**Engineering of global regulator cAMP receptor protein (CRP) in *Escherichia coli*
for improved lycopene production**

Lei Huang¹, Yue Pu¹, Xiuliang Yang², Xiangcheng Zhu³, Jin Cai¹, Zhinan Xu^{1*}

¹ Institute of Bioengineering, Department of Chemical and Biological Engineering, Zhejiang University,
Hangzhou, China

² Shangdong Jincheng Biopharmaceutical Corporation Limited, Zibo, China

³ Hunan Engineering Research Center of Combinatorial Biosynthesis and Natural Product Drug
Discovery, Changsha, China

*Corresponding author: Institute of Bioengineering, Department of Chemical and Biological
Engineering, 38 Zheda Road, Zhejiang University, Hangzhou 310027, China. Fax: +86 571 87951220.
E-mail address: znxu@zju.edu.cn

Abstract

Transcriptional engineering has received significant attention for improving strains by modulating the behavior of transcription factors, which could be used to reprogram a series of gene transcriptions and enable multiple simultaneous modifications at the genomic level. In this study, engineering of the cAMP receptor protein (CRP) was explored with the aim of subtly balancing entire pathway networks and potentially improving lycopene production without significant genetic intervention in other pathways. Amino acid mutations were introduced to CRP by error-prone PCR, and three variants (mcrp26, mcrp159 and mcrp424) with increased lycopene productivity were screened. Combinations of three point mutations were then created via site-directed mutagenesis. The best mutant gene (mcrp26) was integrated into the genome of *E. coli* BW25113-BIE to replace the wild-type *crp* gene (MT-1), which resulted in a higher lycopene production (18.49 mg/g DCW) compared to the original strain (WT). The mutant strain MT-1 was further investigated in a 10-L bench-top fermentor with a lycopene yield of 128 mg/l at 20 h, approximately 25% higher than WT. DNA microarray analyses showed that 396 genes (229 up-regulated and 167 down-regulated) were differentially expressed in the mutant MT-1 compared to WT. Finally, the introduction of the mutant *crp* gene (mcrp26) increased β -carotene production in *E. coli*. This is the first report of improving the phenotype for metabolite overproduction in *E. coli* using a CRP engineering strategy.

Keywords *Escherichia coli*; cAMP receptor protein; error-prone PCR; transcriptional engineering; lycopene

1. Introduction

An important methodology in the biotechnology industry is the production of desirable chemicals from microbes, which is becoming preferred over traditional chemical routes due to the mild reaction conditions and the expected specificity (Martin et al., 2003). However, engineering the production of strains and the development of economical production processes have significantly challenged the commercialization of microbial-derived products (Aristidou and Penttila, 2000). To improve microbial performance, two types of approaches, “random approaches” (Stemmer, 1994; Zhang et al., 1997) and “rational approaches” (Montemiglio et al., 2013; Chen et al., 1996), have been applied to screen new phenotypes. Due to its simplicity and convenience, random mutagenesis has been widely used to optimize strains, although it is a time-consuming and labor-intensive process. However, rational engineering of strain phenotypes requires comprehensive knowledge of complex metabolic networks and a clear structural model of target cells (Patnaik, 2008), which is impractical for common strain improvement.

Current cellular and metabolic engineering approaches almost exclusively rely on the deletion or over-expression of certain genes, which could lead to the accumulation of intermediates or unexpected side products. Because various metabolites in the metabolic networks are interconnected through a large number of biochemical and regulatory reactions, engineering a desired phenotype should be

1 enormously facilitated by multiple simultaneous gene modifications (Alper and
2 Stephanopoulos, 2007).

3 Transcriptional engineering approaches have received much attention for strain
4 improvement in recent years, which relies on the modifications of transcription factor
5 behaviors to reprogram a series of gene transcriptions and enables multiple
6 simultaneous perturbations at the genomic level. It has been reported that strain
7 phenotypes can be improved by engineering RNA polymerase subunits (Alper and
8 Stephanopoulos, 2007; Alper et al., 2006; Klein-Marcuschamer et al., 2009),
9 transcription factors such as zinc finger proteins (Park et al., 2003), H-NS (Hong et al.,
10 2010b), Hha (Hong et al., 2010a), IrrE (Chen et al., 2011), or cAMP receptor protein
11 (CRP) (Chong et al., 2013a). CRP is a well-known global transcription factor that
12 regulates the expression of more than 400 genes belonging to different functional
13 groups of *E. coli* (Shimada et al., 2011). The CRP-cAMP complex can act as a dual
14 regulator by inducing sequence-specific DNA recognition within or near the
15 promoters of genes (Won et al., 2008; Fic et al., 2006). In fact, 95% of regulators in *E.*
16 *coli* are controlled by CRP through direct regulation or co-regulation
17 (Martinez-Antonio and Collado-Vides, 2003). This offers a unique opportunity to
18 induce a simultaneous global transcription-level alteration that could potentially
19 impact cellular properties. The engineering of CRP has been applied to improve
20 various endurance-related phenotypes of *E. coli*, such as the tolerance to biofuel
21 toxicity (Chong et al., 2014; Zhang et al., 2012b; Chong et al., 2013a), osmotic
22 pressure (Zhang et al., 2012a), acetate or low pH (Chong et al., 2013b; Basak et al.,
23 2014), organic solvent toxicity (Basak et al., 2012) and oxidative stress (Basak and
24 Jiang, 2012).

25 However, beyond the tolerance phenotypes, the reframing of the entire *E. coli*
26 transcriptional machinery through CRP engineering to improve the metabolite
27 overproduction has not been reported. Similarly, improving metabolite production
28 should also require multiple simultaneous gene expression changes. Due to the
29 convenient colorimetric screening property of lycopene, its production in *E. coli* has
30 been chosen as a representative metabolic phenotype. The production of lycopene in *E.*

E. coli involves many primary metabolic networks such as the isoprenoid pathway (Jin and Stephanopoulos, 2007; Kang et al., 2005) and other peripheral biosynthetic and housekeeping pathways. The genetic manipulation of only localized branches of the isoprenoid pathway was inefficient in improving lycopene production (Matthews and Wurtzel, 2000). To increase the yield of lycopene, multiple factors should be considered, such as the availability of substrates, the growth inhibition caused by excess intracellular intermediates, and the possible cellular toxicity of accumulated lycopene (Sandmann et al., 1999). Lycopene production is a complex phenotype regulated by various genes, some of which remain unknown or unreachable.

In this study, we explored the possibility that the engineering of CRP can potentially lead to a subtle balance of pathway networks and thus improve the production of lycopene without significant intervention on local enzymes or pathways. Error-prone PCR was applied to introduce mutations into CRP, and the mutants with increased lycopene yield were screened. The best CRP engineering strain derived from *E. coli* BW25113-BIE was further investigated in a 10-L bench-top fermentor. The transcriptional changes between the mutant and wild type were also analyzed using DNA microarrays.

2. Materials and methods

2.1. Strains and plasmids

E. coli DH5 α was used as the host for genetic manipulation. *E. coli* BW25113-BIE and BW25113-CARO were provided by Professor Albermann as gifts (Albermann et al., 2010). A low-copy number plasmid pKSC was derived from pUC18, in which the ampicillin resistant gene was replaced by the kanamycin resistant gene from pET-39b (+) and integrated with the SC101 origin of replication from pSC101 (Zhang et al., 2012a). The plasmid pUC18-*crp* containing the native *crp* gene was constructed and served as the template for error-prone PCR (Zhang et al., 2012a).

2.2. Enzymes and culture media

The restriction enzymes, LA Taq DNA polymerase, and T4 DNA ligase were purchased from Takara (Japan). DNA gel extraction kits, PCR cleanup kits and spin miniprep kits were purchased from Axygen (USA). GeneMorph II Random Mutagenesis Kit was purchased from Stratagene (La Jolla, CA). Luria-Bertani (LB) medium (w/v) containing 0.5% yeast extract, 1% tryptone and 1% NaCl was used as culture media. SOC medium (yeast extract 5 g/l, tryptone 20 g/l, NaCl 10 mM, KCl 2.5 mM, MgCl₂ 10 mM, and glucose 20 mM). MBL medium (w/v) contained 0.5% glucose, 3% yeast extract, 2% tryptone, 0.35% (NH₄)₂HPO₄, 0.35% KH₂PO₄, 0.5% K₂HPO₄, 0.7% MgSO₄ · 7H₂O, and 2.1% NaCl.

2.3. Random and site-directed mutagenesis

Random mutagenesis of CRP was performed using a GeneMorph II Random Mutagenesis Kit. The primers *crp_sense* and *crp_anti* are listed in Table 1. The error-prone PCR was performed with 30 ng pUC18-*crp* using the following program: 3 min at 95 °C, 30 cycles of 45 s at 95 °C, 45 s at 62 °C and 1 min at 72 °C, followed by another 10 min at 72 °C. The resulted fragments were recovered using the QIAquick Gel Extraction Kit and digested by *Bam*H I and *Kpn* I overnight. These fragments were then ligated into the plasmid pKSC via the same sites and transformed into *E. coli* BW25113-BIE competent cells via electroporation to construct the CRP mutant library.

Combinations of point mutations were created by site-directed mutagenesis using the appropriate primers (Table 1) and the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. The resulted plasmids were also transformed into *E. coli* BW25113-BIE, and the mutations were verified via DNA sequencing.

2.4. Phenotype screening of CRP mutants

The mutant library was incubated in 1 mL SOC medium at 37 °C for 3 h. Then, the culture was spread onto the LB agar plates containing 30 µg/mL kanamycin and cultivated at 30 °C for 48 h. The isolated single colony was grown in 5 mL LB medium supplemented with 30 µg/mL kanamycin at 30 °C for 24 h. The cells were harvested from the 1 mL culture by centrifuging at 16,000 g for 30 s and were

1 resuspended in 800 µl acetone. After incubating in darkness for 15 min, the mixture
2 was centrifuged at 16,000 g for 30 sec, and the supernatant was transferred into a
3 clean tube for lycopene quantification. The absorbance of the extracted lycopene
4 solution at 472 nm was measured using a spectrophotometer and calibrated against the
5 lycopene standard (Sigma-Aldrich) to determine the lycopene content. The plasmid
6 containing mutated *crp* gene was recovered from the screened mutant and
7 retransformed into fresh *E. coli* BW25113-BIE to validate that the enhanced lycopene
8 production phenotype was indeed endowed by the engineered *crp* gene. Dry cell
9 weight was determined by baking a washed cell pellet at 105 °C for 24 h in the dark.
10 Experiments were conducted in triplicate to minimize errors and all presented results
11 are average values.

12 2.5. Chromosomal integration of engineered *crp*

13 To replace the wild-type *crp* gene in the genome, the engineered target *crp* was
14 amplified from the isolated mutant via high fidelity PCR using PrimeSTAR® HS
15 DNA polymerase (TAKARA, Japan) and the primers *crp_red_sense* / *crp_red_anti*
16 (Table 1). The purified *crp* fragment was transformed into *E. coli* BW25113-BIE
17 harboring plasmid pKD46. The expression of λ-red enzymes and the preparation of
18 competent cells were carried out as previously described (Sharan et al., 2009).
19 Competent cells were electroporated with 0.2-0.4 µg purified *crp*, and the mixture
20 was resuspended in 1 mL LB medium. After incubating at 30 °C over 6 h with
21 shaking, the cell culture was spread onto LB agar plates containing 30 mg/L
22 kanamycin and incubated at 37 °C overnight. Five to ten colonies per transformation
23 were chosen and validated using PCR to confirm the integration of the engineered *crp*
24 gene.

25 2.6. Batch culture in stirred tank fermentor

26 Next, 10% seed culture (v/v) was inoculated into a 10-L stirred tank fermentor
27 containing 6 L MBL medium, and cultivated at 30 °C with 30 L/min aeration rate and
28 200 rpm stirring speed. The pH of the fermentation broth was maintained at
29 approximately 7.0 via automatic addition of ammonia solution. Culture samples were

collected every two hours to analyze the cell growth and lycopene production. The batch fermentation was terminated after 30 hours of incubation.

2.7. DNA Microarray Analysis

Total cellular RNA was extracted from the cells grown for 15 h using the RNeasy Mini Kit (Qiagen) and treated with RNase-free DNase I (Qiagen). Microarrays were performed using the AffymetrixGenechip *E. coli* Genome 2.0 array. The slides were scanned on a GeneArray® Scanner, and the data were analyzed using Bioconductor software.

3. Results

3.1. Isolation of CRP Variants

After the mutated *crp* gene library was introduced into the *E. coli* BW25113-BIE, the transformants producing intense red pigmentation were selected to recover the plasmids, and the mutation sites in the *crp* gene were confirmed via DNA sequencing. These validated plasmids were then retransformed into fresh *E. coli* BW25113-BIE to eliminate false positives. The plasmid pKSC-*crp* containing the original *crp* gene was introduced into *E. coli* BW25113-BIE as a positive control strain. Three positive recombinant strains containing with different mutated *crp* genes were obtained and denoted M₁ (mcrp26), M₂ (mcrp159) and M₃ (mcrp424) (Table 2).

The lycopene production in these recombinant strains was evaluated. Compared to the positive control that yielded 19.39 mg/g DCW lycopene, M₁, M₂ and M₃ produced 21.74 mg/g, 20.56 mg/g and 20.17 mg/g DCW lycopene, respectively, which correspond to approximately 12.1%, 6% and 4% higher lycopene than the control (Fig. 1). The three point mutations were then combined to create three additional mutants (M₄ (mcrp26/159), M₅ (mcrp26/424) and M₆ (mcrp26/159/424)) to evaluate if the mutations have synergetic effects. M₄, M₅, and M₆ produced 20.36 mg/g, 20.49 mg/g and 20.22 mg/g DCW lycopene, respectively (Fig. 1).

To eliminate possible influence by the backbone of the introduced plasmid on lycopene production, the best strain, M₁, was selected to recover the mutated *crp* gene fragment (mcrp26) for chromosomal integration. After PCR verification, the resulting BW25113-BIE with mcrp26 fused into the genome was denoted “MT-1”, which

produced 18.49 mg/g DCW lycopene, approximately an 8% increase compared to the wild type (17.12 mg/g DCW) (Fig. 2). Thus, the obtained results indicated that the engineered CRP was able to regulate the transcription of genes related to lycopene metabolic pathways. Considering genetic stability, the *crp* integration strain MT-1 was further investigated.

3.2. Growth conditions of the scaled-up culture

The scaled-up production of lycopene was carried out in a 10-L bench-top fermentor. Based on the fermentation profiles, the biggest difference in lycopene production between the WT and MT-1 was approximately 25% after 20 h of cultivation, at which time the cell density of the WT was slightly higher than that of MT-1 (Fig. 3). Thus, under the better scaled-up culture conditions, the effect of engineered CRP on the lycopene production was also magnified.

3.3. Transcriptional analysis of MT-1

To identify possible reasons for the enhanced lycopene production caused by engineered CRP, the transcriptome of MT-1 and WT were analyzed via DNA microarrays (GEO accession no. **GSE59067**). The analyses indicated that the mutated CRP simultaneously changed to the expression of multiple genes (Table S1 in File S1). Specifically, total 396 genes (229 up-regulated and 167 down-regulated) were differentially expressed in MT-1 comparing to those in WT, with P-value threshold of 0.05. Among them, only 35 genes were directly regulated by CRP, of which 20 genes had been up-regulated and 15 genes down-regulated. Approximately 91% of the differentially expressed genes were indirectly regulated by CRP. These results suggest that CRP could influence the expression of other genes through complicated metabolic networks. Moreover, the MEP pathway derives five-carbon isoprene units, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are the building blocks of lycopene. The biosynthetic pathway from the initial substrate glucose to these common precursors involves 18 genes (*pgi*, *pfkA*, *fbaA*, *tpiA*, *gapA*, *pgk*, *gpm*, *eno*, *pykAF*, *dxs*, *dxr*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*), among them only three genes were up-regulated (*pfkA* 0.3-fold, *fbaA* 0.1-fold, and *ispG* 0.3-fold) in MT-1, and none of them were down-regulated.

Further investigations of transcriptional levels using gene ontology analyses revealed that certain genes involved the transportation system (e.g., *araE* and *nikA*), and biofilm formation (e.g., *yfaZ*, *hofB*, and *yihR*) or associated with oxidation-reduction enzymes (e.g., *yggW*, *yqhD*, *ygcU*, *wecC* and *wecG*) and other regulators (e.g., *araC*, *yeiL*, *csrB*, *hns* and *creB*) were varied in the mutant MT-1 (Table S1), which enabled the reframing of genetic control circuits by modifying the entire genomic hierarchy.

3.4. Cross enhancement of other carotenoids production

The effect of engineered CRP on the production of other carotenoids such as β -carotene was further investigated. Using the same strategy, the CRP mutant of the β -carotene-producing strain BW25113-CARO was obtained and evaluated. The results showed that the mutation CRP mutation could also enhance the production of β -carotene, which was approximately 15% improved compared to the control (Fig. 4). Thus, the engineered CRP may be applied for improving the production of other carotenoids in addition to lycopene.

4. Discussion

In this study, we demonstrated that it is possible to enhance lycopene production from the transcription level via the mutation of the global transcription factor CRP. Random mutagenesis of the *crp* gene were generated using error-prone PCR. The CRP variants associated with enhanced lycopene production were isolated by screening the colonies with relatively increased pigmentation compared with the control.

CRP is a homodimeric protein with each subunit consisting of three domains: the N-terminal domain (residues 1-134) is responsible for CRP dimerization and for interacting with the allosteric effector cAMP; the C-terminal domain (residues 140-209) is responsible for interacting with DNA; and the flexible hinge (residues 135-139) covalently couple these two domains (Martinez-Antonio and Collado-Vides, 2003; Zhou and Yang, 2006). Some tolerant phenotypes of *E. coli* have been successfully achieved by the engineering the global regulator CRP (Table 3). Unlike the direct evolution of enzymes in which the amino acid mutations are always located

1 around the active site of enzyme, the mutations in CRP were relatively decentralized
2 (Table 2 and Table 3). Because many crucial interactions between transcription
3 factors and target genes are fairly weak and nonspecific, even a small change can
4 exert significant influence on transcription regulation (Bougdour et al., 2008). In our
5 work, the best isolated lycopene-producing mutant MT-1(M1) had only one amino
6 acid substitution at residue D8 (D8V), which belongs to the N-proximal cAMP
7 binding domain. It was reported that the cAMP-binding capacity could influence all
8 three types of CAP-dependent promoters (Busby and Ebright, 1999), which may alter
9 the transcription of genes related to lycopene production. The amino acid substitutions
10 in M2 and M3 were K52N and G141S, respectively. Residue K52 was reported to
11 have a negative effect on the transcription activation of class II CRP-dependent
12 promoters (Rhodius and Busby, 2000), whereas substitution of the residue with a
13 neutral or negatively charged amino acid (K52N, K52D, or K52L) would improve the
14 binding of CRP to a CRP-dependent promoter (Williams et al., 1991). G141 is located
15 in the D α -helix, which is near the hinge (135-139) and faces the F α -helix (Harman,
16 2001; Kim et al., 1992). The activation of CRP is typically demonstrated to be
17 allosteric, where the cAMP-binding signals are transmitted from the N-terminal
18 domain to the C-terminal domain through the hinge. Based on the crystal structure
19 and mutational analyses, it is proposed that the binding of cAMP aligns the subunits
20 of the CRP dimer by contacting the large domains of both subunits in the middle of
21 the C α -helix. These contacts transmit a change to the hinge end of the D α -helix,
22 altering the relative orientations of the large and small domains by changing the hinge
23 angle. These combinative motions ultimately position the F α -helix for specific
24 contact with the DNA surface (Kolb et al., 1993). Thus, the G141S substitution may
25 affect the transmission of cAMP-binding signals and the affinity between CRP and
26 DNA.

27 The diverse phenotypes of microbial strains are derived from complicated
28 interactions among thousands of gene expression products, which are regulated by
29 different transcription factors. Modifications of these transcriptional factors change
30 their DNA binding specificity, their partner proteins, and their influence on

transcription (Hobert, 2008). Our DNA microarray analyses data revealed that the simple mutation of trans-acting transcription factor CRP reprogramed the transcription profile and altered the cell phenotype accordingly. In our present work, the possible influence of several important genes, which were regulated by the engineered CRP, on the production of lycopene was investigated (Table 4). These genes may be classified to four groups based on their effects on lycopene production in *E. coli*: 1) the availability of substrates (*pfkA*, *fbaA*, *ispG*), 2) the energy supply (*appY*, *araE*), 3) the resistance to toxicity (*rpoS*, *yjiD*), and 4) other regulators (*araC*, *yeiL*, *csrB*, *hns* and *creB*).

The *pfkA* encoded 6-phosphofructokinase I is a key enzyme regulating the glycolysis pathway. At the G6P (glucose-3-phosphate) node of the glycolysis pathway, overproduction of 6-phosphofructokinase I readjusted the metabolic flux trending towards G3P (glyceraldehyde-3-phosphate) instead of PPP (pentose phosphate pathway), and the resulting G3P pool led to enhanced lycopene production. Previous studies also indicated that the precursor balance between pyruvate and G3P plays a major role in the high-level production of isoprenoids in *E. coli* (Choi et al., 2010). Similar to *pfkA*, the *fbaA* gene also served the same function in balancing the substrates. Thus, overexpression of *pfkA* and *fbaA* could increase G3P levels and thereby improve the production of lycopene, which has been reported in previous studies (Choi et al., 2010; Farmer and Liao, 2001). However, *ispG* encodes 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase, which is involved in the sixth step of the methylerythritol phosphate (MEP) pathway. It has been reported that the enhanced expression of *ispG* effectively reduces the efflux of methylerythritol cyclodiphosphate (MEC) inside cells (Zhou et al., 2012), leading to a significant increase in downstream isoprenoid production.

The up-regulated gene *appY* (0.6-fold) encodes a transcriptional activator of two energy metabolism operons *hya* and *cbdAB-appA*, which are induced by anaerobiosis. During aerobic respiration, ubiquinone is an important component of the electron transport chain for energy production. It is composed of a quinine ring and an isoprenyl side chain containing seven molecules of IPP and one DMAPP molecule

(Kainou et al., 2001; Okada et al., 1997). Because the biosyntheses of ubiquinone and lycopene share the same building block molecules of IPP and DMAPP, cells may not be able to synthesize a normal level of ubiquinone in the presence of active lycopene production. Blockage of ubiquinone synthesis has been shown to increase the expression of the *cbdAB-appA* operon (Georgellis et al., 2001). Hence, the up-regulated *appY* gene could induce the expression of *hya* and *cbdAB-appA*, sequentially arousing energy production via the anaerobic metabolic pathway to complement the energy deficiency incurred due to low intracellular levels of ubiquinone, and thus ensuring the overproduction of lycopene.

The MEP pathway is NADPH-dependent, requiring 16 NADPH to produce one lycopene. *AraE* is an arabinose/proton symporter responsible for the uptake of arabinose with low affinity (140-320 μ M) (Daruwalla et al., 1981). Imported arabinose can be catabolized to xylulose-5-phosphate and serve as a substrate for the pentose phosphate pathway, which is an important pathway providing NADPH. Under active lycopene producing condition, G3P competes with PPP for the metabolic flux, so the increased level of xylulose-5-phosphate would supplement the PPP cycle and thus provide efficient NADPH for lycopene production.

The toxicity of high concentrations of lycopene towards cells has been previously observed. It has been documented that carotenoids may accumulate in the cellular membrane bilayer and break the membrane integrity (Sikkema et al., 1995; Sung et al., 2007), in which the functions of the membrane as a matrix for embedded proteins and as a selective barrier were significantly impaired. Thus, the enhanced production of lycopene caused by the engineered CRP may be not only a result of expression level changes in lycopene biosynthesis-related enzymes but also due to an increased stress resistance to intracellular accumulation of hydrophobic lycopene. The *rpoS* encoded stationary phase sigma factor σ S stimulated the transcription of genes expressed in the stationary phase or related to stress responses (Table 4). The inactivation of *rpoS* in *E. coli* resulted in an 85% reduction in carotenoid formation (Sandmann et al., 1990). The protein YjiD inhibited the proteolysis of factor σ S by interacting with RssB, which normally targets σ S for degradation via ClpXP protease

(Bougdour et al., 2008). In addition, YjiD is a major determinant for the steady-state level of RpoS (Merrikh et al., 2009). Therefore, the up-regulation of rpoS and yjiD should to correlate with enhanced lycopene production.

In addition, CRP regulates other transcriptional regulators such as AraC, YeiL, CsrB, H-NS and CreB. Both AraC and YeiL are co-regulators of CRP, whereas CsrB facilitates RNA binding to CsrA to form an RNA-protein complex, which positively regulates G6P isomerase. The H-NS protein plays an important role in the regulation of many genes in response to environmental changes and adaptation to stress. The carbon source-responsive regulator CreB is a DNA-binding transcriptional dual regulator and belongs to the CreBC two-component system (TCS), which controls genes involved in the pentose phosphate pathway. This may explain why CRP regulated 91% of the differentially expressed genes in *E. coli* mutant with enhanced lycopene production.

The physiological pathways of microbial organisms are too sophisticated to decipher the relationship between genes and desired phenotypes. Fortunately, the microarray analyses of transcriptionally engineered strains could reveal some clues regarding the mysterious gene functions associated with the newly acquired important traits. For those most highly up-regulated genes (such as *yghD*, *flgF*, *nikA*, *hisC*, *yeiL*, and *yibD*) and down-regulated genes (such as *malE*, *yjiO*, *yjiU*, and *ybcV*), it is thought that they may be involved in the accumulation of lycopene in *E. coli* (Table S1). For example, YghD is a predicted secretion pathway protein, which may contribute to the release of lycopene from the inner membrane to alleviate the toxicity of lycopene against cell. As the third member of the CRP/FNR family, YeiL contains the main conserved structural elements and may serve as an auxiliary regulator to facilitate CRP/FNR regulation of other genes. However, the exact roles of these highly regulated genes in lycopene biosynthesis require further investigation.

In conclusion, the engineered CRP may directly or indirectly alter the metabolic pathways related to lycopene biosynthesis and increase lycopene production in *E. coli* by reprogramming gene networks and metabolism. Transcriptional control of gene expression has been widely adapted in metabolic engineering and synthetic biology

1 applications to optimize biological systems such as metabolic pathways and genetic
2 circuits. In our work, the implementation of CRP engineering highlighted the
3 potential of this approach for building microbial cell factories with improved titers of
4 other chemicals or biofuels.

5
6 **Acknowledgments** This work was financially supported by the National Natural
7 Science Foundation of China (21006088), the National High Technology Research
8 and Development Program of China (2012AA022105A & 2011AA02A114), the
9 Natural Science Foundation of Zhejiang province of China (Y4100344), the
10 Specialized Research Fund for the Doctoral Program of Higher Education of China
11 (20100101120030), the Fundamental Research Funds for the Central Universities
12 (2014QNA4024).

13 14 **Reference**

- 15 Albermann, C., Trachtmann, N., Sprenger, G.A., 2010. A simple and reliable method
16 to conduct and monitor expression cassette integration into the *Escherichia coli*
17 chromosome. *Biotechnol. J.* 5, 32-38.
- 18 Alper, H., Moxley, J., Nevoigt, E., Fink, G.R., Stephanopoulos, G., 2006. Engineering
19 yeast transcription machinery for improved ethanol tolerance and production.
20 *Science* 314, 1565-1568.
- 21 Alper, H., Stephanopoulos, G., 2007. Global transcription machinery engineering: A
22 new approach for improving cellular phenotype. *Metab. Eng.* 9, 258-267.
- 23 Aristidou, A., Penttilä, M., 2000. Metabolic engineering applications to renewable
24 resource utilization. *Curr. Opin. Biotechnol.* 11, 187-198.
- 25 Basak, S., Geng, H., Jiang, R., 2014. Rewiring global regulator cAMP receptor
26 protein (CRP) to improve *E. coli* tolerance towards low pH. *Journal of*
27 *Biotechnology* 173, 68-75.
- 28 Basak, S., Jiang, R.R., 2012. Enhancing *E. coli* Tolerance towards Oxidative Stress
29 via Engineering Its Global Regulator cAMP Receptor Protein (CRP). *PLoS One*
30 7, e51179.
- 31 Basak, S., Song, H., Jiang, R.R., 2012. Error-prone PCR of global transcription factor
32 cyclic AMP receptor protein for enhanced organic solvent (toluene) tolerance.
33 *Process Biochem.* 47, 2152-2158.
- 34 Boudour, A., Cunniff, C., Baptiste, P.J., Elliott, T., Gottesman, S., 2008. Multiple
35 pathways for regulation of sigma(S) (RpoS) stability in *Escherichia coli* via the
36 action of multiple anti-adaptors. *Mol. Microbiol.* 68, 298-313.
- 37 Busby, S., Ebright, R.H., 1999. Transcription activation by catabolite activator protein

1 (CAP). Journal of Molecular Biology 293, 199-213.

2 Chen, R.D., Greer, A., Dean, A.M., 1996. Redesigning secondary structure to invert
3 coenzyme specificity in isopropylmalate dehydrogenase. Proc. Natl. Acad. Sci.
4 U. S. A. 93, 12171-12176.

5 Chen, T.J., Wang, J.Q., Yang, R., Li, J.C., Lin, M., Lin, Z.L., 2011.
6 Laboratory-Evolved Mutants of an Exogenous Global Regulator, IrrE from
7 *Deinococcus radiodurans*, Enhance Stress Tolerances of *Escherichia coli*. PLoS
8 One 6, e16228.

9 Choi, H.S., Lee, S.Y., Kim, T.Y., Woo, H.M., 2010. In Silico Identification of Gene
10 Amplification Targets for Improvement of Lycopene Production. Appl. Environ.
11 Microbiol. 76, 3097-3105.

12 Chong, H.Q., Geng, H.F., Zhang, H.F., Song, H., Huang, L., Jiang, R.R., 2014.
13 Enhancing *E. coli* Isobutanol Tolerance Through Engineering Its Global
14 Transcription Factor cAMP Receptor Protein (CRP). Biotechnol. Bioeng. 111,
15 700-708.

16 Chong, H.Q., Huang, L., Yeow, J.W., Wang, I., Zhang, H.F., Song, H., Jiang, R.R.,
17 2013a. Improving Ethanol Tolerance of *Escherichia coli* by Rewiring Its Global
18 Regulator cAMP Receptor Protein (CRP). PLoS One 8, e57628.

19 Chong, H.Q., Yeow, J., Wang, I., Song, H., Jiang, R.R., 2013b. Improving Acetate
20 Tolerance of *Escherichia coli* by Rewiring Its Global Regulator cAMP Receptor
21 Protein (CRP). PLoS One 8, e77422.

22 Daruwalla, K.R., Paxton, A.T., Henderson, P.J.F., 1981. Energization of the
23 transport-systems for arabinose and comparison with galactose transport in
24 *Escherichia coli*. Biochem. J. 200, 611-627.

25 Farmer, W.R., Liao, J.C., 2001. Precursor balancing for metabolic engineering of
26 lycopene production in *Escherichia coli*. Biotechnol. Prog. 17, 57-61.

27 Fic, E., Polit, A., Wasylewski, Z., 2006. Kinetic and structural studies of the allosteric
28 conformational changes induced by binding of cAMP to the cAMP receptor
29 protein from *Escherichia coli*. Biochemistry 45, 373-380.

30 Georgellis, D., Kwon, O., Lin, E.C.C., 2001. Quinones as the redox signal for the Arc
31 two-component system of bacteria. Science 292, 2314-2316.

32 Harman, J.G., 2001. Allosteric regulation of the cAMP receptor protein. Biochimica
33 Et Biophysica Acta-Protein Structure and Molecular Enzymology 1547, 1-17.

34 Hobert, O., 2008. Regulatory logic of neuronal diversity: Terminal selector genes and
35 selector motifs. Proc. Natl. Acad. Sci. U. S. A. 105, 20067-20071.

36 Hong, S.H., Lee, J., Wood, T.K., 2010a. Engineering global regulator Hha of
37 *Escherichia coli* to control biofilm dispersal. Microb. Biotechnol. 3, 717-728.

38 Hong, S.H., Wang, X.X., Wood, T.K., 2010b. Controlling biofilm formation,
39 prophage excision and cell death by rewiring global regulator H-NS of
40 *Escherichia coli*. Microb. Biotechnol. 3, 344-356.

41 Jin, Y.S., Stephanopoulos, G., 2007. Multi-dimensional gene target search for
42 improving lycopene biosynthesis in *Escherichia coli*. Metab. Eng. 9, 337-347.

43 Kainou, T., Okada, K., Suzuki, K., Nakagawa, T., Matsuda, H., Kawamukai, M.,
44 2001. Dimer formation of octaprenyl-diphosphate synthase (IspB) is essential for

1 chain length determination of ubiquinone. J. Biol. Chem. 276, 7876-7883.

2 Kang, M.J., Lee, Y.M., Yoon, S.H., Kim, J.H., Ock, S.W., Jung, K.H., Shin, Y.C.,
3 Keasling, J.D., Kim, S.W., 2005. Identification of genes affecting lycopene
4 accumulation in *Escherichia coli* using a shot-gun method. Biotechnol. Bioeng.
5 91, 636-642.

6 Kim, J., Adhya, S., Garges, S., 1992. Allosteric changes in the cAMP receptor protein
7 of *Escherichia coli*: Hinge reorientation. Proc. Natl. Acad. Sci. U. S. A. 89,
8 9700-9704.

9 Klein-Marcuschamer, D., Santos, C.N.S., Yu, H.M., Stephanopoulos, G., 2009.
10 Mutagenesis of the Bacterial RNA Polymerase Alpha Subunit for Improvement
11 of Complex Phenotypes. Appl. Environ. Microbiol. 75, 2705-2711.

12 Kolb, A., Busby, S., Buc, H., Garges, S., Adhya, S., 1993. Transcriptional regulation
13 by cAMP and its receptor protein. Annu Rev Biochem 62, 749-795.

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

17 Martinez-Antonio, A., Collado-Vides, J., 2003. Identifying global regulators in
18 transcriptional regulatory networks in bacteria. Curr. Opin. Microbiol. 6,
19 482-489.

20 Matthews, P.D., Wurtzel, E.T., 2000. Metabolic engineering of carotenoid
21 accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool
22 with expression of deoxyxylulose phosphate synthase. Appl. Microbiol.
23 Biotechnol. 53, 396-400.

24 Merrikh, H., Ferrazzoli, A.E., Lovett, S.T., 2009. Growth Phase and (p)ppGpp
25 Control of IraD, a Regulator of RpoS Stability, in *Escherichia coli*. J. Bacteriol.
26 191, 7436-7446.

27 Montemiglio, L.C., Macone, A., Ardiccioni, C., Avella, G., Vallone, B., Savino, C.,
28 2013. Redirecting P450 EryK Specificity by Rational Site-Directed Mutagenesis.
29 Biochemistry 52, 3678-3687.

30 Okada, K., Minehira, M., Zhu, X.F., Suzuki, K., Nakagawa, T., Matsuda, H.,
31 Kawamukai, M., 1997. The ispB gene encoding octaprenyl diphosphate synthase
32 is essential for growth of *Escherichia coli*. J. Bacteriol. 179, 3058-3060.

33 Park, K.S., Lee, D.K., Lee, H., Lee, Y., Jang, Y.S., Kim, Y.H., Yang, H.Y., Lee, S.I.,
34 Seol, W., Kim, J.S., 2003. Phenotypic alteration of eukaryotic cells using
35 randomized libraries of artificial transcription factors. Nat. Biotechnol. 21,
36 1208-1214.

37 Patnaik, R., 2008. Engineering complex phenotypes in industrial strains. Biotechnol.
38 Prog. 24, 38-47.

39 Rhodius, V.A., Busby, S.J.W., 2000. Transcription activation by the *Escherichia coli*
40 cyclic AMP receptor protein: Determinants within activating region 3. Journal of
41 Molecular Biology 299, 295-310.

42 Sandmann, G., Albrecht, M., Schnurr, G., Knorzer, O., Boger, P., 1999. The
43 biotechnological potential and design of novel carotenoids by gene combination
44 in *Escherichia coli*. Trends Biotechnol. 17, 233-237.

- 1 Sandmann, G., Woods, W.S., Tuveson, R.W., 1990. Identification of carotenoids in
2 *erwinia-herbicola* and in a transformed *Escherichia coli* strain. FEMS Microbiol.
3 Lett. 71, 77-82.
- 4 Sharan, S.K., Thomason, L.C., Kuznetsov, S.G., Court, D.L., 2009. Recombineering:
5 a homologous recombination-based method of genetic engineering. Nature
6 Protocols 4, 206-223.
- 7 Shimada, T., Fujita, N., Yamamoto, K., Ishihama, A., 2011. Novel Roles of cAMP
8 Receptor Protein (CRP) in Regulation of Transport and Metabolism of Carbon
9 Sources. PLoS One 6, e20081.
- 10 Sikkema, J., Debont, J.A.M., Poolman, B., 1995. Mechanisms of membrane toxicity
11 of hydrocarbons. Microbiol. Rev. 59, 201-222.
- 12 Stemmer, W.P.C., 1994. DNA shuffling by random fragmentation and
13 reassembly-in-vitro recombination for molecular evolution. Proc. Natl. Acad.
14 Sci. U. S. A. 91, 10747-10751.
- 15 Sung, W.S., Lee, I.S., Lee, D.G., 2007. Damage to the cytoplasmic membrane and
16 cell death caused by lycopene in *Candida albicans*. J. Microbiol. Biotechnol. 17,
17 1797-1804.
- 18 Williams, R., Bell, A., Sims, G., Busby, S., 1991. The role of 2 surface exposed loops
19 in transcription activation by the *Escherichia coli* CRP and FNR proteins.
20 Nucleic Acids Research 19, 6705-6712.
- 21 Won, H.S., Seo, M.D., Ko, H.S., Choi, W.S., Lee, B.J., 2008. Interdomain interaction
22 of cyclic AMP receptor protein in the absence of cyclic AMP. J. Biochem. 143,
23 163-167.
- 24 Zhang, H.F., Chong, H.Q., Ching, C.B., Jiang, R.R., 2012a. Random mutagenesis of
25 global transcription factor cAMP receptor protein for improved osmotolerance.
26 Biotechnol. Bioeng. 109, 1165-1172.
- 27 Zhang, H.F., Chong, H.Q., Ching, C.B., Song, H., Jiang, R.R., 2012b. Engineering
28 global transcription factor cyclic AMP receptor protein of *Escherichia coli* for
29 improved 1-butanol tolerance. Appl. Microbiol. Biotechnol. 94, 1107-1117.
- 30 Zhang, J.H., Dawes, G., Stemmer, W.P.C., 1997. Directed evolution of a fucosidase
31 from a galactosidase by DNA shuffling and screening. Proc. Natl. Acad. Sci. U.
32 S. A. 94, 4504-4509.
- 33 Zhou, D., Yang, R., 2006. Global analysis of gene transcription regulation in
34 prokaryotes. Cellular and Molecular Life Sciences 63, 2260-2290.
- 35 Zhou, K., Zou, R.Y., Stephanopoulos, G., Too, H.P., 2012. Metabolite Profiling
36 Identified Methylerythritol Cyclodiphosphate Efflux as a Limiting Step in
37 Microbial Isoprenoid Production. PLoS One 7, e47513.

1 List of figure captions:

2 **Fig. 1** Lycopene content of BW25113/pKSC-crp (control), M₁, M₂, M₃, M₄, M₅ and
3 M₆ cultivated in LB medium at 30 °C for 24 h. Lycopene was extracted by acetone
4 from the harvested cells. The dry cell weight was obtained by drying the washed cell
5 pellet at 105 °C for 24 h. The data represents the means of the three separate
6 experiments, and the error bars represent the standard deviation.

7 **Fig. 2** Comparison of lycopene content between MT-1 and WT cultivated in LB
8 medium at 30 °C for 24 h. Lycopene was extracted by acetone from the harvested
9 cells. The dry cell weight was obtained by drying the washed cell pellet at 105 °C for
10 24 h. Bars are used for cell mass (*white bars*), lycopene (*sparse bars*), and lycopene
11 content (*black bars*). The data represents the means of three separate experiments, and
12 the error bars represent the standard deviation.

13 **Fig. 3** Comparison of lycopene production and cell density between MT-1 and WT in
14 fed-batch time course fermentation. 10% seed culture (*v/v*) was inoculated into a 10-L
15 stirred tank fermentor containing 6 L MBL medium and fermented at 30°C with 30
16 L/min aeration rate and 200 rpm stirring speed. Culture samples were collected every
17 two hours to analyze cell growth and lycopene production. MT-1 (red lines), WT
18 (blue lines, the density of culture (empty circles), lycopene production (filled circles).

19 **Fig. 4** β -carotene content of recombinant *E.coli* cultivated in LB medium at 30°C for
20 24 h. pKSC-mcrp26 was introduced into a β -carotene-producing strain
21 BW25113-CARO, and the same strain harboring a pKSC-crp plasmid was set as a
22 control. Bars are used for cell mass (*white bars*), β -carotene (*sparse bars*), and
23 β -carotene content (*black bars*). The data represents the means of three separate
24 experiments, and the error bars represent the standard deviation.

25

1 **Table1. Primers used in this study.**

Primer	Sequence
crp_sense	5'-GAGAGGATCCATAACAGAGGATAACCGCGCATG-3'
crp_anti	5'-AGATGGTACCAAACAAAATGGCGCGCTACCAGGTAACGCGCCA-3'
crp_159_sense	5'-GTGGCAGTGCTGATCAACGACGAAGAGGGTAAAGA-3'
crp_159_anti	5'-TCTTTACCCTCTTCGTCGTTGATCAGCACTGCCAC-3'
crp_424_sense	5'-GCGTTCCTCGACGTGACGAGCCGCATTGCACAGACTC-3'
crp_424-anti	5'-GAGTCTGTGCAATGCGGCTCGTCACGTCGAGGAACGC-3'
crp_red_sense	5'-CAGCGGCGTTATCTGGCTCTGGAGAAAGCTTATAACAGAGGATAA CCGCGCATGGTGCTTGGCAAACCGCAAACAGACCCG-3'
crp_red_anti	5'-GGGGAAACAAAATGGCGCGCTACCAGGTAACGCGCCACTCCGAC GGGATTATGAGCGATTGTGTAGGCTGGAGCTGCTTCG-3'

2

3

4

1 **Table 2. Amino acid substitutions in M₁, M₂, M₃, M₄, M₅ and M₆.**

Strain	Mutated gene	Mutation	Amino acid substitutions
M ₁	mcrp26	A-T	D8V
M ₂	mcrp159	A-C	K52N
M ₃	mcrp424	G-A	G141S
M ₄	mcrp26/159	A-T, A-C	D8V, K52N
M ₅	mcrp26/424	A-T, G-A	D8V, G141S
M ₆	mcrp26/159/424	A-T, A-C, G-A	D8V, K52N, G141S

2

3

1 **Table 3. Domain distribution of *crp* mutants.**

Effect	Domain			Reference
	N-terminal	Hinge	C-terminal	
Biofuel tolerance	H31D, V47E, D53N, G56S, M59T, G71D, Q80L, V86L, T127N	F136I, D138V, D138Y	V176A, G177A, T208N, S179R, H199R	Chong et al. 2013a, Zhang et al. 2012b, Chong et al. 2013b
Acetate or low pH tolerance	F69C	D138Y, V139M		Chong et al. 2013b Basak et al. 2014
Osmotic pressure tolerance	K21I, V43L, F69Y, G71C, G74D, R87H, M114V, Q119N, K130E	D138Y		Zhang et al. 2012a
Oxidative stress	F69C, R82C, T127N	D138V, V139M,	T146L	Basak et al. 2012a
Organic solvent tolerance	T127N	F136I		Basak et al. 2012b
Metabolite improvement	D8V, K52N		G141S	This work

2

3

1 **Table 4. The expression changes of some genes, which might be able to increase**
2 **the lycopene synthesis ability.**

Gene name	Gene description	log ₂ fold-change	p-value
<i>fbaA</i>	fructose-bisphosphate aldolase	0.1	0.0229
<i>pfkA</i>	regulate the glycolysis pathway	0.3	0.0473
<i>ispG</i>	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase	0.3	0.0131
<i>appY</i>	related to anaerobiosis	0.6	0.0473
<i>araE</i>	responsible for providing NADPH	1.1	0.0017
<i>rpoS</i>	related to stress responses	0.5	0.0045
<i>yjiD</i>	inhibit the proteolysis of σ^S	0.5	0.0018
<i>araC</i>	co-regulators of CRP	0.6	0.0011
<i>yeiL</i>	co-regulators of CRP	3.2	0.0020
<i>csrB</i>	regulate G6P isomerase	0.9	0.0002
<i>hns</i>	DNA-binding transcriptional dual regulator	-0.4	0.0384
<i>creB</i>	catabolic regulation response regulator	-0.5	0.0422

3

4

Figure 1

Figure 2

Figure 3

Figure 4