Metabolic Engineering 13 (2011) 194–203

Contents lists available at ScienceDirect

Metabolic Engineering

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

Targeted proteomics for metabolic pathway optimization: Application to terpene production

Alyssa M. Redding-Johanson, Tanveer S. Batth, Rossana Chan, Rachel Krupa, Heather L. Szmidt, Paul D. Adams, Jay D. Keasling, Taek Soon Lee, Aindrila Mukhopadhyay, Christopher J. Petzold *

Lawrence Berkeley National Laboratory, Physical Biosciences Division, Joint BioEnergy Institute, 5885 Hollis Street, 4th Floor, Emeryville, CA 94608, USA

ARTICLE INFO

Article history:
Received 25 August 2010
Received in revised form
15 November 2010
Accepted 20 December 2010
Available online 5 January 2011

Keywords:
Targeted proteomics
Selected-Reaction Monitoring (SRM)
Mevalonate pathway
Amorphadiene production
Metabolic pathway optimization
E. coli

ABSTRACT

Successful metabolic engineering relies on methodologies that aid assembly and optimization of novel pathways in microbes. Many different factors may contribute to pathway performance, and problems due to mRNA abundance, protein abundance, or enzymatic activity may not be evident by monitoring product titers. To this end, synthetic biologists and metabolic engineers utilize a variety of analytical methods to identify the parts of the pathway that limit production. In this study, targeted proteomics, via selected-reaction monitoring (SRM) mass spectrometry, was used to measure protein levels in *Escherichia coli* strains engineered to produce the sesquiterpene, amorpha-4,11-diene. From this analysis, two mevalonate pathway proteins, mevalonate kinase (MK) and phosphomevalonate kinase (PMK) from *Saccharomyces cerevisiae*, were identified as potential bottlenecks. Codon-optimization of the genes encoding MK and PMK and expression from a stronger promoter led to significantly improved MK and PMK protein levels and over three-fold improved final amorpha-4,11-diene titer (>500 mg/L).

1. Introduction

The goal of many metabolic engineering efforts is to establish cheap, stable supplies of compounds ranging from therapeutic drugs to fuels (Herrero et al., 2008; Fortman et al., 2008; Huang et al., 2008; Munoz-Bertomeu et al., 2008; Gosset, 2009; Celinska, 2010; Colling et al., 2010). Inherent to successful metabolic engineering, regardless of the desired product, are methodologies that aid assembly and optimization of biosynthetic pathways in microbes to minimize cost and maximize productivity. To this end, synthetic biologists and metabolic engineers are developing vast libraries of well-characterized, standardized biological parts that can be easily assembled into larger networks (Keasling, 2008; Endy, 2005; Anderson et al., 2010; Martin et al., 2009). Additionally, a variety of directed evolution techniques (Bloom et al., 2005; Alper et al., 2005; Wang et al., 2009) have been successfully applied to overcome pathway imbalances and improve production; however, they rely heavily on rapid screening methods that are not universally applicable (*e.g.*, GFP, lycopene assays). Alternatively, metabolite production has been improved through rational approaches such as gene titration experiments (Pitera et al., 2007; Luetke-Eversloh and Stephanopoulos, 2008; Anthony et al., 2009) or by co-localizing pathway proteins via synthetic scaffolds (Moon et al., 2010; Dueber et al., 2009). Ultimately, the success (or failure) of any optimization strategy is measured by the efficiency of the pathway as reported by final product titer or productivity.

Many different factors may contribute to pathway performance. Problems such as those due to consumption of intermediates by native pathways, mRNA abundance, protein abundance, or enzymatic activity may not be evident by only monitoring product titers. At the metabolite level, monitoring all pathway intermediates aids identification of bottlenecks, where further engineering can improve product titer. Yet, pathways that divert intermediates to native processes at the cost of final product formation may disguise the location of a bottleneck. Furthermore, developing assays for every intermediate is often challenging due to the lack of available standards, and intermediates that degrade rapidly or that are isomers (*e.g.*, IPP/DMAPP, leucine/isoleucine). These complications often leave the engineer guessing at the metabolite levels in parts of the pathway.

In contrast, mRNA analysis can be used to assay entire genomes in a single experiment and is frequently used to verify that the host organism is transcribing a heterologous pathway of interest. Genome-wide microarray analysis also reveals a variety of stress response factors due to the metabolic changes arising from non-native pathways in the cell. However, the transcript level is, at best, a surrogate for the level of the encoded protein (Maier et al., 2009). Factors such as mRNA secondary structure, ribosome binding site strength, codon usage, and transcript

* Corresponding author. Fax: +1 510 486 4252.
E-mail address: cjpetzold@lbl.gov (C.J. Petzold).

1096-7176/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2010.12.005

ribosome occupancy, in addition to the half-life of the protein produced, determine the level of a given protein in the cell at any time. Characteristics of engineered pathways such as the promoter strength, the position of genes in the operon, and spacing between genetic elements also profoundly impact the amount of protein produced. As a result, achieving optimal protein levels in engineered pathways is challenging and often not accomplished.

In multi-gene pathways, careful tuning of protein production is needed to maximize product titers (Alper et al., 2005; Pitera et al., 2007; Anthony et al., 2009). Underscoring the impact of protein abundance on metabolite production, it was recently shown that the intracellular level of PhaB, a protein involved in poly-3-hydroxybutyrate (PHB) production, is more important than precursor availability (Tyo et al., 2010). However, maximizing protein production is not a universal solution to achieving high titers, because heterologous gene expression can cause a significant burden to the cell (Stoebel et al., 2008), especially when the metabolite being produced is toxic to the cell. Fortunately, many methods for modulating protein production are available in *Escherichia coli*. For example, within a given synthetic operon the protein expression level can be varied by changing the ribosome binding site affinity (Salis et al., 2009), the stability of the mRNA intergenic region (Pfleger et al., 2006), or the strength of the promoter driving expression of the gene of interest (Solem and Jensen, 2002; De Mey et al., 2007). Tightly controlled promoter systems are often desirable because leaky expression systems can be detrimental to cell growth when the protein is toxic to the cell (Makrides, 1996; Lee and Keasling, 2008). Successful optimization of pathway enzyme levels benefits greatly from direct monitoring and quantifying proteins in the engineered system.

One challenge facing protein analysis is the lack of high-throughput techniques similar to those available for nucleic acids. Separating a protein of interest from a cellular proteome is difficult and labor intensive because proteins have similar physical attributes (*e.g.*, MW, pI) and are produced at vastly different levels in the cell. Traditionally, selective protein detection has been accomplished via immunoblot analysis. Immunoblots are fast, convenient, and, for the same protein, easily multiplexed. However, custom antibodies must be obtained for every pathway protein when assaying many proteins simultaneously and obtaining accurate quantitative information can be challenging. This can substantially increase the development time and experimental costs (Arnott et al., 2002).

Advances in protein quantification via liquid chromatography-mass spectrometry (LC-MS and LC-MS/MS) have expanded the number of proteins that can be quantified in a cell lysate by several orders of magnitude. Isotopic labeling of the proteins can be used for relative quantification of a few samples, which is typically achieved by differentially growing cells in media containing either heavy- or light-isotope labeled metabolites or by using functional group-specific chemical labels (*e.g.*, iTRAQ). Alternately, mass spectrometric signal intensity can be used for comparison of many samples in so called "label-free" quantification methods. Label-free methods quantify protein levels between samples by comparison of the number of peptides confidently detected (Ishihama et al., 2005), counting the number of MS/MS spectra collected for a given peptide (Zhang et al., 2006), or from raw spectral intensities (Wang et al., 2003). Alternatively, a targeted proteomics approach, via selected-reaction monitoring (SRM) mass spectrometry, can be used to quantify many specific proteins in a sample. The SRM technique has been used for small molecule drug metabolism and pharmacokinetic (DMPK) studies since the early 1990s (Baillie, 2008), but has only recently been applied to peptides. Coupled to LC separation, SRM-based analysis provides rapid detection and quantification of target molecules based on the peptide mass and peptide-specific MS/MS fragment ions. By using two points of mass selection, the background signal and noise are greatly reduced. Furthermore, a significant increase in sensitivity of the target molecules is typically observed, because the entire mass range is not scanned. For peptide analysis, SRM analysis has predominantly been used for the development of diagnostic disease biomarkers in plasma and tissue samples (Anderson and Hunter, 2006; Kuhn et al., 2004; Pan et al., 2009). Recently, Aebersold and co-workers designed SRM transitions for *Saccharomyces cerevisiae* proteins spanning the protein copy number range from the most highly abundant to fewer than 50 copies per cell (Picotti et al., 2009).

The work described herein details the application of targeted proteomics via SRM mass spectrometry for metabolic engineering applications. For initial characterization, we used SRM-based methods to examine the level of red fluorescent protein (RFP) produced from plasmids with a variety of promoters and origins of replication. Secondly, SRM mass spectrometry was used to measure protein levels in *E. coli* engineered to produce the sesquiterpene, amorpha-4,11-diene. *E. coli* containing a high flux mevalonate pathway has the potential to provide a vast range of isoprenoid-based bulk and high value compounds that are typically obtained from petrochemical or plant sources. Amorpha-4,11-diene titers were measured to investigate the impact of vector modification on bottlenecks and to refine hypotheses to maximize production. In particular, these studies revealed that levels of two pathway proteins, mevalonate kinase (MK) and phosphomevalonate kinase (PMK), were particularly low. To overcome these bottlenecks, several different strategies were employed to increase MK and PMK and balance overall protein levels.

2. Material and methods

All chemicals, solvents and media components were purchased and used without modification from Sigma-Aldrich (St. Louis, MO), Fisher Scientific (Pittsburgh, PA), or VWR (West Chester, PA) unless otherwise noted. *E. coli* strains DH10B (Invitrogen, Carlsbad, CA) and DH1 (ATCC) were used for plasmid construction and analytical experiments. For targeted proteomics experiments, mass spectrometric-grade trypsin was obtained from Sigma-Aldrich and prepared according to manufacturer's instructions.

2.1. Strain construction

E. coli DH10B was used as the host for cloning and fluorescent protein production experiments, while *E. coli* DH1 was used as the host for amorpha-4,11-diene production. The original amorpha-4,11-diene-producing system consists of three plasmids: pMevT, pMBIS, and pADS (Martin et al., 2003). The pMevT plasmid harbors three genes for the conversion of acetyl-CoA to mevalonate: acetoacetyl-CoA synthase (AtoB) from *E. coli*, HMG-CoA synthase (HMGS) and HMG-CoA reductase (HMGR) from *S. cerevisiae*. HMGR was truncated at the N-terminus to remove its regulatory component and improve solubility. The pMBIS plasmid harbors the genes encoding five proteins for the conversion of mevalonate to farnesyl diphosphate (FPP): three genes from *S. cerevisiae*, mevalonate kinase (MK), phosphomevalonate kinase (PMK), and phosphomevalonate decarboxylase (PMD), and two genes from *E. coli*, isopentenyl diphosphate isomerase (Idi) and farnesyl diphosphate synthase (IspA). The pADS plasmid contains the gene encoding amorpha-4,11-diene synthase (ADS) from *Artemisia annua*. For this study, all plasmids denoted pBb were prepared as part of the BglBrick standard collection of expression

Table 1
List of *E. coli* base strains and plasmids used in this study.

	Description	Reference
Strain		
DH1	F^- *endA1 hsdR17 (rk−, mk+) supE44 thi-1* λ^- *recA1 gyrA96 relA1*	Hanahan, 1983
DH10B	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara,leu*)7697 *galU galK* λ^- *rpsL nupG*	Invitrogen
ΔAcrAB	DH1 derivative with Δ*acrA* and Δ*acrB*, Amp^r	This study
Plasmid		
pBbA1k-RFP	p15A, Kan^r, P_{trc} , mRFP	This study
pBbA2k-RFP	p15A, Kan^r, P_{tet} , mRFP	This study
pBbA5k-RFP	p15A, Kan^r, P_{lacUV5} , mRFP	This study
pBbA8k-RFP	p15A, Kan^r, P_{BAD}, mRFP	This study
pBbS5k-RFP	pSC101, Kan^r, P_{lacUV5} , mRFP	This study
pBbE5k-RFP	ColE1, Kan^r, P_{lacUV5} , mRFP	This study
pMevT	p15A, Cm^r, P_{lac} , AtoB-HMGS-HMGR	Martin et al., 2003
pMBIS	pBBR1, Tet^r, P_{lac} , MK-PMK-PMD-Idi-IspA	Martin et al., 2003
pADS	pMB1, Amp^r, P_{trc} , ADS	Martin et al., 2003
pBbA5c-MevT-MBIS	p15A, Cm^r, P_{lacUV5} , AtoB-HMGS-HMGR-MK-PMK-PMD-Idi-IspA	This study
pBbA5c-MevT-MBIS(CO)	pBbA5c-MevT-MBIS derivative containing codon-optimized MK and PMK	This study
pBbA5c-MevT-T1-MBIS	p15A, Cm^r, P_{lacUV5}-AtoB-HMGS-HMGR-term-P_{trc}-MK-PMK-PMD-Idi-IspA	This study
pBbA5c-MevT-T1-MBIS(CO)	pBbA5c-MevT-T1-MBIS derivative containing codon-optimized MK and PMK	This study

vectors (Anderson et al., 2010). The BglBrick cloning strategy utilizes four restriction sites, which are *EcoRI, BglII, BamHI and XhoI,* taking advantage of the fact that *BglII* and *BamHI* are compatible (Anderson et al., 2010). Briefly, these plasmids were created by ligating three individual cassettes, one containing a replication origin, one containing an antibiotic marker, and one containing an expression cassette comprised of a repressor, a promoter, and the BglBricked mRFP. Individual genes were cloned to include the BglBrick compatible restriction sites at the ends of the gene, and then were digested and ligated together to form larger gene expression cassettes. Gene expression cassettes could then be digested to give a single piece that was ligated with a digested BglBrick-compatible plasmid backbone. Plasmids used in this study are listed in Table 1.

2.2. Growth conditions and amorpha-4,11-diene analysis

E. coli strains harboring RFP-containing plasmids were grown in M9 minimal medium (Neidhardt et al., 1974) without MOPS. The medium was supplemented with 0.8% glucose and 50 μg/mL kanamycin. Starter cultures were grown overnight at 37 °C, and all samples were shaken in rotary shakers at 200 rpm (Kühner, Basel, Switzerland). Fresh cultures were inoculated at an optical density (OD_{600}) of 0.1, and allowed to continue growth at 37 °C. Samples were induced at an OD_{600} of 0.3 with a final concentration of 25–500 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) or 25 nM anhydrotetracycline (aTc), depending on the system being studied, and were sampled at several time points post-induction for proteomics experiments. For samples harboring amorpha-4,11-diene production plasmids, the strains were grown in either M9 media or EZ Rich Defined medium (Teknova, Hollister, CA). EZ-Rich medium is based on the M9 recipe but includes amino acids, thus minimizing cellular adaptation time, so it was selected for production experiments. The medium was supplemented with 1% glucose and appropriate antibiotics (Table 1). Production cultures were inoculated at an optical density (OD_{600}) of 0.1 and placed in shakers at 30 °C, as the lower temperature improved amorphadiene production. Samples were induced at an OD_{600} of 0.3 with a final concentration of 100 μM IPTG and were sampled at 4 h for mRNA or 24 h for both protein and amorpha-4,11-diene measurements. OD measurements were conducted on a UV–vis spectrophotometer (Beckman, Fullerton, CA) operating at 600 nm. Fluorescence measurements were taken in a Synergy 4 Microplate Reader (BioTek Instruments) using a 553 nm emission and a 583 nm excitation frequency. For amorpha-4,11-diene production assays, a dodecane layer (20% by volume) was added to the culture upon induction to trap amorpha-4,11-diene and was sampled and analyzed using gas-chromatography mass-spectrometry (GC–MS) as reported previously (Anthony et al., 2009).

2.3. Proteomics sample preparation

For proteomic analysis, cells were pelleted by centrifugation at 8000*g* (4 °C) for 5 min, the supernatant was decanted, and cells were flash frozen in liquid nitrogen and stored at −80 °C until use. The cells were thawed on ice and vortexed to homogenize the solution. An aliquot (100 μL) of the concentrated cell pellet was placed into a fresh 1.7 mL tube, and protein extraction was done using chloroform/methanol precipitation as described previously (Wessel and Flugge, 1984), with the exception that protein pellets were allowed to dry in a vacuum concentrator (ThermoSavant) for 30–60 min. Ammonium bicarbonate (500 μL of a 50 mM solution in 10% acetonitrile) was added, and the protein pellets were resuspended by vortexing and sonication in a sonic bath. Protein concentration was determined with the DC Protein reagent (BioRad, Hercules, CA) according to manufacturer's instructions. Protein (50 μg) was taken from each sample and diluted to a final concentration of 1 mg/mL in 50 mM ammonium bicarbonate. To reduce the disulfide bonds in the samples, 5 mM tris(2-carboxyethyl)phosphine (TCEP) was added to the solution and incubated at room temperature for 30 min. The proteins were alkylated by adding iodoacetic acid, prepared in 100 mM NaOH, to a final concentration of 200 mM and the reaction proceeded for 30 min in the dark. After alkylation, trypsin (1 μg/μL) was added to the sample to a final concentration of 1:50 (trypsin:-sample) and incubated at 37 °C overnight. Standard proteins were dissolved in 50 mM ammonium bicarbonate and subjected to the same reduction, alkylation, and trypsinization conditions as cellular lysates. Prior to analysis, bovine serum albumin digests were spiked into the cellular lysate at a concentration of 17 fmol/μL to serve as an internal standard.

2.4. Targeted proteomics analysis

Peptide samples were analyzed on a LC-MS/MS system consisting of an Eksigent TEMPO nanoLC-2D coupled to an Applied

Biosystems (Foster City, CA) 4000Q-Trap mass spectrometer running with Analyst™ 1.5 (Applied Biosystems) operating in SRM mode. The peptide samples were loaded onto a Pepmap100 μ-guard column (Dionex-LC Packings, Sunnyvale, CA) and were washed for 20 min using a 15 μL/min flow rate and 100% buffer A (98% (v/v) acetonitrile, 0.1% (v/v) formic acid, balance H_2O) prior to injection onto a Dionex Pepmap100 analytical column (75 μm i.d., 150 mm length, 100 Å, and 3 μm) with a 45-min method at a flow rate of 300 nL/min. The pressure in the column was allowed to equilibrate by running 95% buffer A with 5% buffer B (98% (v/v) acetonitrile, 0.1% (v/v) formic acid, balance H_2O) for two min. Peptide separation occurred during a 15 min ramp to 70% buffer A. The column was washed by a ramping the column to 20% (v/v) buffer A over 3 min and maintaining that composition for 10 min. Column re-equilibration occurred by ramping back to 95% A in 2 min and maintaining that composition for 13 min. The analytical column was interfaced to the mass spectrometer by using the MicroIonSpray head with nebulizing gas and a 10-μm Picotip emitter (New Objective, Woburn, MA) operating in the positive mode (2250–2400 V).

Peptide transitions for each protein were selected after data were collected operating in Information Dependent Acquisition (IDA) mode and consisting of a SRM survey scan and two product ion (MS/MS) scans. To ensure specificity of targeted peptides, the Q3 ions were selected by choosing from the product ion scan high-abundant *y*-series fragments above a molecular weight of 600 *m/z*. Three or four SRM transitions were selected for each individual targeted protein. Peptide transitions were verified by acquiring full MS/MS scans for each selected SRM to reconfirm the target peptide sequence. The retention time of the peptide was used for added confirmation. MS/MS spectra were collected for 2 s over a mass range of 100–1600 *m/z* with Q1 resolution set to low and rolling collision energy. MS/MS spectra were processed with the Paragon algorithm (ProteinPilot 2.0, Applied Biosystems) and searched against a database consisting of all the ORFs from *E. coli* K12 obtained from MicrobesOnline (www.microbesonline.org) (Dehal et al., 2009). Additionally, the database contained all exogenous proteins of interest, and common contaminants such as human keratins, trypsin, and commonly used standard proteins (β-galactosidase, bovine serum albumin, cytochrome c, etc.). The data were searched with the following settings: protease digestion with trypsin, cysteine blocking with iodoacetic acid, confidence level was set to 95% (ProtScore = 1.3), and the Paragon algorithm was set to rapid, so as to consider only fully tryptic peptides with no post-translational modifications. Once all SRM transitions were optimized and validated, an SRM-only scan method was used.

2.5. SRM protein quantitative analysis

MultiQuant™ version 1.2 software (Applied Biosystems) was used to determine the peak area for all SRM transitions. Automatic peak integration using the Intelliquant algorithm was used with peak smoothing set to 3. As three to four transitions were selected for every protein, peak areas were summed for each protein to give a total peak area. Absolute intensities between samples cannot be directly compared due to differences in instrument and column performance; consequently, total peak area for each protein was normalized to the total peak area for the internal standard, BSA, to give a normalized total peak area for each protein, similar to a method reported earlier (Silva et al., 2006). For experiments involving the three plasmid (3p) or two plasmid (2p) system producing amorpha-4,11-diene, normalization was done using the chloramphenicol resistance protein. Each normalized total peak area was averaged for three replicate analyses, and the standard deviation was calculated.

2.6. Transcript quantitative analysis

Samples for mRNA collection were allowed to grow 4 h following the addition of inducer, which was during late-exponential phase growth. The equivalent of 5 mL of cells at an OD_{600} of one were harvested in a microfuge tube, flash frozen and stored at −80 °C. To extract mRNA, the cell pellet was resuspended directly in buffer RLT from the RNeasy kit (Qiagen) and cold, 0.5 mm glass beads (Biospec, Bartlesville, OK) were added. The mixture was subject to bead beating for 30 s followed by resting on ice for 1 min for 3 cycles. The remaining protocol was followed as directed in the RNeasy kit. Purified mRNA was treated with Turbo DNA-*free*™ kit to remove genomic DNA contamination (Invitrogen), and was then transcribed into cDNA using the SuperScript III First Strand Synthesis System (Invitrogen) and random hexamer primers, as directed. Primers were designed to amplify 70 bp at the 5′- or 3′-end of the gene. Real-time PCR (RT-PCR) reactions were performed in a StepOnePlus RT-PCR System (Applied Biosystems) using the PerfeCTa SYBR Green SuperMix, ROX (Quanta Biosciences) with a total reaction volume of 25 μL per well. For all RT-PCR reactions, *arcA* was used as a control to allow for normalization between samples. The *arcA* gene was selected from a compendium of microarray analyses (MicrobesOnline) because *arcA* was consistently observed and had little variability under numerous conditions.

3. Results and discussion

3.1. Promoter characterization

To test the SRM quantification method, we constructed several plasmids having identical plasmid backbones, having the promoter as the only variable element driving the expression of red fluorescent protein (RFP). Four promoters commonly used for heterologous protein production in *E. coli*—P_{lacUV5}, P_{trc}, P_{tet}, and P_{BAD}—were selected for study. The P_{lacUV5} and P_{trc} systems were induced with IPTG, the P_{BAD} construct was induced with arabinose, and the P_{tet} construct was induced with aTc. The *trc* promoter provided the highest level of RFP as measured by SRM (Fig. 1) and total

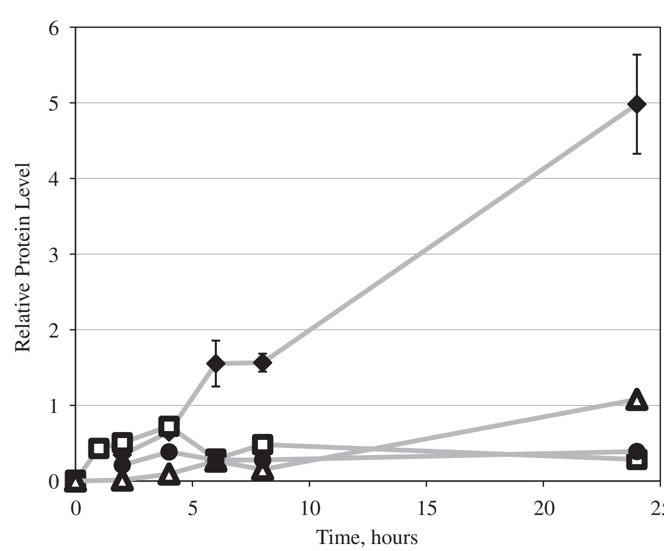

Fig. 1. Relative protein data for RFP from various promoters. *E. coli* strain DH10B harbored pBbA1k-RFP, pBbA2k-RFP, pBbA5k-RFP, and pBbA8k-RFP which contain the P_{trc} (◆), P_{tet} (□), P_{lacUV5} (Δ), and P_{BAD} (●) promoters, respectively. Time points are shown by symbols and straight lines were used to connect points. Results are the average of three technical replicates, with error bars representing the standard deviation from the mean. *Note*: the error bars for P_{BAD}, P_{tet} and P_{lacUV5} fall within the symbol, and are consequently not visible.

fluorescence measurements (data not shown), while the *lac*UV5 promoter produced the lowest level of RFP. The *lac*UV5 promoter also showed the slowest expression dynamics of the promoter systems tested, requiring six hours before substantial protein was detected (Fig. 1). In contrast, the *tet* promoter produced significant amounts of RFP after one hour. Upon induction, the strains with plasmids under control of *trc* and *lac*UV5 promoters demonstrated a consistent increase in RFP levels over the course of the experiment. However, the *tet* promoter system behaved inconsistently; it produced the maximum level of RFP at four hours post-induction and showed decreasing levels thereafter despite growth comparable to the other systems. Analogously, fluorescence measurements of RFP produced from the P_{tet} system indicated that fluorescence levels plateau in this system at approximately four hours, which is consistent with decreasing production per cell. This observation could be explained by the presence of a sub-population of cells that had been cured of the RFP plasmid. Consequently, cells were examined under bright field and fluorescence microscopy, which confirmed that all of the cells in the population were producing RFP after 24 h.

To establish whether the profile observed for P_{tet}–dependent RFP production resulted from decreased levels of aTc, an additional 25 nM aTc was added to the culture five hours after the first induction point (Fig. 2). As shown in Fig. 2, the resulting levels of RFP at 20 h were approximately four-fold higher than the case, where aTc was added solely at T_0, suggesting that the aTc concentration in the cell decreases with time. Based upon reports that the native *E. coli* AcrAB-TolC efflux pump system exports aTc (Le et al., 2006), the P_{tet} RFP plasmid was transformed into a Δ*acrAB* strain. The Δ*acrAB* strain produced a similar fluorescence profile and RFP protein levels as compared to the wild-type strain (data not shown). It is still possible that other transporters are involved; however, no other pump has currently been reported to specifically export anhydrous tetracycline. The chemical breakdown of aTc into epianhydrotetracycline has also been documented, although the kinetics of degradation was only provided for a low pH and high temperature system (Palmer et al., 2010). It seems more likely that aTc breakdown is occurring in these cells, whether through metabolic activity or chemical degradation. Overall, the promoter study demonstrated that the SRM method reveals dynamic behavior of protein production from various promoters, which track well with fluorescence measurements.

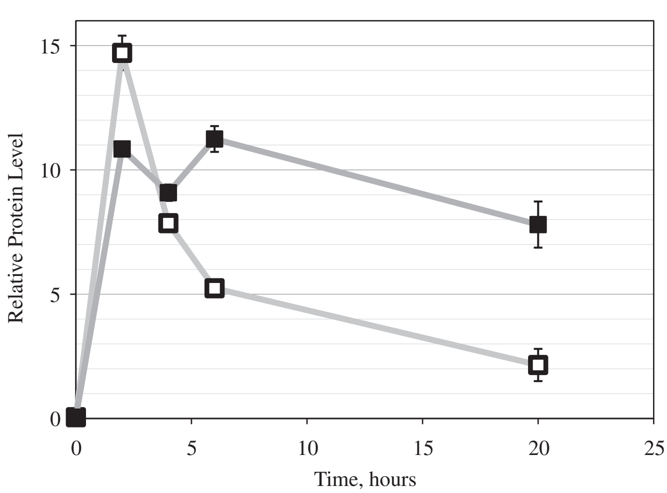

Fig. 2. RFP production from P_{tet} with and without additional aTc. *E. coli* strain DH10B harbored pBbA2k-RFP. In one case, 25 nM aTc was added at T0 (□), as in Fig. 1. In the second case, an additional 25 nM aTc added at 5 h (■). Time points are shown by symbols and straight lines were used to connect points. Results are the average of three technical replicates with error bars representing the standard deviation from the mean.

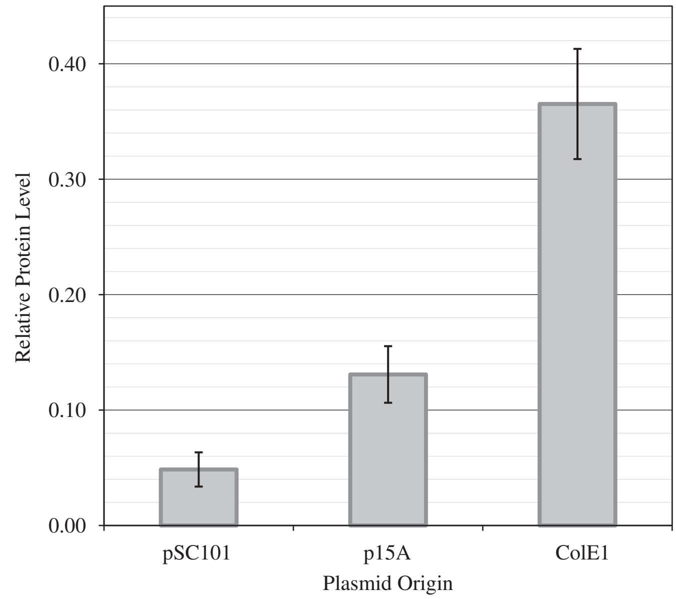

Fig. 3. RFP production study utilizing plasmids containing three different origins of replication. *E. coli* DH10B was harboring plasmids pBbS5k-RFP (pSC101 origin), pBbA5k-RFP (p15A origin), or pBbE5k-RFP (ColE1 origin). All plasmids contained P_{lacUV5} and *kan*^{r}. Results are the average of three technical replicates with error bars representing the standard deviation from the mean.

3.2. Origin of replication comparison

Another important strategy used to vary the expression level of a heterologous protein *in vivo* is to change the plasmid copy number, which is controlled by the origin of replication, and can vary from as few as 2–3 copies (*e.g.* pSC101*) to greater than 500 copies per cell (*e.g.* pUC) (Kues and Stahl, 1989; Sambrook, 1989; Lutz and Bujard, 1997; Hiszczynska-Sawicka and Kur, 1997; Grabherr and Bayer, 2002). Here, the SRM method was used to determine the levels of protein produced by varying the plasmid copy number via three distinct origins of replication: pSC101, p15A, and ColE1. As before, the remainder of the plasmid backbone remained constant, containing the RFP gene under control of P_{lacUV5}. To test the protein levels resulting from these different plasmids, samples were prepared as described above, and the same amount of total protein was analyzed in all cases (Fig. 3).

As expected, SRM analysis of protein production levels indicates that the pSC101 plasmid produces the lowest level of protein, consistent with reports in the literature (Lutz and Bujard, 1997; Kues and Stahl, 1989). RFP protein production from the plasmid containing the p15A origin was approximately 2.7 times that of the pSC101 system, and the ColE1 system produced approximately 7.5 fold more protein compared to pSC101. These relative protein ratios are very similar to those reported by using a protein activity-based screen using the luciferase reporter (Lutz and Bujard, 1997), and are equivalent within the experimental error of these two independent techniques. Based upon the results obtained using the RFP plasmids that varied in promoter strength and plasmid copy number, the protein signal and dynamic range observed indicate that the SRM workflow provides a sensitive protocol to identify biologically relevant changes in protein levels.

3.3. Application to amorpha-4,11-diene production strains

Based on these preliminary experiments we applied the SRM technique to study *E. coli* that had been engineered to produce high levels of the isoprenoid, amorpha-4,11-diene. Martin et al.

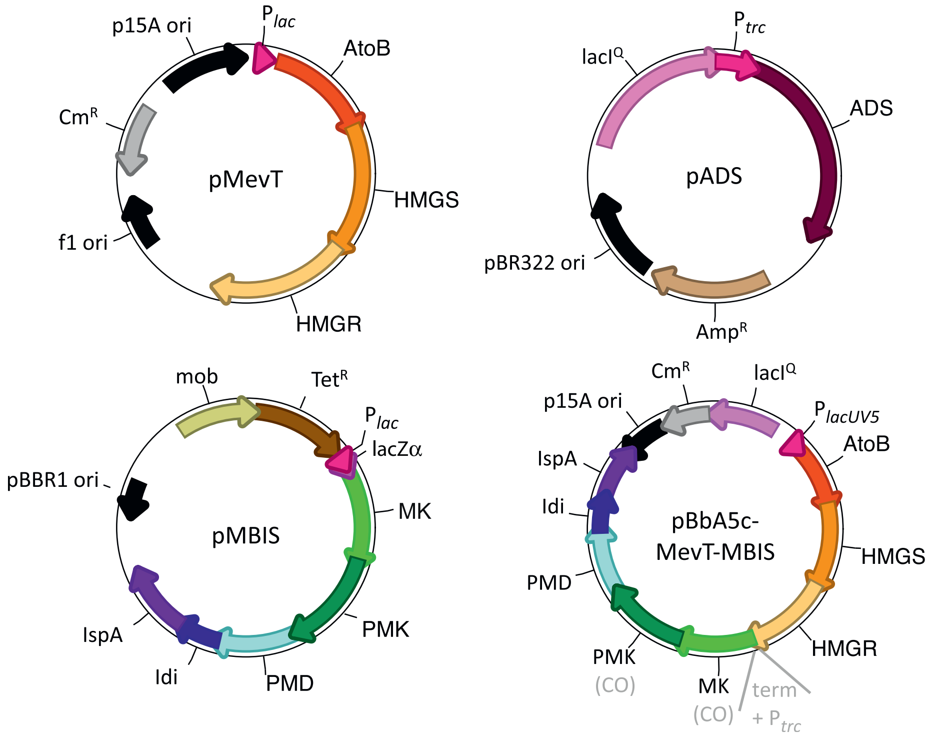

Fig. 4. Plasmid maps for the three- and two-plasmid systems. The plasmids pMevT, pMBIS, and pADS were transformed to generate the 3p strain, while the plasmids pBbA5c-MevT-MBIS and pADS were transformed to generate the 2p strain (Table 2). Modifications made to the pBbA5c-MevT-MBIS plasmid to generate plasmid variants are shown in gray.

(2003) engineered the original strain by incorporating part of the mevalonate biosynthesis pathway from *S. cerevisiae* (HMGS, HMGR, MK, PMK, and PMD), the amorpha-4,11-diene synthase gene from *A. annua* (ADS), and additional copies of several native *E. coli* genes (*atoB*, *idi*, and *ispA*) into *E. coli* DH1. The nine-enzyme pathway was constructed on three plasmids called pMevT, pMBIS, and pADS (Fig. 4); the strain harboring all three plasmids is referred to as 3p (Table 2). Plasmids are known to affect the cell in complex ways (Ricci and Hernández, 2000; Ow et al., 2009). Maintaining multiple plasmids has been hypothesized to increase the metabolic burden on the cell in terms of replication as well as the total number of antibiotic resistance markers the cell must consequently produce (Rozkov et al., 2004). Thus, expressing the pathway genes from fewer plasmids may yield higher amorpha-4,11-diene production. To this end, the genes constituting the entire mevalonate pathway (*atoB* to *ispA*) were condensed onto a single plasmid, pBbA5c-MevT-MBIS (Fig. 4), and a strain harboring this plasmid along with pADS was constructed, referred to as 2p (Table 2). GC–MS analysis showed that the two-plasmid system, 2p, produced virtually no amorpha-4,11-diene in the absence of inducer (0 μM IPTG) and production increased in an inducer-dependent manner. In contrast, the 3p strain produced nearly 40% of its maximal production without inducer, indicating that the proteins in 3p are poorly regulated (Fig. 5). This is likely due to the fact that the plasmids in the 3p strain use the P_{lac} promoter while the 2p system uses P_{lacUV5} (Stefano et al., 1980). Further, the 2p strain did not demonstrate improved total amorpha-4,11-diene production relative to the 3p strain, although the maximal production for these two strains occurred at different inducer concentrations. To identify bottlenecks in the pathway, the levels of the pathway enzymes from these two systems were analyzed (Fig. 6).

SRM transitions for the pathway proteins were selected from the most abundant peptides that provided unique identification of each protein (see Section 2). The majority of the proteins were

Table 2
List of strain names and the corresponding plasmids. For a full plasmid description, see Table 1.

Strain	Plasmids
3p	pMevT, pMBIS, pADS
2p	pBbA5c-MevT-MBIS, pADS
2p-CO	pBbA5c-MevT-MBIS(CO), pADS
2p-T1	pBbA5c-MevT-T1-MBIS, pADS
2p-T1-CO	pBbA5c-MevT-T1-MBIS(CO), pADS

readily detected by independent data acquisition (IDA) analysis and subsequently optimized. As shown in Fig. 6, there are many differences between mevalonate pathway protein levels in the 2p and 3p systems. SRM measurements indicated that the levels of most proteins in the mevalonate pathway were higher in the 2p strain compared to the 3p strain, with the exception being PMD. Notably, the first three proteins (AtoB, HMGS, and HMGR) in the pathway were at least six-fold higher in the 2p strain as compared to the 3p strain. This observation was unexpected considering that amorpha-4,11-diene production levels were equal, or lower, in the 2p strain. Based on this data, we speculated that there was a bottleneck present downstream of HMGR in the pathway.

The most striking aspect of the SRM data was the poor production of mevalonate kinase (MK) and phosphomevalonate kinase (PMK). Initial concerns about MK and PMK levels focused on potential technical problems (*e.g.*, sample preparation, SRM transition optimization.) Alternate cellular lysis and protein solubilization techniques (*e.g.*, 8 M Urea) also failed to yield MK or PMK peptides in targeted and non-targeted analysis, demonstrating that the proteins were not forming inclusion bodies or insoluble protein aggregates, rather they were not being produced. To facilitate optimization of SRM transitions for these proteins, His-tagged versions of these proteins were enriched using Ni-NTA column purification and analyzed in IDA mode.

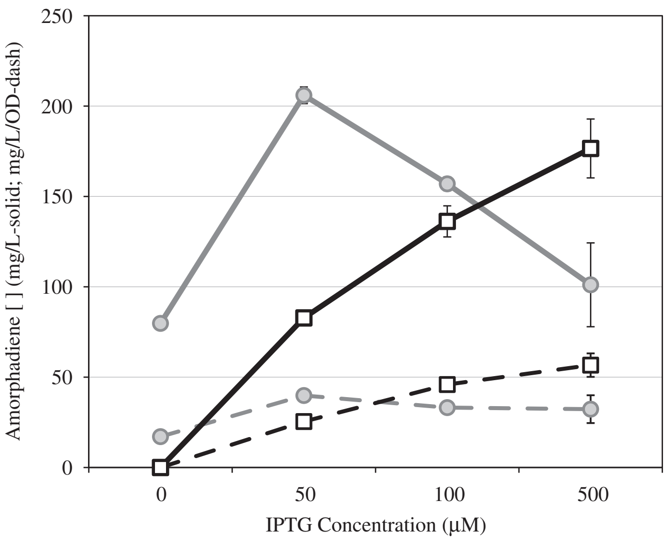

Fig. 5. Amorpha-4,11-diene production data (EZ rich media, 30 °C growth) from various inducer concentrations in the 3p (○) and 2p (□) strains. Lines were used to connect data series, with solid lines denoting total production and dashed lines denoting specific production, which takes into account the optical density of the culture (mg/L/OD). Results are the average of three biological replicates, with error bars representing the standard deviation from the mean.

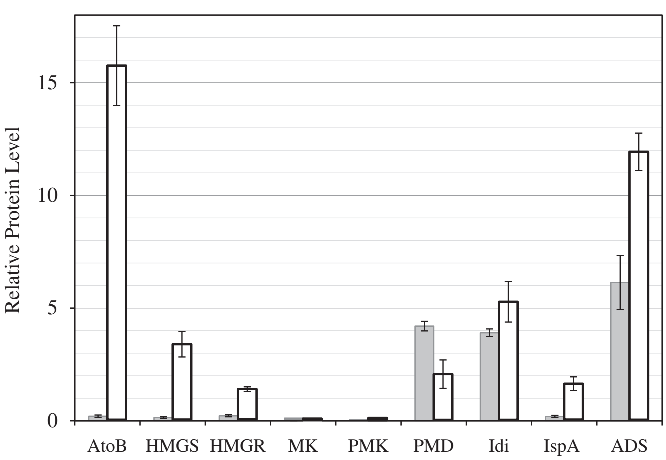

Fig. 6. Relative protein levels (M9 media, 37 °C growth, and 500 μM IPTG) for amorpha-4,11-diene pathway proteins in the 3p (■) and 2p (□) strains. Results are the average of three technical replicates, with error bars representing the standard deviation from the mean.

Once SRM transitions for MK and PMK peptides were optimized, small quantities of the enriched proteins were spiked into the cellular lysate and identified, which demonstrated that they were detectable in complex samples (data not shown). Taken together, these results point to problems in the engineered pathway.

The very low levels of MK and PMK could be due to several factors, including rapid protein degradation or poor translation. RT-PCR analysis of the genes in the mevalonate pathway was used to investigate whether poor mRNA stability was limiting protein production. RT primers were designed to amplify a 70 base-pair region at both the 5′ and 3′ end of the MK and PMK transcripts, and at the 3' end of *arcA* transcript that was used as a control. The results for 5′- and 3′-MK, 5′- and 3′-PMK and 3′-PMD are shown (Fig. 7). The contrast between the protein and mRNA data is quite striking; based upon mRNA data alone, one would anticipate that MK and PMK proteins would be expressed at detectable levels, as the mRNA levels are comparable to that of PMD (Fig. 7). The mRNA for MK is less abundant at the 5′-end than 3′-end, an

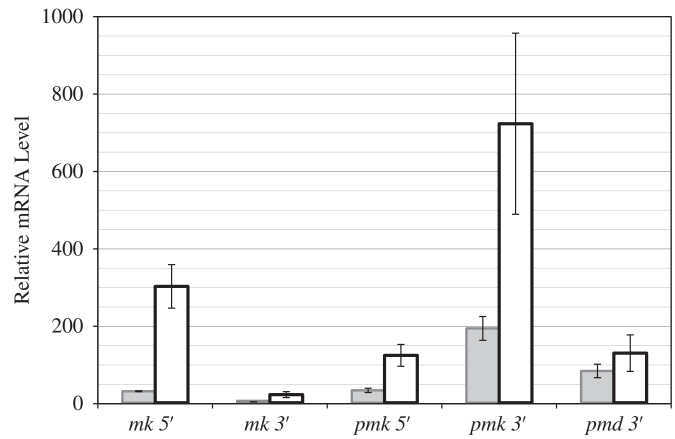

Fig. 7. Relative mRNA levels from RT-PCR (M9 media, 37 °C growth, and 500 μM IPTG) for selected genes from the 3p (■) and 2p (□) strains. Cells were harvested at 4 hours post-induction for mRNA extraction. Standard curves were used to calculate the efficiency of primer amplification. Results are the average of three technical replicates, with error bars showing the standard deviation from the mean.

indicator of RNA degradation, which may explain the low level of protein observed. In contrast, the PMK gene showed an unusual trend where the 5′-end of the transcript is present at lower levels than the 3′-end. The combination of these two different profiles may be explained by the presence of an internal promoter within the open reading frame and/or a preferential RNase E cleavage site (Carpousis et al., 2009). Interestingly, PMD, Idi, and IspA in both the 3p and 2p strains are produced without induction and the levels do not change significantly upon addition of the inducer. This suggests that gene expression is not under the control of the *lac*UV5 promoter and supports the hypothesis that an internal promoter is present.

Because it appeared that the mRNA coding sequence was non-ideal for the MK and PMK gene expression, plasmid pBbA5c-MevT-MBIS(CO) was constructed, which has the same structure as pBbA5c-MevT-MBIS plasmid but contains versions of MK and PMK that have been codon-optimized for expression in *E. coli*. Interestingly, the native yeast PMD sequence produced high levels of protein in *E. coli* despite having a similar rare codon distribution as that of MK and PMK. Codon-optimization did improve the levels of MK and PMK proteins (Fig. 8a) and improved the level of amorpha-4,11-diene production (Fig. 8b). Strikingly, most of the mevalonate pathway proteins were produced at lower levels in these constructs, even while amorpha-4,11-diene production increased. In contrast to the 3p and 2p strains, the protein levels for PMD, Idi, and IspA did respond to changes in inducer concentration in these constructs containing codon-optimized MK and PMK, indicating that they were actually under the control of the *lac*UV5 promoter. Since a small improvement in MK and PMK protein levels gave a corresponding increase in production, it was hypothesized that increasing these protein levels further may result in additional production. Thus, a *trc* promoter was added prior to the MK gene (Fig. 4). A terminator sequence was also inserted prior to this second promoter on the plasmid to ensure that only one promoter would drive the expression of the downstream genes. Increasing the promoter strength had a dramatic impact on the level of the downstream proteins (Fig. 8a), and also gave a 330% improvement in amorpha-4,11-diene production compared to the 2p strain at 24 h post-induction (Fig. 8b).

Based on the improved protein production that was observed in the codon-optimized constructs, it was of interest to determine if codon-optimization had altered the transcript profile of these

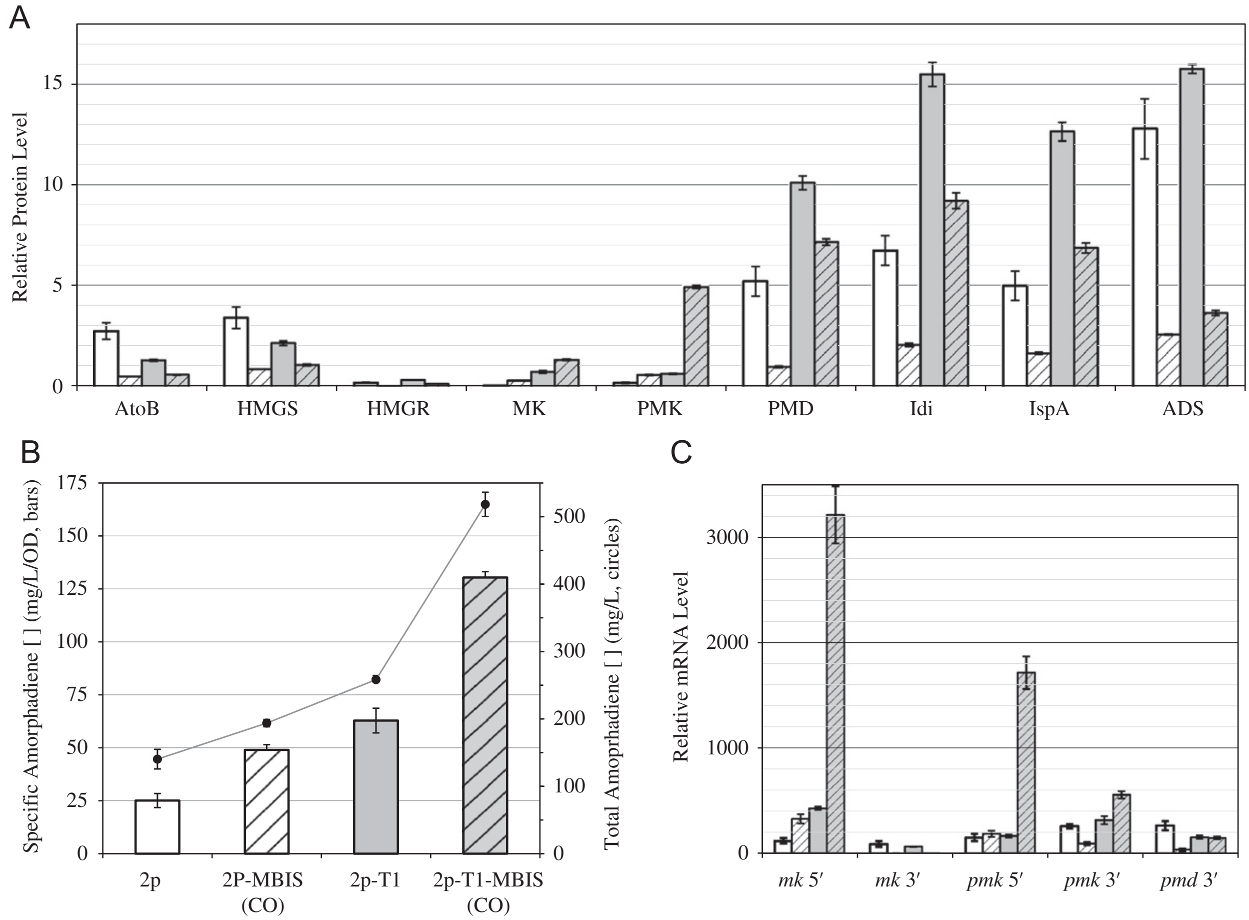

Fig. 8. Strains 2p (□), 2p-CO ( ), 2p-T1 ( ), and 2p-T1-CO ( ) containing variations of the pBbA5c-MevT-MBIS plasmid were investigated (EZ rich media, 30 °C growth, and 100 μM IPTG) for: (A) Relative protein levels. Values are shown as the average from three technical replicates, with error bars calculated as the standard deviation from the mean. (B) Amorpha-4,11-diene production. Both specific concentration (bars), relative to OD_{600} measurements (mg/L/OD), as well as total concentration (circles) are shown. Values are shown as the average from three biological replicates, with error bars calculated as the standard deviation from the mean. (C) Relative mRNA levels from RT-PCR results. Standard curves were used to calculate the efficiency of primer amplification. Values shown are the average of three technical replicates, with error bars calculated as the standard deviation from the mean.

genes (Fig. 8c). The data show that codon-optimization of the genes did result in increased levels of MK transcript at the 5′-end, but did not increase the total level of PMK transcript; however, the level of the codon-optimized PMK transcript was more consistent between the 5′- and 3′-end than previously observed. Introduction of the *trc* promoter alone (without codon-optimization of MK and PMK genes) improved the total level of MK transcript, but did not improve levels of PMK or PMD, again suggesting that there was a problem in the coding region of the PMK gene. In contrast, when the effects of codon-optimization and the additional *trc* promoter were combined, a substantial increase in the total level of transcript for both MK and PMK was obtained.

Comparison of the protein data with the mRNA data revealed some striking differences between the transcript data and the protein data. For example, codon-optimization of the PMK gene resulted in no change in the level of transcript produced for this gene, yet the protein level increased by 2.7 fold. Yet, when the *trc* promoter was introduced the PMK transcript increased by 10 fold, and gave a corresponding 7.5 fold increase in protein. In contrast, the MK transcript increased by 7.5 fold, but the corresponding protein level increased by only 1.8 fold. These data highlight the complexity of transitioning between transcript and protein levels *in vivo*. The protein data (Fig. 8a) demonstrate that making changes to small stretches of the mRNA coding sequence, such as introducing a downstream terminator and second promoter, can cause great variability in the levels of the upstream proteins. However, the results from amorpha-4,11-diene production (Fig. 8b) suggest that increasing the amount of the MK and PMK proteins is more important for increasing amorpha-4,11-diene titer than the downstream proteins. Future work will include improving the low levels of HMGR. Increased expression of this protein may also yield greater amounts of amorpha-4,11-diene.

4. Conclusions

In many metabolic engineering projects, precise control of expression of various components comprising a heterologous metabolic pathway is needed to avoid excess burden to the cell or the production of toxic intermediates. Information regarding the level of each pathway protein in the cell provides complementary information to that obtained from metabolite or mRNA measurements. Furthermore, these measurements facilitate characterization of parts for synthetic biology applications and identification of post-transcriptional or post-translational factors that impact production. This type of information, in combination with production data and cellular growth rate, enables the development of clearer and more focused hypotheses for directed pathway engineering. In the targeted proteomics studies described here, an SRM-based mass spectrometric method was used to determine protein levels from the *E. coli* strains engineered to produce amorpha-4,11-diene, resulting in key information that had not been revealed by any other cellular measurement.

Using the SRM technique, we were able to identify regions of the nine-gene amorpha-4,11-diene pathway where increased protein expression directly resulted in improved titer. Overall, the improvements in protein production led to a greater than 300% improvement in final titer within 24 h of production in shake-flask studies. Further, the ability to directly analyze the protein levels allowed for accurate interpretation of the results of pathway optimization. Dramatic differences in the level of mevalonate pathway proteins were observed between the 3p and 2p strains, without a corresponding increase in total productivity of the strain. While the majority of the pathway proteins were detected at an appreciable level, potential bottlenecks were evident for mevalonate kinase (MK) and phosphomevalonate kinase (PMK) from *S. cerevisiae*. Subsequent quantitative PCR analysis revealed that mRNA from both MK and PMK exhibited abnormal characteristics. Codon-optimization of MK and PMK genes and expression from a stronger promoter led to significant improvements in the protein levels and corresponding amorpha-4,11-diene production. These results begin to demonstrate that optimization of individual gene sequences is hampered by a complex interplay between the transcription and translation machinery in living systems. Further, determination of protein production levels constitutes an important metric to facilitate pathway optimization of metabolically engineered organisms and enable characterization of parts for use in synthetic biology.

Disclosure statement

Jay D. Keasling has financial interests in Amyris Biotechnologies and LS9, INC.

Acknowledgments

This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy. Purified mevalonate kinase and phosphomevalonate kinase were kindly provided by Amyris Biotechnologies. The authors thank Mario Ouellet for help with RT-PCR experiments, Dr. Mary Dunlop for donating the Δ*acrAB* strain, and Dr. Leonard Katz for manuscript review.

References

Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., 2005. Tuning genetic control through promoter engineering. Proc. Natl. Acad. Sci. USA 102, 12678–12683.

Anderson, J.C., Dueber, J., Leguia, M., Wu, G., Goler, J., Arkin, A., Keasling, J., 2010. BglBricks: a flexible standard for biological part assembly. J. Biol. Eng. 4, 1.

Anderson, L., Hunter, C.L., 2006. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol. Cell. Proteomics 5, 573–588.

Anthony, J.R., Anthony, L.C., Nowroozi, F., Kwon, G., Newman, J.D., Keasling, J.D., 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, 13–19.

Arnott, D., Kishiyama, A., Luis, E., Ludlum, S., Marsters, J., Stults, J., 2002. Selective detection of membrane proteins without antibodies—a mass spectrometric version of the Western blot. Mol. Cell. Proteomics 1, 148–156.

Baillie, T.A., 2008. Metabolism and toxicity of drugs. two decades of progress in industrial drug metabolism. Chem. Res. Toxicol. 21, 129–137.

Bloom, J., Meyer, M., Meinhold, P., Otey, C., MacMillan, D., Arnold, F., 2005. Evolving strategies for enzyme engineering. Curr. Opin. Struct. Biol. 15, 447–452.

Carpousis, A., Luisi, B., McDowall, K., 2009. Endonucleolytic initiation of mRNA decay in *Escherichia coli*. In: Proceedings of the Molecular Biology of RNA Processing and Decay in Prokaryotes, Progress in Molecular Biology and Translational Science.

Celinska, E., 2010. Debottlenecking the 1,3-propanediol pathway by metabolic engineering. Biotechnol. Adv. 28, 519–530.

Colling, J., Groenewald, J., Makunga, N., 2010. Genetic alterations for increased coumarin production lead to metabolic changes in the medicinally important *Pelargonium sidoides* DC (Geraniaceae). Metab. Eng. 12, 561–572.

De Mey, M., Maertens, J., Lequeux, G., Soetaert, W., Vandamme, E., 2007. Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. BMC Biotechnol. 7, 34.

Dehal, P.S., Joachimiak, M.P., Price, M.N., Bates, J.T., Baumohl, J.K., Chivian, D., Friedland, G.D., Huang, K.H., Keller, K., Novichkov, P.S., Dubchak, I.L., Alm, E.J., Arkin, A.P., 2009. MicrobesOnline: an integrated portal for comparative and functional genomics. Nucleic Acids Res., D396–D400.

Dueber, J., Wu, G., Malmirchegini, G., Moon, T., Petzold, C., Ullal, A., Prather, K., Keasling, J., 2009. Synthetic protein scaffolds provide modular control over metabolic flux. Nat. Biotechnol. 27, 753–759.

Endy, D., 2005. Foundations for engineering biology. Nature 438, 449–453.

Fortman, J., Chhabra, S., Mukhopadhyay, A., Chou, H., Lee, T., Steen, E., Keasling, J., 2008. Biofuel alternatives to ethanol: pumping the microbial well. Trends Biotechnol. 26, 375–381.

Gosset, G., 2009. Production of aromatic compounds in bacteria. Curr. Opin. Biotechnol. 20, 651–658.

Grabherr, R., Bayer, K., 2002. Impact of targeted vector design on ColE1 plasmid replication. Trends Biotechnol. 20, 257–260.

Hanahan, D., 1983. Studies of transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166, 557–580.

Herrero, O., Ramon, D., Orejas, M., 2008. Engineering the *Saccharomyces cerevisiae* isoprenoid pathway for *de novo* production of aromatic monoterpenes in wine. Metab. Eng. 10, 78–86.

Hiszczynska-Sawicka, E., Kur, J., 1997. Effect of *Escherichia coli* IHF mutations on plasmid p15A copy number. Plasmid 38, 174–179.

Huang, B., Guo, J., Yi, B., Yu, X., Sun, L., Chen, W., 2008. Heterologous production of secondary metabolites as pharmaceuticals in *Saccharomyces cerevisiae*. Biotech. Lett. 30, 1121–1137.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., Mann, M., 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol. Cell. Proteomics 4, 1265–1272.

Keasling, J.D., 2008. Synthetic biology for synthetic chemistry. ACS Chem. Biol. 3, 64–76.

Kues, U., Stahl, U., 1989. Replication of plasmids in gram-negative bacteria. Microbiol. Mol. Biol. Rev. 53, 491.

Kuhn, E., Wu, J., Karl, J., Liao, H., Zolg, W., Guild, B., 2004. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and C-13-labeled peptide standards. Proteomics 4, 1175–1186.

Le, T.T., Guet, C.C., Cluzel, P., 2006. Protein expression enhancement in efflux-deleted mutant bacteria. Protein Express. Purif. 48, 28–31.

Lee, S.K., Keasling, J.D., 2008. Heterologous protein production in *Escherichia coli* using the propionate-inducible pPro system by conventional and auto-induction methods. Protein Express. Purif. 61, 197–203.

Luetke-Eversloh, T., Stephanopoulos, G., 2008. Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression. Metab. Eng. 10, 69–77.

Lutz, R., Bujard, H., 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res. 25, 1203.

Maier, T., Güell, M., Serrano, L., 2009. Correlation of mRNA and protein in complex biological samples. FEBS Lett. 583, 3966–3973.

Makrides, S., 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiol. Rev. 60, 512–538.

Martin, L., Che, A., Endy, D., 2009. Gemini, a bifunctional enzymatic and fluorescent reporter of gene expression. PLoS One 4, e7569.

Martin, V., Pitera, D., Withers, S., Newman, J., Keasling, J., 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. Nat. Biotechnol. 21, 796–802.

Moon, T., Dueber, J., Shiue, E., Prather, K., 2010. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. Metab. Eng. 12, 298–305.

Munoz-Bertomeu, J., Ros, R., Arrillaga, I., Segura, J., 2008. Expression of spearmint limonene synthase in transgenic spike lavender results in an altered monoterpene composition in developing leaves. Metab. Eng. 10, 166–177.

Neidhardt, F.C., Bloch, P.L., Smith, D.F., 1974. Culture medium for enterobacteria. J. Bacteriol. 119, 736–747.

Ow, D.S., Lee, D., Tung, H., Lin-Chao, S., 2009. Plasmid Regulation and Systems-Level Effects on Escherichia coli Metabolism. In: Proceedings of the Systems Biology and Biotechnology of *Escherichia coli*. pp. 273–294.

Palmer, A.C., Angelino, E., Kishony, R., 2010. Chemical decay of an antibiotic inverts selection for resistance. Nat. Chem. Biol. 6, 105–107.

Pan, S., Aebersold, R., Chen, R., Rush, J., Goodlett, D., McIntosh, M., Zhang, J., Brentnall, T., 2009. Mass spectrometry based targeted protein quantification: methods and applications. J. Proteome Res. 8, 787–797.

Pfleger, B., Pitera, D., Smolke, C., Keasling, J., 2006. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nat. Biotechnol. 24, 1027–1032.

Picotti, P., Bodenmiller, B., Mueller, L., Domon, B., Aebersold, R., 2009. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. Cell 138, 795–806.

Pitera, D.J., Paddon, C.J., Newman, J.D., Keasling, J.D., 2007. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. Metab. Eng. 9, 193–207.

Ricci, J.C.D., Hernández, M.E., 2000. Plasmid effects on *Escherichia coli* metabolism. Crit. Rev. Biotechnol. 20, 79.

Rozkov, A., Avignone-Rossa, C., Ertl, P., Jones, P., O'Kennedy, R., Smith, J., Dale, J., Bushell, M., 2004. Characterization of the metabolic burden on *Escherichia coli* DH1 cells imposed by the presence of a plasmid containing a gene therapy sequence. Biotechnol. Bioeng. 88, 909–915.

Salis, H.M., Mirsky, E.A., Voigt, C.A., 2009. Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. 27, 946–950.

Sambrook, J., 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor N.Y.

Silva, J., Gorenstein, M., Li, G., Vissers, J., Geromanos, S., 2006. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. Mol. Cell. Proteomics 5, 144–156.

Solem, C., Jensen, P.R., 2002. Modulation of gene expression made easy. Appl. Environ. Microbiol. 68, 2397–2403.

Stefano, J., Ackerson, J., Gralla, J., 1980. Alterations in two conserved regions of promoter sequence lead to altered rates of polymerase binding and levels of gene-expression. Nucleic Acids Res. 8, 2709–2723.

Stoebel, D., Deant, A., Dykhuizen, D., 2008. The cost of expression of *Escherichia coli* lac operon proteins is in the process, not in the products. Genetics 178, 1653–1660.

Tyo, K.E., Fischer, C.R., Simeon, F., Stephanopoulos, G., 2010. Analysis of polyhydroxybutyrate flux limitations by systematic genetic and metabolic perturbations. Metab. Eng. 12, 187–195.

Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., Church, G.M., 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894–898.

Wang, W., Zhou, H., Lin, H., Roy, S., Shaler, T.A., Hill, L.R., Norton, S., Kumar, P., Anderle, M., Becker, C.H., 2003. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. Anal. Chem. 75, 4818–4826.

Wessel, D., Flugge, U., 1984. A method for the quantitative recovery of protein in dilute-solution in the presence of detergents and lipids. Anal. Biochem. 138, 141–143.

Zhang, B., VerBerkmoes, N., Langston, M., Uberbacher, E., Hettich, R., Samatova, N., 2006. Detecting differential and correlated protein expression in label-free shotgun proteomics. J. Proteome Res. 5, 2909–2918.