports and phosphorylates glucose using ATP as phosphate donor, thus increasing intracellular PEP and increasing the availability of intracellular GAP. As shown in Fig. 3, overexpression of the GGS in the wild type using T7 promoter (The resulting strain was named as MGG, Table 2) resulted in an increase in AD titer (mg/L) by more than threefold (from 10 to 37 mg/L) as compared to wild type (MGG01, Table 2). With the PTS deletion and the overexpression of the GGS, AD titer (The resulting strain was named as PGG,

Table 2) was further increased by 70 % to ~60 mg/L as compared to that of MGG. The specific yield of AD increased from about 7 mg/L/OD (MGG) to ~12 mg/L/OD (PGG). The results indicated that the glucose uptake pathway played a significant part in the AD production where the ATP-dependent GGS (PGG strain, Table 2) was superior to the PEP-dependent PTS (MGG01 strain, Table 2). Noteworthy, the cell density was similar for the three strains in most of the conditions except that the cell density of MGG and PGG was very low when high concentration of inducer was used. This suggested that the high expression of the GGS could be toxic to cells.

Differential control of multivariant modules encoding critical enzymes in the glucose uptake and the DXP and AD synthesis pathways

A two-module optimization of the upstream DXP pathway and heterologous downstream terpenoid-forming pathway increased Taxol precursor titer significantly (Ajikumar et al. 2010). As the replacement of the PTS with the GGS increased AD yield, we hypothesized that the global pathway optimization of the GGS and the DXP and AD synthesis pathways could further increase AD production. It was reported that overexpression of *dxs*, *ispDF*, and *idi* improved the yields of carotenoids (Alper et al. 2005; Yuan et al. 2006) and taxol precursor (Ajikumar et al. 2010). In addition, *ADS* has been identified as a critical step for the AD production (Anthony et al. 2009; Pitera et al. 2007). All these important enzymes and glucose uptake transporters were grouped into four modules. The first module encoded *galP-glk*, which regulates glucose uptake. The second module encoded *dxs*, which diverts the carbon flux from central metabolic pathway into the DXP pathway. The operon *ispDF-idi* was encoded in the third module where *ispDF* facilitates the conversion of intermediates of the DXP pathway and *idi* is crucial for balancing the ratio of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Lastly, *ispA* and *ADS* genes were encoded in the fourth module where enzyme *ispA* condensates IPP and DMAPP into FPP and FPP is cyclized into AD by enzyme *ADS* (Fig. 1).

All the four modules were controlled simultaneously under the native T7 promoter or its variants (TM1, TM2, and TM3) in a combinatorial manner. As a result, the total combination of four-module system with four different promoters has $4^4 =$

256 constructs. To limit the number of conditions to be tested, 14 constructs (Fig. 4d) were analyzed based on the experimental design principle (the details about these constructs were explained in "Materials and methods"). As shown in Fig. 4, through the differential control of the four modules, AD titer varied significantly by more than 190-fold from 0.36 mg/L (strain P1313 with 0.03 mM IPTG) to 68.9 mg/L (strain P2121 with 0.01 mM IPTG).

Regression analysis of the four-module expression system

To identify the global optimal condition and to gain an insight into the behavior of the four-module expression system, a linear regression model was used to analyze the experimental data in Fig. 4. The relative expression of each module (Table 3) was calculated from the transcriptional efficiencies of promoters and copy numbers of plasmids, as described in "Materials and methods". As the relative expression of each module is independent of the inducer dosage, the averages of AD titers at different inducer dosages were used in the regression analysis. The experimental data and regression model were shown in three-dimensional figures for each two modules (Fig. S1). For the ease of visualization and analyses, the data were also demonstrated as two-dimensional heat maps for each two modules using a linear regression model (Fig. 5). Based on the analyses, the optimal ranges for the four

Table 3 Relative expression levels of each module in the 14 constructs of four-module expression system

No.	Strains[a]	Relative expression levels of each module			
		galP-glk	dxs	ispDF-idi	ispA-ADS
1	P1231	0.17	0.21	0.09	0.52
2	P2231	0.08	0.24	0.10	0.58
3	P3231	0.04	0.25	0.11	0.61
4	P1332	0.31	0.16	0.16	0.37
5	P2332	0.15	0.20	0.20	0.45
6	P3332	0.07	0.22	0.22	0.50
7	P1121	0.12	0.37	0.15	0.37
8	P2121	0.05	0.39	0.16	0.39
9	P3121	0.02	0.41	0.16	0.41
10	P2111	0.04	0.32	0.32	0.32
11	P3111	0.02	0.33	0.33	0.33
12	P1313	0.20	0.10	0.59	0.10
13	P2313	0.09	0.12	0.67	0.12
14	P3313	0.04	0.12	0.71	0.12

[a] The strains were named based on their sequence of promoters; numbers "1," "2," and "3" stood for TM1, TM2, and TM3 promoters, respectively (e.g., P1231 referred to a construct in which galP-glk, dxs, idi, and ispA-ADS were controlled by TM1, TM2, TM3, and TM1 promoters, respectively)

modules were well accounted by the 14 chosen constructs. The predicted optimal ranges of relative expression levels of the four modules (galP-glk, dxs, ispDF-idi, and ispA-ADS) were 0.04~0.1, 0.25~0.4, 0.1~0.4, and 0.25~0.44, respectively. Thus, a relatively lower expression of the first module (galP-glk) was found to be beneficial for AD production. Higher expression levels of the second (dxs) and the fourth modules (ispA-ADS) were required. Interestingly, for the third module (ispDF-idi), both low expression (~0.15) and high expression (~0.35) resulted in higher AD production. Especially in Fig. 5b, there were two optimal ranges for the third module (ispDF-idi) when the relative expression level of the first module (galP-glk) is within 0.04~0.1. Further studies demonstrated the operon ispDF in the third module was less important and may even be adverse for AD production in certain conditions (Fig. S2).

Further improvement of AD production facilitated by regression analysis

From the regression analysis of four-module expression system data, we concluded that the first module encoding galP-glk was beneficial when expressed at a relatively lower level; modules encoding dxs and ispA-ADS were required to be expressed at relatively higher levels, and the expression of ispDF in third module was unnecessary. Thus, we constructed strains where the genes dxs, idi, and the operon ispA-ADS were under the control of the native T7 promoter, and the operon galP-glk was controlled by TM1, TM2, or TM3 promoter (Fig. 1, the details of the three strains were shown in Table 2). The relative expression levels of the four modules for the three strains were shown in Table 4. As a result, the highest AD titer in strain P2000 increased by ~3-fold when compared to that of P2121, an increase from 68.9 to 201.2 mg/L (Fig. 6). In addition, the specific yield was also enhanced significantly from ~12.2 to ~18.0 mg/L/OD, validating the prediction from regression analysis. Interestingly, similar to other 14 strains of the four-module expression systems, the cell density of strain P1000, P2000, and P3000 was inversely correlated with the inducer dosages (Fig. 6b).

Discussion

In E. coli, isoprenoid biosynthesis is reported to be limited by imbalance of GAP and PYR, or specifically, the GAP is less abundant than the PYR. As PEP and GAP are interconvertible with three reversible enzymatic reactions, the overexpression of ppsA (catalyzing PYR to PEP) or pck (converting oxaloacetate to PEP) can increase the availability of intracellular GAP by increasing intracellular PEP and thus enhanced the production of lycopene (a C40 isoprenoid) (Farmer and Liao 2000; Farmer and Liao 2001). Similarly, the AD production was
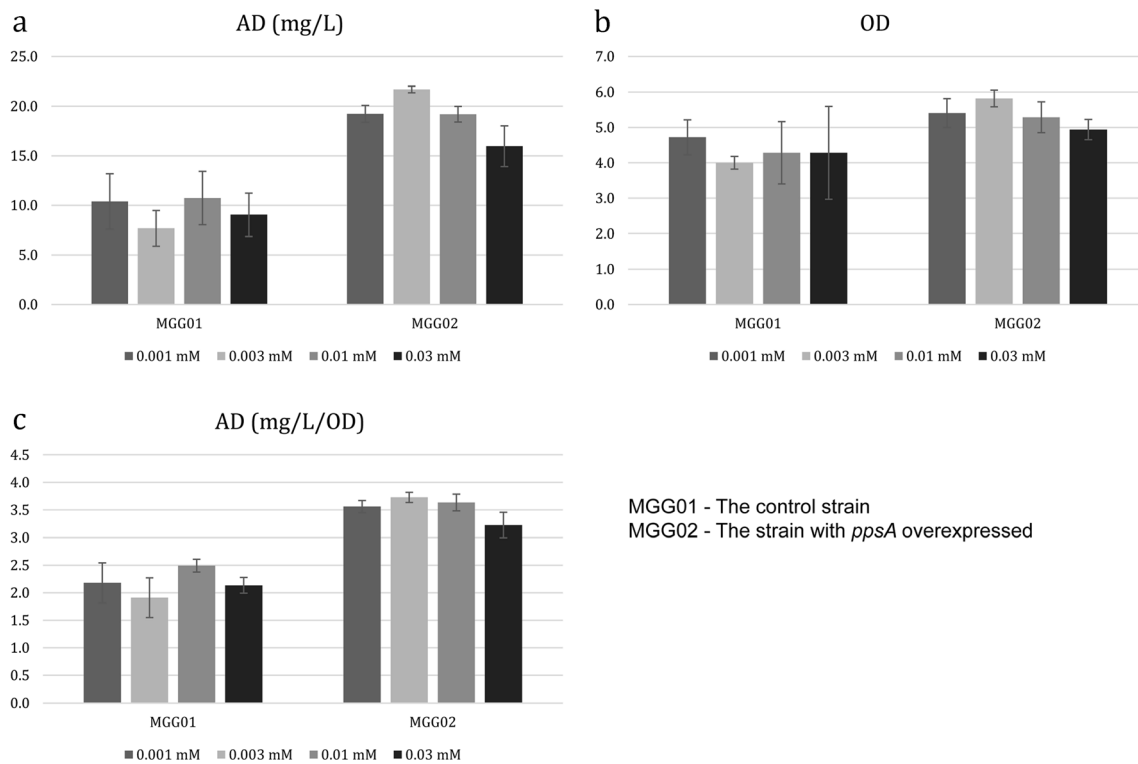
Fig. 2 Overexpression of *ppsA* increased AD production. **a** AD titer (mg/L). **b** Cell density of different strains. **c** Specific yield of AD (mg/L/DO). *OD* optical density of culture at 600 nm. MGG01 strain was the control strain without overexpressing *ppsA*, whereas *ppsA* was overexpressed in MGG02 strain
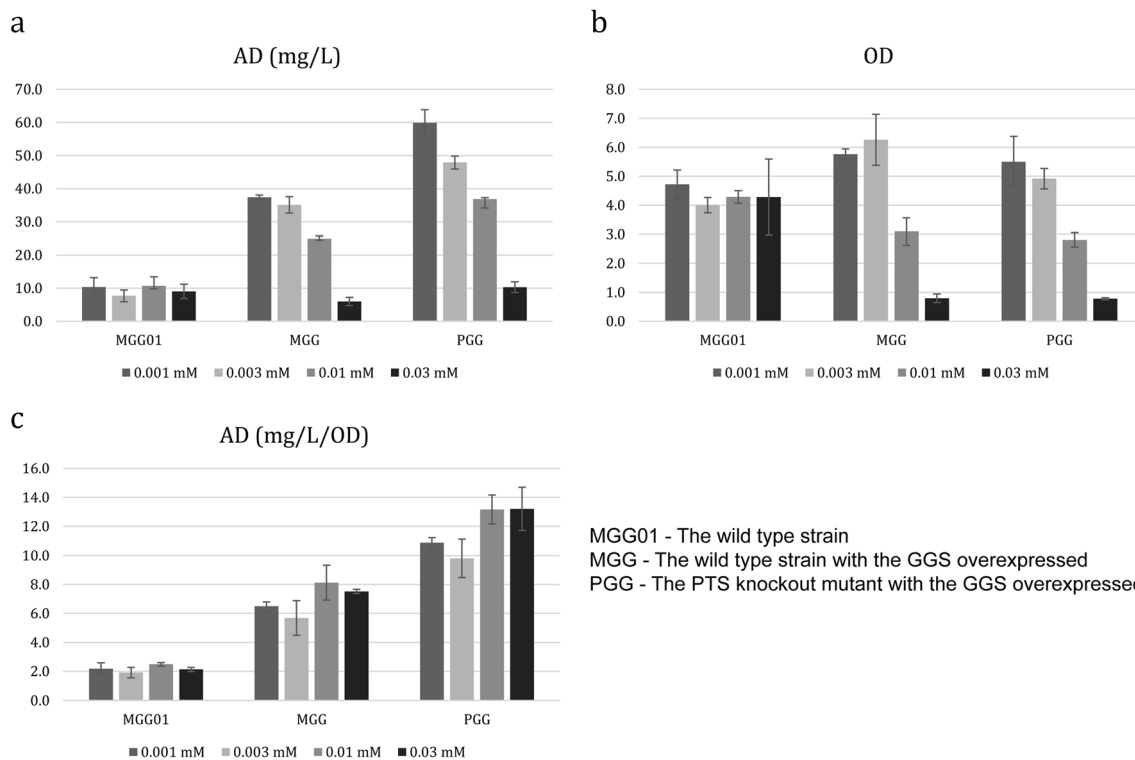

Fig. 3 The effect of different glucose uptake pathways (the GGS and/or the PTS) on AD production. MGG01 was the wild-type strain which used the PTS to import glucose. MGG was the wild-type strain with the GGS overexpressed which used both the PTS and the GGS to import the glucose, and PGG was the PTS knockout mutant with the GGS overexpressed; thus, the glucose was imported solely by the GGS

predictably increased when *ppsA* was overexpressed, indicative of the importance of balancing GAP and PYR (Fig. 2). This inspired us to explore the productivity of the PTS knockout strain, as inactivating the PTS can increase intracellular PEP and thus may increase intracellular GAP. Thus, it was hypothesized that the deletion of the PTS could potentially redirect more carbon flux from the central pathway to the synthesis of isoprenoids through the DXP pathway. In *E. coli*, the PTS is the main glucose transport system and is also responsible for carbon catabolite repression (CCR) and chemotaxis (Hernandez-Montalvo et al. 2003). Thus, deletion of the PTS is also pleiotropic and its mutant has several distinct characteristics over the wild-type strain, such as enhancement of intracellular PEP concentration, coexistence of glycolytic and gluconeogenic pathways, and the simultaneous utilization of different carbon sources (Flores et al. 2002). The PTS mutant is able to co-utilize xylose and glucose (Nichols et al. 2001) and produces higher titers of various compounds

such as recombinant proteins (Chou et al. 1994), aromatics (Flores et al. 1996), lysine (Lindner et al. 2011), L-phenylalanine (Baez-Viveros et al. 2007), and succinate (Zhang et al. 2009b). However, the uptake rate of glucose in PTS mutant is known to be low (Flores et al. 2007; Flores et al. 2002) and consequently the growth rate is exceptionally low. Therefore, in this study, the GGS was overexpressed in the PTS-deficient mutant in an attempt to increase the biomass when grown on glucose.

As predicted, the use of the GGS (consuming ATP) to increase glucose uptake in the PTS-deficient mutant was found to be beneficial for AD production (Fig. 3). Considering the benefit of using the GGS glucose uptake pathway, we hypothesized that global pathway balancing of the glucose uptake and the DXP and AD synthesis pathways may further improve the AD production. Although AD titer varied more than 190-fold within different conditions, the best construct (P2121) only improved AD titer by 15 % than that of
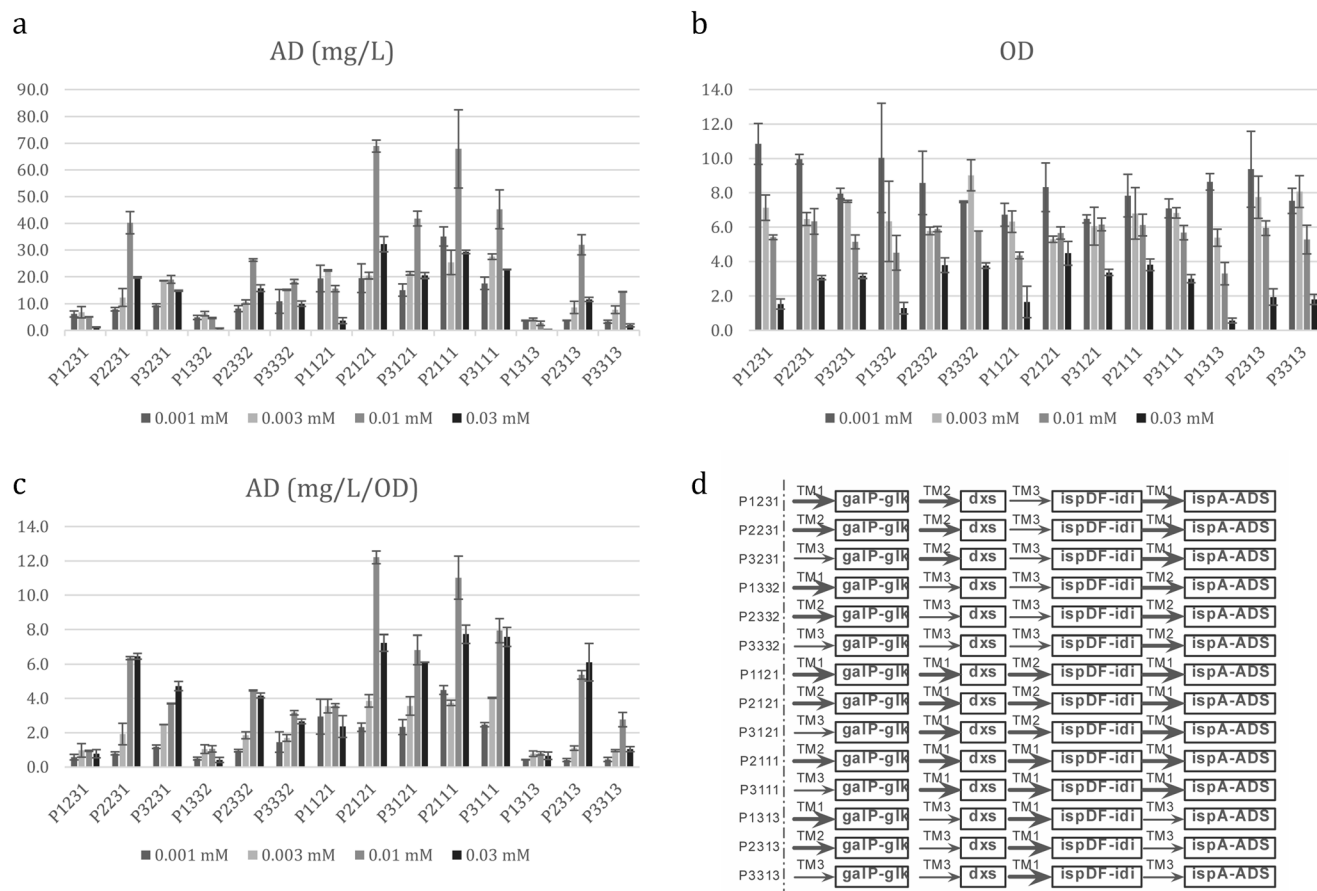

Fig. 4 Optimization of AD production by simultaneously controlling four modules. TM1, TM2, and TM3 promoters were used to control the expression of the four modules: *galP-glk*, *dxs*, *ispDF-idi*, and *ispA-ADS*. The strains were named based on their sequence of promoters, and numbers *1*, *2*, and *3* stood for TM1, TM2, and TM3 promoters, respectively, e.g., P1231 referred to a construct in which the four modules *galP-glk*, *dxs*, *ispDF-idi* and *ispA-ADS* were under the control

of TM1, TM2, TM3, and TM1 promoters, respectively. The transcriptional efficiencies of TM1, TM2, and TM3 were 92, 37, and 16 % of that of T7 promoter, respectively. **a** AD titer (mg/L). **b** Cell density of different strains. **c** Specific yield of AD (mg/L/DO). **d** Illustration of different constructs (the *thickness* of *arrows* represented the relative transcriptional efficiencies of promoters)
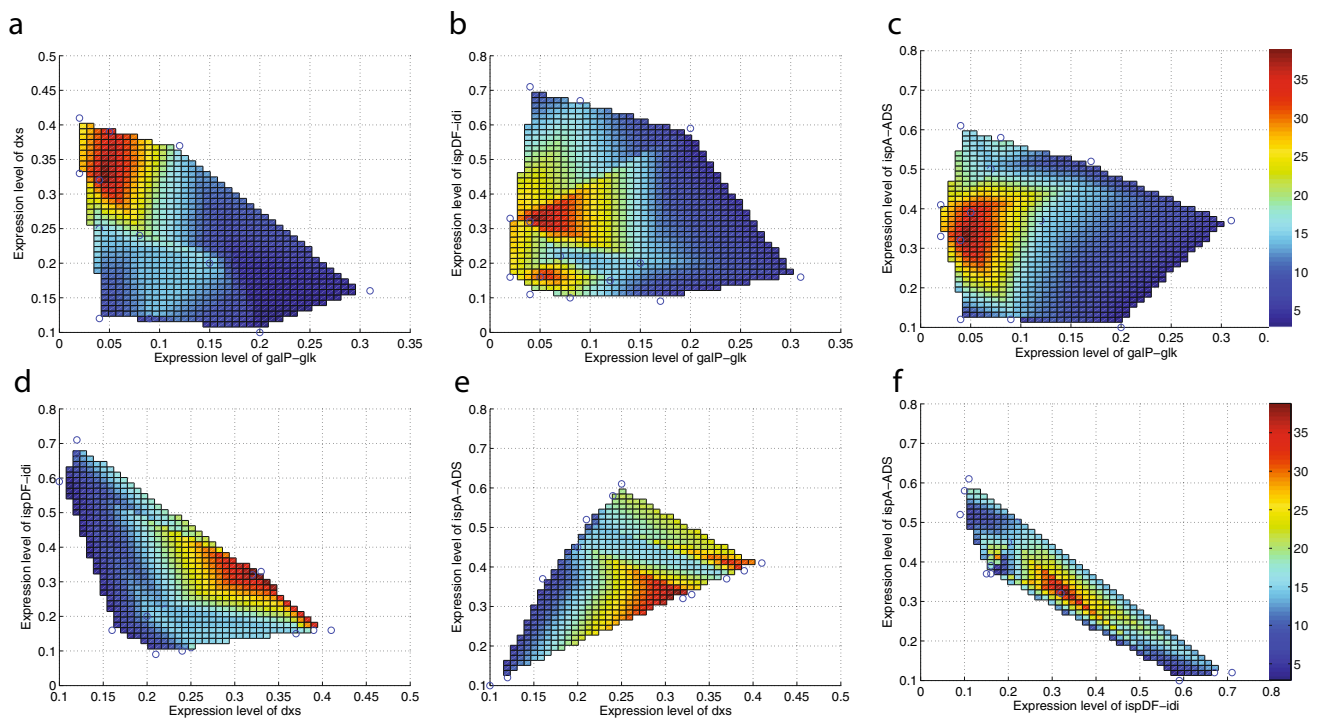
Fig. 5 Regression analysis of the four-module expression system. The regression of AD titers on the relative expression levels of every two modules (**a** *galP-glk* and *dxs*; **b** *galP-glk* and *ispDF-idi*; **c** *galP-glk* and *ispA-ADS*; **d** *dxs* and *ispDF-idi*; **e** *dxs* and *ispA-ADS*; **f** *ispDF-idi* and *ispA-ADS*). Both abscissa and ordinate referred to the relative expressional levels of their modules. Relative expression levels were calculated based on the promoter transcriptional efficiencies and plasmid copy numbers (as described in "Materials and methods"). *Blue dots* in the figures were experimental data as consistent with Fig. S1

PGG strain. And it was also observed that high expressions of the GGS improved specific yield of AD but inhibited cell growth in PGG and MGG (Fig. 3) and the 14 strains of the four-module expression system (Fig. 4). The decreased cell density at high induction level is likely to be due to the stress of high expression of galactose permease and glucose kinase (Lu et al. 2012) or sugar phosphate stress due to accumulation of glucose-6-phosphate or fructose-6-phosphate (Morita et al. 2003), and this requires further investigation. Nevertheless, this limitation can be mitigated with tight control of the

Table 4 Relative expression levels of each module in constructs P1000, P2000, and P3000

No.	Strains[a]	Relative expression levels of each module			
		galP-glk	*dxs*	*idi*	*ispA-ADS*
1	P1000	0.09	0.30	0.30	0.30
2	P2000	0.04	0.32	0.32	0.32
3	P3000	0.02	0.33	0.33	0.33

[a] The strains were named based on their sequence of promoters; number "0," "1," "2," and "3" stood for T7, TM1, TM2, and TM3 promoters, respectively (e.g., P1000 referred to a construct in which *galP-glk*, *dxs*, *idi*, and *ispA-ADS* were controlled by TM1, T7, T7, and T7 promoters, respectively)

expressions of *galP* and *glk* using the combinatorial modular expression system described herein.

To better understand the behavior of the four-module expression system, the experimental data were analyzed using a linear regression model. Based on the regression analysis (Fig. 5), the optimal ranges of expression levels for the first, second, and fourth modules were identified using the panel of constructs. However, the third module was predicted to have multiple optimal ranges (~0.15 and ~0.35). Further experiment indicated that excluding *ispDF* resulted in a slightly enhanced AD production as compared to the inclusion of *ispDF* in our conditions (Fig. S2). This finding was consistent with previous reports that enhancement in production of various metabolites requires only *dxs* and *idi* but not *ispDF* to be overexpressed or modulated (Jin and Stephanopoulos 2007; Morrone et al. 2010; Sun et al. 2014) and that the overexpression of *ispDF* operon even decreased isoprenoid production (Zou et al. 2013). The enzymes *ispD* and *ispF* were reported to be more soluble than the enzymes *dxs* and *idi* and may have higher activities (Zhou et al. 2012); hence, the overexpression of *ispDF* may be unnecessary when the pathway was optimized, and further study is required to address this issue. The fourth module was required to be expressed at higher levels as the activity of heterologous *ADS* from *Artemisia annua* is known to be poor in *E. coli* (Chen et al. 2013; Green et al. 2009). Consistent with the prediction of the

Fig. 6 Further improvement of AD production facilitated by regression analysis. **a** AD titer (mg/L). **b** Cell density of different strains. **c** Specific yield of AD (mg/L/DO). In strains P1000, P2000, and P3000, the galP-glk was placed under the control of TM1, TM2, and TM3 promoters, respectively. And all the other genes were placed under the control of T7 promoter. The strains were named based on their sequence of promoters, and numbers *0*, *1*, 2, and 3 stood for T7, TM1, TM2, and TM3 promoters, respectively, e.g., P1000 referred to a construct in which galP-glk, dxs, idi, and ispA-ADS were controlled by TM1, T7, T7, and T7 promoters, respectively
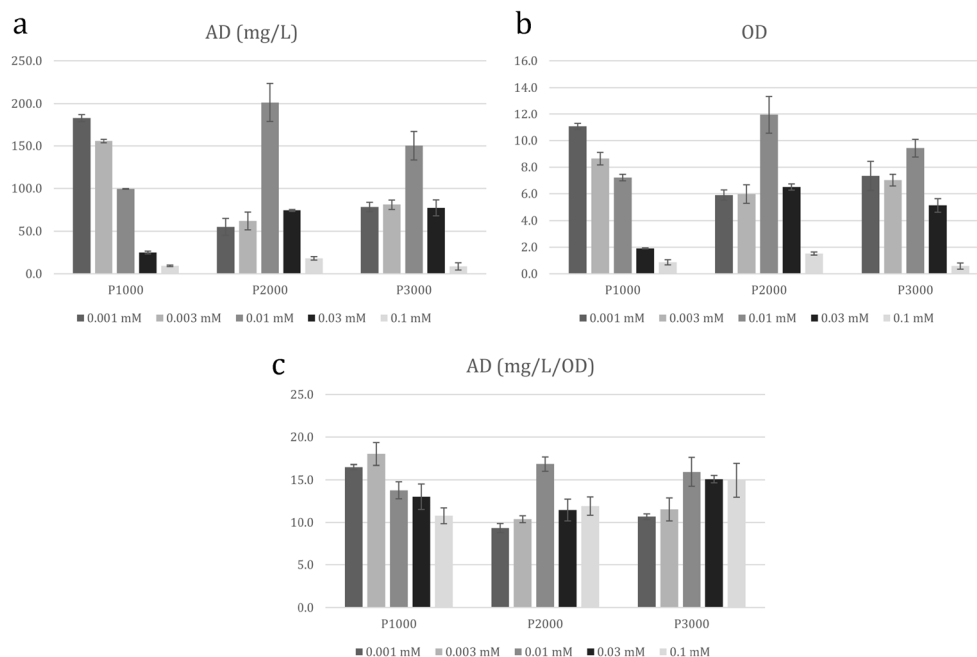
regression model, the combination of a relatively lower expression of galP-glk and higher expression of dxs and ispA-ADS enhanced AD production. These predictions were validated experimentally using strains P1000, P2000, and P3000 where AD titer increased another threefold as compared to the highest titer in the 14 constructs (overall ~19-fold increase over that of wild type). To our knowledge, this was the first study that simultaneously controlled four expression modules. These results demonstrated the successful application of EDASPO for the pathway optimization, especially when more than three modules were involved. And the major advantage of EDASPO lies in its predictability with a small portion of the full combination; thus, it significantly decreases the time cost and the work amount.

In this study, a chemically defined medium (CDM), instead of complex media (CM), was used to produce AD. It was worth noting that the 20 g/L glucose added into the CDM was more than enough (other nutrition was limiting), and there was glucose left in the media (data not shown) when the AD was harvested for all the experiments. In most cases, CDM has a relatively low production unless through extensive medium optimization. However, CDM has a few advantages. For example, CDM is more cost-effective as CM ingredients (such as yeast extract, tryptone, and peptone) are much more expensive than inorganic nitrogen salts (such as ammonium salts or ammonia). Furthermore, unlike CM whose composition varies with batches and suppliers, CDM is more reliable and scalable, making it a preferred medium for fed-batch high-cell density cultivations (Li et al. 2014). In addition, the complicated interactions among complex components in CM may interfere with the pathway optimization and kinetic study, thus making data more difficult to interpret (Zhang et al. 2009a). It

is interesting to note that the AD yield (18 mg/L/OD) using the DXP pathway was among some of the highest yields reported to date (Zhang et al. 2013; Zhou et al. 2013). However, as compared to the highest AD yield using the mevalonate (MVA) pathway in *Saccharomyces cerevisiae* (242 mg/L/OD) (Westfall et al. 2012), our yield was ~13-fold lower. Preliminary study of optimizing the MVA pathway using EDASPO resulted in a comparable AD yield of more than 200 mg/L/OD in *E. coli* (unpublished observation), demonstrating the broader utility of this approach.

References

Ajikumar PK, Xiao WH, Tyo KE, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G (2010) Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. Science 330(6000):70–4

Alper H, Miyaoku K, Stephanopoulos G (2005) Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. Nat Biotechnol 23(5): 612–6

Anthony JR, Anthony LC, Nowroozi F, Kwon G, Newman JD, Keasling JD (2009) Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. Metab Eng 11(1):13–19

Baez-Viveros JL, Flores N, Juarez K, Castillo-Espana P, Bolivar F, Gosset G (2007) Metabolic transcription analysis of engineered *Escherichia coli* strains that overproduce L-phenylalanine. Microb Cell Fact 6:30

Chatterjee R, Millard CS, Champion K, Clark DP, Donnelly MI (2001) Mutation of the ptsG gene results in increased production of

succinate in fermentation of glucose by *Escherichia coli*. Appl Environ Microbiol 67(1):148–154

Chen X, Zhang C, Zou R, Zhou K, Stephanopoulos G, Too HP (2013) Statistical experimental design guided optimization of a one-pot biphasic multienzyme total synthesis of amorpha-4,11-diene. PLoS One 8(11):e79650

Chou CH, Bennett GN, San KY (1994) Effect of modified glucose uptake using genetic engineering techniques on high-level recombinant protein production in *Escherichia coli* dense cultures. Biotechnol Bioeng 44(8):952–60

Cunningham FX Jr, Sun Z, Chamovitz D, Hirschberg J, Gantt E (1994) Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium *Synechococcus sp* strain PCC7942. Plant cell 6(8):1107–21

Datsenko KA (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97(12):6640–6645

De Anda R, Lara AR, Hernandez V, Hernandez-Montalvo V, Gosset G, Bolivar F, Ramirez OT (2006) Replacement of the glucose phospho-transferase transport system by galactose permease reduces acetate accumulation and improves process performance of *Escherichia coli* for recombinant protein production without impairment of growth rate. Metab Eng 8(3):281–90

Farmer WR, Liao JC (2000) Improving lycopene production in *Escherichia coli* by engineering metabolic control. Nat Biotechnol 18(5):533–7

Farmer WR, Liao JC (2001) Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. Biotechnol Prog 17(1):57–61

Flores N, Xiao J, Berry A, Bolivar F, Valle F (1996) Pathway engineering for the production of aromatic compounds in *Escherichia coli*. Nat Biotechnol 14(5):620–623

Flores S, Gosset G, Flores N, de Graaf AA, Bolivar F (2002) Analysis of carbon metabolism in *Escherichia coli* strains with an inactive phos-photransferase system by 13C labeling and NMR spectroscopy. Metab Eng 4(2):124–137

Flores N, Leal L, Sigala JC, de Anda R, Escalante A, Martinez A, Ramirez OT, Gosset G, Bolivar F (2007) Growth recovery on glucose under aerobic conditions of an *Escherichia coli* strain carrying a phosphoenolpyruvate:carbohydrate phosphotransferase system deletion by inactivating arcA and overexpressing the genes coding for glucokinase and galactose permease. J Mol Microbiol Biotechnol 13(1–3):105–16

Gorke B, Stulke J (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol 6(8):613–24

Green S, Squire CJ, Nieuwenhuizen NJ, Baker EN, Laing W (2009) Defining the potassium binding region in an apple terpene synthase. J Biol Chem 284(13):8661–9

Hernandez-Montalvo V, Martinez A, Hernandez-Chavez G, Bolivar F, Valle F, Gosset G (2003) Expression of galP and glk in a *Escherichia coli* PTS mutant restores glucose transport and increases glycolytic flux to fermentation products. Biotechnol Bioeng 83(6):687–94

Jin YS, Stephanopoulos G (2007) Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. Metab Eng 9(4):337–47

Jones KL, Kim S-W, Keasling JD (2000) Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metab Eng 2(4):328

Kreth J, Lengeler JW, Jahreis K (2013) Characterization of pyruvate uptake in *Escherichia coli* K-12. PLoS One 8(6):e67125

Li Z, Nimtz M, Rinas U (2014) The metabolic potential of *Escherichia coli* BL21 in defined and rich medium. Microb Cell Fact 13(1):45

Lindner SN, Seibold GM, Henrich A, Kramer R, Wendisch VF (2011) Phosphotransferase system-independent glucose utilization in *Corynebacterium glutamicum* by inositol permeases and glucoki-nases. Appl Environ Microbiol 77(11):3571–81

Lu J, Tang J, Liu Y, Zhu X, Zhang T, Zhang X (2012) Combinatorial modulation of galP and glk gene expression for improved alternative glucose utilization. Appl Microbiol Biotechnol 93(6):2455–2462

Martínez K, de Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolívar FG (2008) Coutilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phospho-transferase system. Microb Cell Fact 7(1):1

Morita T, El-Kazzaz W, Tanaka Y, Inada T, Aiba H (2003) Accumulation of glucose 6-phosphate or fructose 6-phosphate is responsible for destabilization of glucose transporter mRNA in *Escherichia coli*. J Biol Chem 278(18):15608–14

Morrone D, Lowry L, Determan MK, Hershey DM, Xu M, Peters RJ (2010) Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. Appl Microbiol Biotechnol 85(6): 1893–906

Nichols N, Dien B, Bothast R (2001) Use of catabolite repression mutants for fermentation of sugar mixtures to ethanol. Appl Microbiol Biotechnol 56(1–2):120–125

Patnaik R, Roof WD, Young RF, Liao JC (1992) Stimulation of glucose catabolism in *Escherichia coli* by a potential futile cycle. J Bacteriol 174(23):7527–32

Pfleger BF, Pitera DJ, Smolke CD, Keasling JD (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat Biotechnol 24(8):1027–1032

Pitera DJ, Paddon CJ, Newman JD, Keasling JD (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. Metab Eng 9(2):193–207

Sun T, Miao L, Li Q, Dai G, Lu F, Liu T, Zhang X, Ma Y (2014) Production of lycopene by metabolically-engineered *Escherichia coli*. Biotechnol Lett 36(7):1515–22

Too HP, ZOU R, Stephanopoulos GN (2014) Univariant extrinsic initiator control system for microbes and an in vitro assembly of large recombinant dna molecules from multiple components. USA Patent WO2014077782 A1, Nov 15, 2013

Tsuruta H, Paddon CJ, Eng D, Lenihan JR, Horning T, Anthony LC, Regentin R, Keasling JD, Renninger NS, Newman JD (2009) High-level production of amorpha-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. PLoS One 4(2): e4489

Westfall PJ, Pitera DJ, Lenihan JR, Eng D, Woolard FX, Regentin R, Horning T, Tsuruta H, Melis DJ, Owens A, Fickes S, Diola D, Benjamin KR, Keasling JD, Leavell MD, McPhee DJ, Renninger NS, Newman JD, Paddon CJ (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. Proc Natl Acad Sci U S A 109(3):E111–8

Yi J, Draths KM, Li K, Frost JW (2003) Altered glucose transport and shikimate pathway product yields in *E. coli*. Biotechnol Prog 19(5): 1450–1459

Yuan LZ, Rouvière PE, Larossa RA, Suh W (2006) Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. Metab Eng 8(1):79–90

Zhang G, Mills DA, Block DE (2009a) Development of chemically defined media supporting high-cell-density growth of *lactococci*, *enterococci*, and *streptococci*. Appl Environ Microbiol 75(4): 1080–7

Zhang X, Jantama K, Moore JC, Jarboe LR, Shanmugam KT, Ingram LO (2009b) Metabolic evolution of energy-conserving pathways for succinate production in *Escherichia coli*. Proc Natl Acad Sci U S A 106(48):20180–5

Zhang C, Chen X, Zou R, Zhou K, Stephanopoulos G, Too HP (2013) Combining genotype improvement and statistical media optimization for isoprenoid production in *E. coli*. PLoS One 8(10):e75164

Zhou K, Zou R, Stephanopoulos G, Too HP (2012) Enhancing solubility of deoxyxylulose phosphate pathway enzymes for microbial isoprenoid production. Microb Cell Fact 11:148

Zhou K, Zou R, Zhang C, Stephanopoulos G, Too HP (2013) Optimization of amorphadiene synthesis in *bacillus subtilis* via transcriptional, translational, and media modulation. Biotechnol Bioeng 110(9):2556–61

Zou R, Zhou K, Stephanopoulos G, Too HP (2013) Combinatorial engineering of 1-deoxy-D-xylulose 5-phosphate pathway using crosslapping in vitro assembly (CLIVA) method. PLoS One 8(11):e79557