Production of shikimic acid from *Escherichia coli* through chemically inducible chromosomal evolution and cofactor metabolic engineering

Yan-Yan Cui¹, Chen Ling², Yuan-Yuan Zhang¹, Jian Huang¹ and Jian-Zhong Liu¹*

**Abstract**

**Background:** Shikimic acid (SA) produced from the seeds of Chinese star anise (*Illicium verum*) is a key intermediate for the synthesis of neuraminidase inhibitors such as oseltamivir (Tamiflu®), an anti-influenza drug. However, plants cannot deliver a stable supply of SA. To avoid the resulting shortages and price fluctuations, a stable source of affordable SA is required. Although recent achievements in metabolic engineering of *Escherichia coli* strains have significantly increased SA productivity, commonly-used plasmid-based expression systems are prone to genetic instability and require constant selective pressure to ensure plasmid maintenance. Cofactors also play an important role in the biosynthesis of different fermentation products. In this study, we first constructed an *E. coli* SA production strain that carries no plasmid or antibiotic marker. We then investigated the effect of endogenous NADPH availability on SA production.

**Results:** The *pps* and *csrB* genes were first overexpressed by replacing their native promoter and integrating an additional copy of the genes in a double gene knockout (*aroK* and *aroL*) of *E. coli*. The *aroGfb*, *aroB*, *aroE* and *tktA* gene cluster was integrated into the above *E. coli* chromosome by direct transformation. The gene copy number was then evolved to the desired value by triclosan induction. The resulting strain, *E. coli* SA110, produced 8.9-fold more SA than did the parental strain *E. coli* (*Δ*aroKΔaroL). Following qRT-PCR analysis, another copy of the *tktA* gene under the control of the 5P* tac* promoter was inserted into the chromosome of *E. coli* SA110 to obtain the more productive strain *E. coli* SA110. Next, the NADPH availability was increased by overexpressing the *pntAB* or *nadK* genes, which further enhanced SA production. The final strain, *E. coli* SA116, produced 3.12 g/L of SA with a yield on glucose substrate of 0.33 mol/mol.

**Conclusion:** An SA-producing *E. coli* strain that carries neither a plasmid nor an antibiotic marker was constructed by triclosan-induced chromosomal evolution. We present the first demonstration that increasing NADPH availability by overexpressing the *pntAB* or *nadK* genes significantly enhances SA production.

**Keywords:** Shikimic acid, *Escherichia coli*, Chemically induced chromosomal evolution, NADPH, Transhydrogenase, NAD kinase

---

**Background**

Shikimic acid is a key intermediate for the synthesis of the neuraminidase inhibitor oseltamivir (Tamiflu®), an anti-influenza treatment [1]. The main commercial source of SA is seeds of the *Illicium* plant, such as *I. verum* or *I. anisatum*. However, the conventional method of producing SA from *I. verum* is typically low-yield and costly. Therefore, researchers have developed several metabolic engineering approaches to overproduce SA in *E. coli* [2,3]. These approaches are based on genetic modifications of central carbon metabolism and SA pathways. Specifically, the aromatic amino acid pathway is blocked after the SA production stage by deleting the *aroK* and *aroL* genes encoding shikimate kinase I and II. To increase the carbon flux from central carbon metabolism entering the aromatic amino acid pathway, researchers have amplified the feedback resistant 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase genes *aroF* or *aroG* [4-8]. These modifications are commonly complemented...
with over-expression of the aroB and aroE genes [4-8]. Over-expression of the tktA gene (encoding transketolase) enhances the availability of erythrose 4-phosphate (E4P), and consequently increases the SA titer from 38 g/L to 52 g/L [6]. Over-expression of the pps gene (encoding phosphoenolpyruvate (PEP) synthase) has elevated the SA titer to 66 g/L [7]. SA production can also be increased by inactivating the phosphoenolpyruvate: carbohydrate phosphotransferase system (PTS) operon, combined with expressing the ATP dependent uptake and phosphorylation system comprising the glucose facilitator and the glucokinase from Zymomonas mobilis [7,8]. E. coli engineered in this way yielded SA concentrations as high as 87 g/L in minimum medium supplemented with yeast extract [7]. The constitutive and synchronous expression of a six-gene synthetic operon (aroG<sup>SR</sup>, aroB, aroD, aroE, tktA and zwf), in a laboratory-evolved strain bearing simultaneous PTS and pykF inactivations, was recently reported to increase the SA yield on glucose to 42% mol/mol, which represents the highest reported yield [9].

However, all of the above studies use plasmids for gene expression. Among the drawbacks of plasmid-based expression systems are structural and segregational instability, and allele segregation [10-12]. These plasmid instabilities cause genetic instability, with decreased productivity of the desired compound. Tyo et al. [13] reported that plasmid-carrying strains lose poly-3-hydroxybutyrate productivity after 40 rounds of subculturing. Moreover, the markers used for selecting and maintaining plasmids in hosts during cultivation are usually antibiotic resistance genes. However, antibiotics are both costly and banned from food and pharmaceutical production processes. In addition, the potential spread of antibiotic-resistant marker to natural microbes requires serious consideration. The likely outcome of this scenario is rapid emergence of multidrug-resistant organisms (e.g., superbacteria) [14,15]. To overcome these drawbacks of plasmid constructs, Tyo et al. [13] developed a plasmid-free method that achieves high copy numbers of the desired genes, termed chemically induced chromosomal evolution (CIChE). However, the λInCh genomic integration protocol used in CIChE is complicated and time-consuming, and Tyo et al.’s CIChE strains still carry an antibiotic resistance marker (chloramphenicol resistance) [13]. In our previous paper [16], we reported a modified CIChE that overcomes the drawbacks of Tyo et al.’s original CIChE protocol. The resulting strain, developed by triclosan-induced chromosomal evolution, carries neither a plasmid nor an antibiotic marker. Thus, our first task in this study was to construct an SA-producing E. coli strain by triclosan-induced chromosomal evolution.

Cofactors are also known to play an important role in the biosynthesis of different fermentation products. Once the enzyme levels are no longer limiting, cofactor availability can become limiting and productivity must be boosted by cofactor manipulation. In silico flux analysis has identified a potentially important role for intracellular NADPH concentration in SA biosynthesis [17]. Thus, we also investigated the effect of endogenous NADPH availability on SA production by over-expressing the genes involved in NADPH synthesis.

**Results and discussion**

**Inactivation of genes encoding shikimate kinase I and II, and chromosomal promoter replacement**

To construct an SA-accumulating host strain, we initially focused on preventing the conversion of SA to chorismic acid in E. coli. To this end, we deleted the aroK and aroL genes to obtain E. coli BW25113 (ΔaroKΔaroL). PEP synthase (encoded by the pps gene) converts pyruvic acid to PEP. Carbon storage regulator CsrA controls glycogen synthesis and modulates the levels of three enzymes that directly participate in PEP metabolism: pyruvate kinase, PEP carboxykinase and PEP synthase (which synthesize PEP from oxaloacetate and pyruvate, respectively). The first of these is positively regulated, while both PEP enzymes are negatively regulated, by CsrA [18,19]. Because CsrA activity is antagonized by CsrB [18], PEP availability was increased by replacing the native promoters of the pps and csrB genes with the P<sub>lacQ1</sub> promoter in E. coli BW25113 (ΔaroKΔaroL), yielding E. coli BW25113 (ΔaroKΔaroL, P<sub>pps::PlacQ1</sub>, P<sub>csrB::PlacQ1</sub>). SA production in this strain was 53.3% higher than in E. coli BW25113 (ΔaroKΔaroL) ((0.23 ± 0.02) g/L vs (0.15 ± 0.04) g/L; P < 0.01; Table 1). To further improve PEP availability, an additional copy of the pps and csrB genes controlled by the T5 promoter was inserted into E. coli BW25113 (ΔaroKΔaroL, P<sub>pps::PlacQ1</sub>, P<sub>csrB::PlacQ1</sub>), yielding E. coli BW25113 (ΔaroKΔaroL, P<sub>pps::PlacQ1</sub>, P<sub>csrB::PlacQ1</sub>, P<sub>PPS::P<sub>Pr</sub></sub>P<sub>ars</sub>, P<sub>csr::PlacQ1</sub>) to obtain E. coli BW25113 (ΔaroKΔaroL, P<sub>pps::PlacQ1</sub>, P<sub>csrB::PlacQ1</sub>, P<sub>TPS::PPS</sub>, P<sub>TPS::csr</sub>). SA production in this strain was 73.3% higher than in E. coli BW25113 (ΔaroKΔaroL) ((0.26 ± 0.02) g/L vs (0.15 ± 0.04) g/L; P < 0.01; Table 1). These modifications also resulted in an increment in yield (mol SA/mol glucose) of about 80%. Over-expressing the pps gene increased DAHP production almost twofold, to approach the theoretical maximum [20]. DAHP is the precursor of SA and the condensation product of PEP and E4P, mediated by DAHP synthase. They thought that over-expressing the pps gene drives pyruvate to be recycled back to PEP [20]. Yakandawala et al. [18] achieved an approximately twofold increase in phenylalanine production by over-expressing the csrB gene. SA is the precursor of phenylalanine. Chandran et al. reported that over-expressing the pps gene raised the SA titer of 26% (from 52 to 66 g/L) [7]. Thus, the increments in SA titer and yield could be the consequence of overexpression of both pps and csrB genes.
Quantitative real-time PCR analysis demonstrates that the above modifications indeed enhanced the transcription levels of the \textit{pps} and \textit{csrB} genes (Figure 1). The replacement of the native promoter resulted in the upregulations of the transcription levels of the \textit{pps} and \textit{csrB} genes about 6.9- and 2.8-fold ($P<0.01$), respectively. Integrating an additional copy of the two genes further upregulated the transcription levels of the \textit{pps} and \textit{csrB} genes about 10.4- and 20.9-fold ($P<0.01$), respectively.

**Chemically induced chromosomal evolution**

Many papers reported that plasmid-based over-expression of feedback-resistant DAHP synthases (coded by \textit{aroF}^{fbr} or \textit{aroG}^{fbr}), shikimate dehydrogenase (coded by \textit{aroE}), transketolase (coded by \textit{tktA}), and DHQ synthase enzymes (coded by \textit{aroB}) [2-9]. To overcome the drawbacks of plasmid expression systems, the \textit{aroG}^{fbr}, \textit{tktA}, \textit{aroB} and \textit{aroE} gene cluster was inserted into the chromosome of \textit{E. coli} BW25113 (\textit{ΔaroKΔaroL}, \textit{P}_{pps::PlacQ1,P_{csrB::PlacQ1}, P_{T5-pps}, P_{T5-csrB}}) by direct transformation. The transformed strain was evolved to higher gene copy number by exposure to increasing triclosan concentrations. Figure 2 plots the SA production in CICChE strains resistant to different triclosan concentrations. The maximum SA concentration ($1.34 \pm 0.15$) g/L was obtained by the CICChE strains resistant to 2 $\mu$M triclosan. When the triclosan concentration exceeded 2 $\mu$M, SA production of the CICChE strains did not further increase. Thus, the \textit{recA} gene of the CICChE strain resistant to 2 $\mu$M triclosan was deleted to obtain \textit{E. coli} SA110. Homologous recombination, which can potentially reduce the copy number, is inhibited in this strain. Figure 3 shows that the copy numbers and transcription levels of the \textit{aroE} gene of the CICChE strains increase with triclosan concentration during chromosomal evolution. At a triclosan concentration of 2 $\mu$M, the copy number reached about 8

**Table 1 Shikimic acid (SA) production by different strains**

| Stain                              | OD$_{600}$ | SA concentration (g/L) | Yield (mol SA/mol glucose) |
|------------------------------------|------------|------------------------|---------------------------|
| BW25113 (ΔaroKΔaroL)               | 7.18 ± 0.95| 0.15 ± 0.04            | 0.05                      |
| BW25113 (ΔaroKΔaroL, \textit{P}_{pps::PlacQ1,P_{csrB::PlacQ1}}) | 5.82 ± 0.04| 0.23 ± 0.02            | 0.08                      |
| BW25113 (ΔaroKΔaroL, \textit{P}_{pps::PlacQ1,P_{csrB::PlacQ1}, P_{T5-pps}, P_{T5-csrB}}) | 5.82 ± 0.57| 0.26 ± 0.02            | 0.09                      |
| SA110                              | 5.91 ± 1.34| 1.34 ± 0.15            | 0.21                      |
| SA112                              | 5.85 ± 0.46| 1.70 ± 0.01            | 0.25                      |

**Figure 1** Transcription levels of the \textit{pps} and \textit{csrB} genes in the different strains.
in the CIChE strain. When the triclosan concentration was above 2 μM, the gene copy numbers and transcription levels still increased; however, the SA production of the CIChE strains did not increase. The transcription levels of the other three genes show the similar trend (Additional file 1: Figure S1). The results indicated that there is an optimal copy number and transcription level of the gene cluster for efficient production of SA.

We also analyzed the transcription levels of the *E. coli* SA110 genes by qRT-PCR, and compared the data with those of *E. coli* BW25113 (ΔaroKΔaroL) (Figure 4). The transcription levels of the *tkta*, *aroG* flb, *aroB* and *aroE* genes in *E. coli* SA110 were about 1.6, 14.7, 35.2 and 8.1 times higher \((p < 0.05)\) than in *E. coli* BW25113 (ΔaroKΔaroL). The transcription level of the *tkta* gene was the lowest. It may be because the distance of the *tkta* gene from the promoter is the longest [21]. They thought that the abundance of mRNA decreased monotonically with the increasing distance of the gene from the promoter in *E. coli* [21]. The gene ranked in front will be translated primarily, followed by the translation of the subsequent genes along with the mRNA transcription for polycistronic operons in *E. coli*. However, the transcription level of the second gene (the *aroB* gene) from the promoter was the highest. This could be the consequence of codon usage biases.

![Figure 2](image1.png)

**Figure 2** Shikimic acid production of CIChE strains at different triclosan concentrations.

![Figure 3](image2.png)

**Figure 3** Gene copy number (line) and transcription level (column) of the *aroE* gene in CIChE strains.
Because the tktA gene was least upregulated among these genes, we examined whether tktA is the bottleneck for SA production in E. coli SA110. An additional copy of the tktA gene under the control of the 5P_tac promoter was inserted into the chromosome of E. coli SA110 to obtain E. coli SA112. This modification raised the SA titer to (1.70 ± 0.01) g/L and the SA yield to 0.25 mol/mol (Table 1). Figure 4 shows that this integration upregulated the transcription level of the tktA gene. The transcription level of the tktA gene in E. coli SA112 was much higher than that in E. coli SA110. This may be because the additional integrated tktA gene was directly controlled by the stronger 5P_tac promoter.

**Effect of NADPH availability**

To investigate the effect of NADPH availability on SA production, transhydrogenase (encoded by sthA or pntAB) and NAD kinase (encoded by nadK) were first amplified by plasmid-based over-expression in E. coli SA112. The results are presented in Table 2. Overexpression of the pntAB or nadK genes increased both SA production and the intracellular NADPH concentration. E. coli generates NADPH in one of three ways: the pentose phosphate (PP) pathway, isocitrate dehydrogenase and transhydrogenase. Two isoforms of transhydrogenase exist in E. coli, membrane-bound transhydrogenase (PntAB) and soluble transhydrogenase (SthA). During aerobic batch growth with glucose, 35–45% of the NADPH required for biosynthesis is produced via PntAB, while the PP pathway and isocitrate dehydrogenase generate 35–45% and 20–25%, respectively [22]. ATP-dependent NAD kinase encoded by nadK catalyzes the phosphorylation of NAD to NADP. Overexpression of nadK increases the size of the NADP pool, which potentially increases the abundance of NADPH. It can be seen in Table 2 that overexpression of the sthA gene increased SA titer and reduced the intracellular NADPH concentration. Moreover, the SA titer and intracellular NADPH concentration of the strain overexpressing sthA was lower than in strains overexpressing the pntAB or nadK genes. PntAB catalyzes the transfer of reducing power from NADH to NADP⁺ in an energy-dependent manner at low intracellular NADPH levels, while SthA mainly catalyzes the reoxidation of NADPH in an energy-independent manner when NADPH is abundant [22]. Thus, if the pntAB or sthA gene is overexpressed, the intracellular NADPH concentration is increased or decreased, respectively. As shown in Table 3, overexpressing the sthA gene indeed caused a reduction in intracellular NADPH concentration, with consequent reduction of the SA titer.

Because plasmid-based overexpression of sthA showed lower the SA titer compared to overexpression of the pntAB or nadK genes, to alleviate the metabolic burden caused by the plasmid, the pntAB and nadK genes were integrated into the chromosome of E. coli SA112 to obtain E. coli SA114 and SA116, respectively. Chromosomal overexpression enhanced the intracellular NADPH concentration and SA titer relative to plasmid overexpression (Table 2), and also improved the cell growth. It can be seen in Figure 5 that the transcription levels of the pntAB and nadK genes were enhanced about 3.3- and 3.6-fold (P < 0.01) after chromosomal integration, respectively.

This study demonstrated a strong correlation between NADPH availability and SA production. In the SA biosynthesis pathway, shikimate dehydrogenase (encoded by the

---

**Table 2 Effect of over-expressing genes involved in NADPH synthesis on SA production in E. coli SA112**

| Strain          | OD600  | SA concentration (g/L) | Yield (mol SA/mol glucose) | NADPH (μM) |
|-----------------|--------|------------------------|----------------------------|------------|
| SA112 (pMP5)    | 4.79 ± 0.64 | 1.34 ± 0.09 | 0.22 | 1.93 ± 0.01 |
| SA112 (pMPsthA) | 5.78 ± 0.26 | 1.42 ± 0.07 | 0.24 | 1.61 ± 0.01 |
| SA112 (pMPpntAB)| 5.50 ± 0.11 | 1.89 ± 0.01 | 0.27 | 2.23 ± 0.01 |
| SA112 (pMPnadK) | 5.20 ± 0.34 | 1.92 ± 0.01 | 0.29 | 2.55 ± 0.01 |
| SA114           | 6.55 ± 0.30 | 2.99 ± 0.01 | 0.31 | 4.51 ± 0.01 |
| SA116           | 6.40 ± 0.51 | 3.12 ± 0.01 | 0.33 | 5.77 ± 0.01 |

---

*Figure 4 Transcription levels of genes of interest in the CICHE strain E. coli SA110 and SA112 compared with those in E. coli BW25113 (ΔaroKΔaroL).*
aroE gene) catalyzes the reduction of 3-dehydroshikimate to shikimate, which requires NADPH as a cofactor. These CIChE strains (e.g. SA110, 112, 114, 116), in which the aroE gene was overexpressed and therefore required more NADPH for SA biosynthesis. Although NADPH can be provided by the oxidative part of the PP pathway, this branch is unfavorable for SA production. Stoichiometric analysis shows that the oxidative part of the PP pathway is not required in the reaction scheme. Glucose enters the nonoxidative part of the PP pathway, where it provides E4P for maximum theoretical SA yield [20]. Accordingly, the overall reaction for SA biosynthesis from glucose is: 7 glucose + 6 ATP + 6 NAD + 6 NADPH → 6 SA + 6 ADP + 6 NADH + 6 NADH + 6 NADP + 12H⁺ + 6 HPO₄²⁻. Increased intracellular NADPH via overexpression of the pntAB or nadK genes favored SA production (Table 2), suggesting that more carbon flux was channeled into the nonoxidative part of the PP pathway for E4P biosynthesis. In silico flux analysis has demonstrated a potentially important role for intracellular NADPH concentration in SA biosynthesis. Under different genetic and environmental conditions, E. coli cells consume the same quantities of NADPH to maximize their SA production [17]. According to some studies, maintaining NADPH availability improves metabolite production. For instance, overexpressing E. coli pntAB in Corynebacterium glutamicum enhances NADPH availability, in turn increasing the intracellular levels of L-lysine [23] and L-ornithine [24]. Overexpression of nadk in E. coli increases the NADPH/NADP ratio, thereby enhancing thymidine biosynthesis [25]. In C. glutamicum, overexpressing the NAD kinase gene improves L-lysine [26] and L-ornithine [24] production. Simultaneous chromosomal overexpression of transhydrogenase (pntAB) and NAD kinase (yfjB) genes in E. coli increases the NADPH supply and improves anaerobic isobutanol production [27].

The SA yield obtained in the present study is below the highest value (0.42 mol/mol) reported in the literature [9], achieved by the constitutive and synchronous expression of a six-gene synthetic operon, in a laboratory-evolved strain bearing simultaneous PTS and pykF inactivations in a 1 L bioreactor containing 100 g/L of glucose and 30 g/L of yeast extract. However, the SA yield obtained in this research is on part with the 33% yield reported by Chandran et al. [7]. They reported the highest SA titer (87 g/L). It remains far from the theoretical maximum (86% mol/mol) [20], implying a need for further research to improve this value. Currently, we are exploring whether inactivation of PTS, pykF and the oxidative part of the PP pathway ensures a high SA yield. The results of these modifications will be reported in a subsequent paper.

| Strain          | OD₆₀₀   | SA concentration (g/L) | Yield (mol SA/mol glucose) | NADPH (μM)   |
|-----------------|---------|------------------------|---------------------------|--------------|
| SA116(pMP5)     | 5.56 ± 0.76 | 1.85 ± 0.07            | 0.23                      | 1.67 ± 0.02  |
| SA116(pMPsthA)  | 6.73 ± 0.11 | 1.69 ± 0.11            | 0.20                      | 0.74 ± 0.04  |

Table 3 Effect of overexpressing the sthA gene on SA production in E. coli SA116

**Figure 5** Transcription levels of the pntAB (open bar) and nadK (black bar) in the different strains.
Table 4 Strains and plasmids used in this study

| Strain/plasmid | Description | Source or reference |
|----------------|-------------|---------------------|
| **Strain**     |             |                     |
| E. coli DH5α   | supE44 Δ(lacZΔM15-argF) U169 (O80lacZ ΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1 | Invitrogen |
| E. coli BW25113 | lacI, mbr14, ΔlacZΔM15, hsdR14, ΔaraBAD4113, ΔaraBAD4113D78 | 28 |
| E. coli BW25113 (ΔaroKΔaroL) | E. coli BW25113, ΔaroKΔaroL | This study |
| E. coli BW25113 (ΔaroKΔaroL, Ppps::PlacQ1, PcsrB::PlacQ1) | E. coli BW25113, ΔaroKΔaroL, replacement of the native promoter of the pps and csrB gene with the PlacQ1 promoter | This study |
| E. coli BW25113 (ΔaroKΔaroL, Ppps::PlacQ1, PcsrB::PlacQ1, P15-pps, P15-csrB) | E. coli BW25113 (ΔaroKΔaroL, Ppps::PlacQ1, PcsrB::PlacQ1) with an additional chromosomal copy of the pps and csrB genes under the control of the T5 promoter | This study |
| SA110          | CICHe strain of the aroGfbr, tktA, aroB and aroE gene cluster resistance to 2 μM triclosan from E. coli BW25113 (ΔaroKΔaroL, Ppps::PlacQ1, PcsrB::PlacQ1), ΔrecA | This study |
| SA112          | SA110 with an additional chromosomal copy of the tktA gene under the control of the 5P_lac promoter (5 tandem repeats of the core-tac-promoter) | This study |
| SA114          | SA112 with an additional chromosomal copy of the pntAB genes under the control of the 5P_lac promoter | This study |
| SA116          | SA112 with an additional chromosomal copy of the nadK gene under the control of the 5P_lac promoter | This study |
| **Plasmid**    |             |                     |
| pHKKF3T5b      | CICHe integration expression vector, attP                                                                            | 16 |
| pHKKT5b        | Integration expression plasmid, attP, P5_lac promoter, Kan’                                                          | 33,34 |
| pP21KTSb       | Integration expression plasmid, attP, site, P5_lac promoter, Kan’                                                      | 33,34 |
| pHKBEGB        | pHKBEGB derivative containing aroGfbr, tktA, aroB and aroE genes                                                     | This study |
| pP21KTSb-csrB-pps | pP21KTSb derivative containing the csrB and pps genes                                                                | This study |
| pBAD24         | pBAD24 derivative containing the csrB and pps genes                                                                   | Lab storage |
| pBAD-csrB-pps  | pBAD24 derivative containing the aroB and aroE genes                                                                   | This study |
| pBE6           | pBE6 derivative containing aroB and aroE genes                                                                        | This study |
| pMP5           | pMP5 derivative containing the aroB and aroE genes                                                                     | This study |
| pHKKT5b-tktA   | pHKKT5b derivative containing the tktA gene                                                                          | This study |
| pMPSthA        | pMPSthA derivative containing the sthA gene                                                                          | This study |
| pMPpntAB       | pMPpntAB derivative containing the pntAB gene                                                                         | This study |
| pHKSacb        | Integration expression plasmid, attP_lac site, Stac promoter, Kan’                                                   | This study |
| pP21KTSb-csrB-pps | pP21KTSb derivative containing the csrB and pps genes                                                               | This study |
| pHKBEGB        | pHKBEGB derivative containing the pntAB gene                                                                         | This study |
| pHKBEGBT       | pHKBEGBT derivative containing the pntAB gene                                                                         | This study |
| pP21KTSb-csrB-pps | pP21KTSb derivative containing the csrB and pps genes                                                               | This study |
| pHKKT5b-tktA   | pHKKT5b derivative containing the tktA gene                                                                          | This study |
| pHKKSacb       | Integration expression plasmid, attP_lac site, Stac promoter, Kan’                                                   | This study |
| pP21KTSb-pntAB | pP21KTSb derivative containing the pntAB gene                                                                         | This study |
| pHKSacb        | pHKSacb derivative containing the pntAB gene                                                                         | This study |
| Primer | Sequence and purpose* |
|--------|-----------------------|
| LP1    | 5'-ATGACACAACCCTTTTCATGATGCTGGGCCTCGGGCTGTGGTAGAAAACACAGCGATGTTAGGGCTGGAG-3', deletion of *aroL* |
| LP2    | 5'-TCGACAAATGATCGCTGTCGGGCGCCGCATTCCAGAACATCCAGATCGGGAATTGCTGGAG-3', deletion of *aroL* |
| P1     | 5'-GGTTCCTGGGCACACGCCCT-3', Diagnostic PCR for the deletion of *aroL* |
| P2     | 5'-ATTCTCATGACACCGGCCCTT-3', Diagnostic PCR for the deletion of *aroL* |
| KP1    | 5'-ATGACAGGAGAACGCAATATTTCTCCTGATTTGGGCTATGCCCAGAAAGCGCAATATTGTAGGCTGGAG-3', deletion of *aroK* |
| KP2    | 5'-TTAGTCTCGCTTCCAGACATAAATCTTGTTTGCAACACCAACTTTTAGGCCCTAAACGGCTGACATGGGAATTAG-3', deletion of *aroL* |
| P3     | 5'-GCGAGAGGGGTATTCCAT-3', Diagnostic PCR for the deletion of *aroK* |
| P4     | 5'-GTCGCCCCAGAGTAACGGAC-3', Diagnostic PCR for the deletion of *aroK* |
| GP1    | 5'-GCCTGACAGGAGGCGCTAAATGATTCGAAATACAGAGCAACAGCCATT-3', *PstI*, PCR for *aroG* |
| GP2    | 5'-CGGCATGCTTACCCGGAGCCGCCTTCCATT-3', *SpfI*, PCR for *aroG* |
| G15P1  | 5'-CTCAATAATGATCCAACCCACAT-3', site-specific mutagenesis of *aroG* |
| G15P2  | 5'-AAAATCATGCCTGGCCGAGCGGCG-3', site-specific mutagenesis of *aroG* |
| EP1    | 5'-GCGAATCCAGGAGTAAATAATGGAACATCATGCTTTT-3', *EcoRI*, PCR for *aroE* |
| EP2    | 5'-CGCCGGCCGCCCTTACCGGGAGAAATATTCTCTTCT-3', *NotI*, PCR for *aroE* |
| BP1    | 5'-CGCAGCAGGAGGGAATATGAGGAGATTGTCCGTAC-3', *NotI*, PCR for *aroB* |
| BP2    | 5'-GCCATGCTTACCCGGAGAACACAGC-3', *NcoI*, PCR for *aroB* |
| TP1    | 5'-GCACGATGAGGTAATATGATGTCTTCACGGAAAGAGCTTCT-3', *SpfI*, PCR for *nltA* |
| TP2    | 5'-CGAGACTCTTACAGCAGTCTCTTGTTCCTC-3', *SacI*, PCR for *tktA* |
| PntF   | 5'-CAGGGTACCTCATCAATCAAAAGCCT-3', *KpnI*, PCR for *pntAB* |
| PntR   | 5'-CGTCTGACGTGTCATAGCAGCCTTCTCAG-3', *PstI*, PCR for *pntAB* |
| Shf    | 5'-TTTGGTACCGGTAACCCTCAGCATTGC-3', *KpnI*, PCR for *stfA* |
| ShR    | 5'-GGGGTCGACGCGCCCTTACGGATTAAAGATTTT-3', *PstI*, PCR for *stfA* |
| NadF   | 5'-GCGGGGTACCAGATGATATGATTTCTCAAGGTC-3', *KpnI*, PCR for *nadK* |
| NadR   | 5'-GCCGTCTAGATGATAAATTTTTTGCACCA-3', *XbaI*, PCR for *nadK* |
| StacF  | 5'-ACGGGTGTAAGGGACACGCGCATG-3', *MluI*, PCR for the *Sp*<sub>lac</sub> promoter |
| StacR  | 5'-CCGGCCATGGGATCCAACTAGATGCTGATTTTCTCTCTTATGATG-3', *SpfI*, PCR for the *Sp*<sub>lac</sub> promoter |
| CsrF   | 5'-CGGAGCTCTGAGGTTAAATATGATAGTCCAGAAGGATGGAATTAC-3', *SacI*, PCR for the *csrB* and *pps* gene cluster |
| PpsR   | 5'-CGGTTACTTTTTCTCTTGCTACCAGCAGCAGC-3', *KpnI*, PCR for the *csrB* and *pps* gene cluster |
| QEF    | 5'-GGATGCCGCGGAAATATCCACACC-3', qPCR for the *aroE* gene |
| QER    | 5'-ACTACTGGCACTCCTTCTCTG-3', qPCR for the *aroE* gene |
| QBF    | 5'-AACGAAACCCTCTGCTCTG-3', qPCR for the *aro8* gene |
| QBR    | 5'-AAGGCGCAACCAGCTATGAT-3', qPCR for the *aro8* gene |
| QGF    | 5'-CGTTGCTGAGGTTAAATATGATAGTCCAGAAGGATGGAATTAC-3', qPCR for the *aro8* gene |
| QGR    | 5'-ACGCTGACGACACATCCAC-3', qPCR for the *aro8* gene |
| QTF    | 5'-TTGCTGGGCGCTGCTTCTT-3', qPCR for the *tktA* gene |
| QTR    | 5'-CGACGCTGAAATTGGGCTGAC-3', qPCR for the *tktA* gene |
| QNF    | 5'-TGGAAATCAACCGTGGCAAATCT-3', qPCR for the *nakK* gene |
| QNR    | 5'-TGGAAATCAACCGTGGCAAATCT-3', qPCR for the *nakK* gene |
| QPF    | 5'-AGGGCGATGCAGGTTCCAGAC-3', qPCR for the *pntAB* gene |
| QPR    | 5'-ATTGCACCTGGTATTCCGCTGAG-3', qPCR for the *pntAB* gene |
| QPPF   | 5'-GACATCTTCTGCTGAGCAACAC-3', qPCR for the *ppsA* gene |
| QPPR   | 5'-TTACCGGCGGACCATCTTCT-3', qPCR for the *ppsA* gene |
Table 5 Primers used in this study (Continued)

| Name  | Sequences                                      | Description                          |
|-------|------------------------------------------------|--------------------------------------|
| QCSF  | 5'-CTGGATGAAGCGAAGGATG-3', qPCR for the csrB gene |                                      |
| QCSR  | 5'-ATTGGTCTCTGCTACACAC-3', qPCR for the csrB gene |                                      |
| QCF   | 5'-TTGTCGGCGGTGTTGATGTC-3', qPCR for the cysG gene |                                      |
| QCR   | 5'-ATGCGGTGAAGCTGGAAATACG-3', qPCR for the cysG gene |                                      |

*Restriction enzyme sites are underlined.

Conclusion

To overcome the drawbacks of plasmid-based expression systems, the aroG<sup>fr</sup>, aroB, aroE and ttk<sub>A</sub> gene cluster was integrated into the *E. coli* chromosome by direct transformation. The gene copy number was then evolved by their corresponding restriction enzymes and inserted into the chromosome of the *E. coli* strain. The effect of NADPH availability on SA production was also investigated. NADPH availability and SA production were found to be strongly correlated. Plasmid-based or chromosomal overexpression of the pnt<sub>AB</sub> or n<sub>adK</sub> genes enhanced the intracellular NADPH concentration and consequently the SA titer. This is the first report of an engineered SA-producing strain of *E. coli* that lacks both a plasmid and an antibiotic marker. Using this strain, no resistance-conferring compound was required during the fermentation process.

Methods

Strains, plasmids and primers

The strains and plasmids used in this study are listed in Table 4. *E. coli* DH5α was used for plasmid construction. The parent strain for SA production was *E. coli* BW2511 [28]. The primers used in this study are listed in Table 5.

Gene knockout and integration

Gene knockouts and replacement of the native promoter were carried out by PCR product recombination [28] using the pSIM6 plasmid [29] expressing the lambda red recombination system and pKD3 [28] as the template for PCR. Gene knockouts were verified by colony PCR using appropriate primers (listed in Table 5).

Chemically induced chromosomal evolution

The aro<sub>E</sub>, aro<sub>B</sub> and aro<sub>G</sub> genes were amplified from K12 genomic DNA using the corresponding primer pairs (Table 5) and cloned into pMD18-T to obtain pMD-aro<sub>E</sub>, pMD-aro<sub>B</sub> and pMD-aro<sub>G</sub>, respectively. The feedback-inhibition-resistant (fbr) aro<sub>G</sub> gene, which contained an Asp-146-Asn substitution, was obtained using the corresponding primer pairs (Table 5) and then inserted into the corresponding sites of pMP5 to form pMP5<sub>aroG</sub>. The p<sub>nt</sub>AB<sub>fr</sub> fragment was digested from pMD-aro<sub>G</sub><sup>fr</sup> and inserted into the PsI<sub>e</sub>/Sphi sites of pBEB to obtain pBEB<sub>G</sub>. The aro<sub>G</sub><sup>fr</sup>, aro<sub>B</sub> and aro<sub>E</sub> genes were amplified from pBEB using restriction enzymes EcoRI and Sphi. The resulting gene cluster was cloned into the *E. coli* chromosome by direct transformation, as described by Chen et al. [16] and Chiang et al. [14].

CIChE of the above construct was carried out by subculturing the resulting strains in 5 mL Super Optimal Broth (SOB) medium with increasing concentrations of triclosan in 15 mL culture tubes, as described by Chen et al. [16] and Tyo et al. [13]. The strains were grown to stationary phase in 1 μM triclosan. Fifty milliliters of the culture was subcultured into a new culture tube, in which the triclosan concentration was doubled from 1 to 2 μM and allowed to grow to stationary phase. The process was repeated until the desired concentration (as high as 32 μM) was reached. The rec<sub>A</sub> gene of the CIChE strain was then deleted.

Plasmid construction

The multiple cloning site (ATGCATGACGTCGGGCCC GCATGCCACGTGGAGCTCGGTACCATAAAAACGAA AGGCCGTAGTCGAAGACCTGGGCTTCGTTTTAT CAATTCGTCAAGCAGGGGCTTCGAGTCTAGAGTCG ACCCGGCG) was synthesized by Takara Biotechnology. This fragment was digested along with p5TG [32] using NsiI/SacI, and ligated with p5GT to form pMP5. The s<sub>th</sub>A, pnt<sub>AB</sub> and n<sub>adK</sub> genes were amplified from *E. coli* using the corresponding primer pairs (Table 5) and then inserted into the corresponding sites of pMP5 to form pMP<sub>sth</sub>A, pMP<p><sub>pnt</sub>AB</p> and pMP<p><sub>nad</sub>K</p>, respectively.

Chromosomal integration

The p<sub>ps</sub> and csr<sub>B</sub> gene cluster was amplified from pBAD-csr<sub>B</sub>-pps using the primers CsrF and PpsR and then inserted into the *SacI*/*KpnI* sites of pP21KT5b [33,34] to obtain pP21KT5b-csr<sub>B</sub>-pps.
To enhance the expression of the genes on the integrate vector, the T5 promoter of the integrate expression plasmid was replaced with the 5P_{tac} promoter (five tandem repeats of the core-tac-promoter) [32]. The 5P_{tac} promoter was amplified using primers 5TacF and 5TacR from p5TG [32]. The resulting fragment was digested along with pHKKT5b or pP21KT5b [33,34] using MluI and SphI, and was ligated into these plasmids to form pHKK5Tacb or pP21K5Tacb, respectively. The tktA gene was excised from pHKEBGT by SphI and SacI and ligated into the SphI/SacI sites of pP21K5Tacb to obtain the integrate vector pP21K5Tacb-tktA. The pntAB and nadK genes were excised from pMPpntAB and pMPnadk by their corresponding restriction enzymes and then inserted into the corresponding sites of pP21K5Tacb to form the integrate vectors pP21K5Tacb-pntAB and pP21K5Tacb-nadK, respectively.

Assisted by a helper plasmid, pH121, expressing the phage integrase [31], the above integrate vectors (pP21KT5b-csrB-pps, pP21K5Tacb-tktA, pP21K5Tacb-pntAB and pP21K5Tacb-nadK) were inserted into the bacterial attachment (attB) site of E. coli by direct transformation, as described by Chen et al. [16] and Chiang et al. [14].

Quantitative real-time PCR (qRT-PCR)
Total RNA from E. coli cells grown for 60 h in shake flasks was isolated using an RNA extraction kit (Dongsheng Biotech, Guangzhou, China), following the manufacturer's instructions. The first-strand cDNA was synthesized using an All-in-One™ First-Strand cDNA Synthesis Kit (GeneCopoeia, Guangzhou, China). The qRT-PCR was performed with the All-in-One™ qPCR Mix kit (GeneCopoeia) on an iCycler iQ5 Real Time PCR system (Bio-Rad Laboratories, California, USA). The template was 100 ng of cDNA. The PCR conditions were: 95°C for 10 min, followed by 45 cycles of denaturation at 95°C for 10 s, annealing at 60°C for 20 s, and extension at 72°C for 15 s. The primers for qRT-PCR are presented in Table 5. Data were analyzed by the 2^{-ΔΔCt} method described by Livak and Schmittgen [35], and normalized by cysG gene expression.

Gene copy numbers were measured by qPCR on genomic DNA isolated from the appropriate CiChE strains. qPCR was performed as above. The primers QEF and QER (Table 5) were used to measure the copy number of aroE.

SA production
For SA production, E. coli cells were precultured overnight in a falcon tube containing 5 mL Luria Broth (LB) at 37°C. The main cultures were incubated in the fermentation medium (pH 7.0) containing (g/L): glucose 10, peptone 1, proline 1.24, KH_{2}PO_{4} 3, Na_{2}HPO_{4} 7H_{2}O 13, MgCl_{2}-6H_{2}O 0.24, NaCl 0.5, NH_{4}Cl 1, CaCl_{2} 0.1. The main cultures were inoculated with a starting OD_{600} of 0.1 and incubated at 37°C for 72 h in a rotary shaking incubator at 150 rpm.

Assay
Cell growth was measured by optical density at 600 nm and converted to dry cell weight (DCW; g/L) using a standard curve. SA concentration was determined by HPLC using a Shimadzu system (LC-20A, Shimadzu, Japan) equipped with an Inertsil ODS-SP column (5 μm, 4.6 × 150 mm, GL Sciences Inc, Tokyo, Japan). The mobile phase was 5 mmol/L H_{2}SO_{4}, with a flow rate of 0.5 mL/min, at 50°C. SA was detected by a photodiode array detector (SPD-M20A) operating at 210 nm, and quantified by a standard curve constructed from serial dilutions of a SA standard stock solution (1 mg/ml, J&K Scientific Ltd, Beijing, China). Glucose concentration was determined by glucose oxidase using a glucose assay kit (Shanghai Rongsheng Biotech Corporation, Shanghai, China).

NADPH assay
Following aerobic cultivation of E. coli on a rotary shaker (150 rpm) at 37°C for 54 h, the cells were harvested by centrifugation and washed twice with water. Intracellular NADPH was extracted and quantified using the Enzychrom™ NADP+/NADPH Assay kit (BioAssay Systems, Hayward, CA), following the manufacturer’s instructions.

Statistical analysis
All experiments were conducted in triplicate, and data were averaged and presented as the mean ± standard deviation. Significant differences were determined by one-way analysis of variance followed by Tukey's test, using the OriginPro (version 7.5) package. Statistical significance was defined as p < 0.05.

Additional file

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
YYC carried out most of the experiments. CL amplified some of the genes. YYZ constructed some of the plasmids. JH deleted the aroI and aroL genes. JZL developed the concept and designed the method, led the project and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgments
We are grateful to both the National Natural Science Foundation of China (Grant No. 30970089, 20876181, 21272689), the Natural Science Foundation of Guangdong Province (No. S201101001396) and the Key Project of the Scientific and Technical Innovation of Higher Education of Guangdong Province (2012CXZD0002) for their financial support.
Author details

1 Biotechnology Research Center and Biomedical Center, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, P R China. *Current address: School of Materials Science and Engineering, South China University of Technology, Guangzhou 510641, P R China.

Received: 20 November 2013 Accepted: 7 February 2014

References

1. Kim CU, Lew W, Williams MA, Liu H, Zhang L, Swaminathan S, Bischoffberger N, Chen WS, Mendel DB, Tai CY, Laver WG, Stevens RC: Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carboxylic sialic acid analogs with potent anti-influenza activity. J Am Chem Soc 1997, 119:681–690.

2. Krämer M, Bongaerts J, Bovenberg R, Kremer S, Müller U, Orf S, Wubbolts M, Raeven L: Producing shikimate on different carbon sources. Appl Microbiol Biotechnol 2011, 324:1–27.

3. Escalante A, Calderon R, Valdivia A, de Anda R, Hernandez G, Hernandez OT, Gosset G, Bolivar F: Metabolic engineering for the production of shikimic acid in an evolved Escherichia coli strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. Microb Cell Fact 2010, 9:21.

4. Ahrn J, Chung BS, Lee D-Y, Park M, Karimia IA, Jung J-K, Lee H: NADPH-dependent pgk-gene knockout Escherichia coli metabolism producing shikimate on different carbon sources. FEBS Microbiol Lett 2011, 324:10–16.

5. Ramsay LM, Belin D, Carson MJ, Beckwith J: One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 2000, 97:6460–6465.

6. Tupkino IA, Kuznetsov SG, Court DL: Recombining: a homologous recombination-based method of genetic engineering. Nat Protoc 2009, 4:206–223.

7. Miller J, Wang J, Geng Y, Li Y, Wang Q, Liang Q, Qi Q: High NADPH/NADP ratio improves thymidine production by a metabolically engineered Escherichia coli strain. J Biotechnol 2010, 149:24–32.

8. Liu J, Zhou L, Zhong X, Ma Y: Activating transhydrogenase and NAD kinase in combination for improving isobutanol production. Metab Eng 2013, 16:1–10.

9. Datserka KO, Wanner BL: One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 2000, 97:6460–6465.

10. Sharan SK, Thomason LC, Kuznetsov SG, Court DL: Recombining: a homologous recombination-based method of genetic engineering. Nat Protoc 2009, 4:206–223.

11. Millard M, Belin D, Carson MJ, Beckwith J: Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 1995, 177:4121–4130.

12. Haldimann A, Wanner BL: Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. J Bacteriol 2001, 183:6384–6393.

13. Lu M, Wang J, Geng Y, Li Y, Wang Q, Liang Q, Qi Q: A strategy of gene overexpression based on tandem repetitive promoters in Escherichia coli. Microb Cell Fact 2012, 11:19.

14. Liu J, Huang MT, Cui Y, Chen YY: A series of expression plasmids for chromosomal integration and evolution. China patent 201210060042.

15. Huang MT, Chen YY, Liu JZ: Chromosomal engineering of Escherichia coli for the efficient production of coenzyme Q10. Chin J Chem Eng. In press.

16. Li, Vink K, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT Method. Methods 2001, 25:402–408.

Cui et al. Microbial Cell Factories 2014, 13:21
http://www.microbialcellfactories.com/content/13/1/21

Page 11 of 11

doI:10.1186/1475-2859-13-21

Cite this article as: Cui et al.: Production of shikimic acid from Escherichia coli through chemically inducible chromosomal evolution and cofactor metabolic engineering. Microbial Cell Factories 2014 13:21.