Metabolic Engineering 12 (2010) 112–122

Contents lists available at ScienceDirect

Metabolic Engineering

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

Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid

Pornkamol Unrean, Cong T. Trinh, Friedrich Srienc*

Department of Chemical Engineering and Materials Science, and BioTechnology Institute, University of Minnesota, 240 Gortner Laboratory, 1479 Gortner Ave, St. Paul, MN 55108, USA

ARTICLE INFO

Article history:
Received 5 December 2008
Received in revised form
21 August 2009
Accepted 6 November 2009
Available online 26 November 2009

Keywords:
Elementary mode analysis
Metabolic network analysis
Inverse metabolic engineering
carotenoid production
Escherichia coli metabolism

ABSTRACT

We applied elementary mode analysis to a recombinant metabolic network of carotenoid-producing *E. coli* in order to identify multiple-gene knockouts for an enhanced synthesis of the carotenoids diapolycopendial (DPL) and diapolycopendioic acid (DPA). Based on the model, all inefficient carotenoid biosynthesis pathways were eliminated in a strain containing a combination of eight gene deletions. To validate the model prediction, the designed strain was constructed and tested for its performance. The designed mutant produces the carotenoids at significantly increased yields and rates as compared to the wild-type. The consistency between model prediction and experimental results demonstrates that elementary mode analysis is useful as a guiding tool also for the rational strain design of more complex pathways for secondary metabolite production.

1. Introduction

Construction of desirable, better producing strains can be achieved through introduction of pathway modifications that direct the metabolic flux towards the desired products. However, the identifications of the targets of gene manipulations resulting in a specific cellular metabolism to function in a favorable direction is challenging due to the complexity of the interconnected cell reaction network. To address this question rigorously, a quantitative understanding of the cellular metabolic network is needed. In this paper, we describe a further application of elementary mode analysis (EMA) for characterizing the metabolism of a cell and for guiding the rational design of metabolic engineering manipulations.

Elementary mode analysis uses nullspace and convex analysis under steady-state balance operation of the network to decompose the complex metabolic network of a cell into a set of unique and indivisible pathways, called elementary modes (Pfeiffer et al., 1999; Schuster et al., 2000). As a result, EMA generates all possible pathways that can exist in a cell. This pathway information along with the direct correspondence between metabolic reaction network and genetic network allows the prediction of the cellular phenotype resulting from genotype perturbations. EMA has been previously used for understanding the yield range of metabolic pathways in a network or for predicting cell phenotype in a genetically modified organism. For example, EMA has been applied for examining the effects of gene manipulations and of changes in environmental conditions on *Escherichia coli* cell growth (Stelling et al., 2002). EMA has also proven to be a powerful tool in metabolic engineering. The method can be used for designing the intermediary metabolism of a cell to optimize cell growth and biomass production (Carlson and Srienc, 2004a, 2004b; Trinh et al., 2006) or to optimize the formation of primary metabolites such as ethanol (Trinh et al., 2008). However, previous examples of EMA applications do not include the optimization of recombinant cells containing a complex pathway for a secondary metabolite. To demonstrate the use of EMA in such system with a secondary metabolite-producing pathway, we have applied this method to the metabolic network of transgenic carotenoid expressing *E. coli*.

Carotenoids are natural color pigments consisting of long, unsaturated hydrocarbon chains of 30–50 carbons that are synthesized in certain plants and microorganisms. Carotenoids can have a variety of functions. They are the light-harvesting pigments in the photosynthetic process, UV-protective compounds, regulators of membrane fluidity and antioxidant materials to protect cells against harmful oxygen radicals. Recently, it has been found that carotenoids play an important role in the prevention of cancer and of certain chronic degenerative diseases including heart disease, immune deficiency and aging (Vershinin, 1999). Carotenoids are currently used as nutrient supplements, pharmaceuticals and food colorants. To date, more than 600 carotenoids have been characterized. Most of them exist as biosynthetic intermediates. Therefore, they occur only in trace

* Corresponding author. Fax: +1 612 625 1700.
E-mail address: srienc@umn.edu (F. Srienc).

1096-7176/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2009.11.002

amounts in natural sources and are difficult to extract in sufficient amounts to be useful for applications. Therefore, carotenoid production has been attempted using microbial fermentation. Carotenogenic recombinant *E. coli* have been engineered for expressing various carotenoids including zeaxanthin, lycopene and β-carotene (Albrecht et al., 1997; Ruther et al., 1997). However, carotenoid production levels in recombinant *E. coli* are still not high enough for an economical large scale production.

Previous attempts of improving carotenoid production in recombinant *E. coli* mostly relied on a non-systematic approach of genetic manipulation. This included over-expression of genes of the carotenoid biosynthesis pathway or deletion of competing central metabolic genes to redirect precursors towards enhanced production of carotenoids (Albrecht et al., 1999; Wang et al., 1998; Farmer and Liao, 2001). However, this strategy is inefficient due to the lack of direction and an extensive amount of time and cost required. In addition, this approach may not capture all genetic modifications needed for an efficient carotenoid formation. In this study, we applied EMA to rationally design an *E. coli* mutant strain for an enhanced production of the carotenoids diapolycopendial (DPL) and diapolycopendioic acid (DPA). EMA was systematically used to examine gene deletion effects on carotenoid production of *E. coli* as well as to identify a set of multiple-gene knockouts which results in a higher product yield. The rationally designed mutant was experimentally constructed and the kinetics of the production of DPL and DPA was determined to verify the model predictions.

2. Materials and methods

2.1. Bacterial strains and plasmids

E. coli MG1655 and its mutant derivatives were used as the hosts for synthesis of carotenoids DPL and DPA. Individual strains containing, respectively, the genetic knockouts Δ*ldhA*::*kan*+, Δ*frdA*::*kan*+, Δ*poxB*::*kan*+, Δ*pta*::*kan*+, Δ*adhE*::*kan*+, Δ*pykF*::*kan*+, Δ*zwf*::*kan*+ and Δ*maeB*::*kan*+ were obtained from the *E. coli* knockout library, the Keio collection (Baba et al., 2006). *E. coli* cells containing multiple-gene knockouts generated in this study were constructed using a generalized P1 transduction technique after transferring the desired deletion gene from the knockout library into the MG1655 recipient strain (Ausubel et al., 1995). Table 1 summarized all strains used in this study. Primers designed to regions outside and inside of the targeted knockout gene shown in Table 2 were used in conjunction with PCR and gel electrophoresis to verify the knockout construction. The inside primers verify that the structural gene is absent in the chromosome while the outside primers validate that the targeted gene is indeed deleted. The strains were later transformed with a carotenoid plasmid, pACMNOx (Mijts et al., 2005) which has been kindly provided by Prof. C. Schmidt-Dannert (Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis MN, USA). Included on the plasmid is the chloramphenicol resistance gene for selection.

2.2. Growth medium

The bacterial cells were cultivated in M9 minimal medium (2.0 g/l NH_4Cl, 1 g/l NaCl, 25.6 g/l $Na_2HPO_4 \cdot 7H_2O$, 6.0 g/l KH_2PO_4, 0.01 g/l $CaCl_2$, 0.24 g/l $MgSO_4$ (Sigma, St. Louis, MO, USA)) supplemented with trace metals and minerals (5.5 mg/l $CaCl_2 \cdot 2H_2O$, 1.0 mg/l $MnCl_2 \cdot 4H_2O$, 1.7 mg/l $ZnCl_2$, 0.43 mg/l $CuCl_2$, 0.60 mg/l $CoCl_2 \cdot 6H_2O$, 0.60 mg/l $NaMoO_4 \cdot 2H_2O$, 8.42 mg/l $FeSO_4 \cdot 7H_2O$, 10 mg/l Fe(NH_4)-citrate, 0.38 mg/l $Na_2B_4O_7 \cdot 10H_2O$, 1 mg/l of thiamine (Sigma, St. Louis, MO, USA)). The medium contained 5 g/l of glucose as a carbon source. The addition of 50 μg/ml chloramphenicol was used to maintain plasmid in the carotenoid-producing cell culture.

2.3. Growth in batch shake-flasks and bioreactors

In shake-flask growth experiment, 1 ml of seed culture grown overnight from transformed colonies was inoculated in a 250 ml baffled shake flask containing 50 ml of M9 medium supplemented with glucose. The culture was grown in a Labline shaking incubator at 200 rpm agitation speed. Temperature was maintained at 30 °C. Cell samples (5 ml) were periodically collected for carotenoid measurement. Batch bioreactor experiments were conducted in a 10 l Braun bioreactor (Biostat MD, B. Braun Biotech International, Melsungen, Germany) containing 5 l of the same culture medium as the shake flask experiments. The culture was inoculated with 1% (v/v) seed culture grown in aerobic baffled shake-flasks. The reactor was operated at 1 vvm aeration rate. Agitation speed was set at 300 rpm and temperature was controlled at 30 °C. pH was controlled at 6.9 with 28% (v/v) NH_4OH and 40% (v/v) H_3PO_4 (Sigma, St. Louis, MO, USA). An online off-gas mass spectrometer (prima δ-B MS; Thermo Onix, Houston, TX, USA) was used for determining oxygen consumption and carbon dioxide formation in the culture. Cell samples and supernatants were collected every 2–4 h for glucose, carotenoid and cell growth measurements.

2.4. Analytical techniques

Cell growth was monitored via optical density at a wavelength of 600 nm in a 1 cm cuvette using a Hewlett Packard 8452A Diode Array spectrophotometer (Palo Alto, CA, USA). Cell mass was determined from a standard curve correlating optical density to cell dry weight. To measure cell dry weight, samples were collected from the culture, washed with cold deionized water

Table 1
Strains used in this study.

Strain	Derived from parental strain	Modifications	References
Wild-type	MG1655	None	Bachmann (1996)
JW1375	BW25113	Δ*ldhA*::*kan*+	Baba et al. (2006)
JW4115	BW25113	Δ*frdA*::*kan*+	Baba et al. (2006)
JW0855	BW25113	Δ*poxB*::*kan*+	Baba et al. (2006)
JW2294	BW2229	Δ*pta*::*kan*+	Baba et al. (2006)
JW1228	BW25113	Δ*adhE*::*kan*+	Baba et al. (2006)
JW1666	BW25113	Δ*pykF*::*kan*+	Baba et al. (2006)
JW1841	BW25113	Δ*zwf*::*kan*+	Baba et al. (2006)
JW2447	BW25113	Δ*maeB*::*kan*+	Baba et al. (2006)
CRT028	MG1655	Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*Δ*maeB*::*kan*-	This study

Table 2
Primer sequences designed for testing specific deleted genes.

Deleted genes	Left primer	Right primer	Position[a]
ldhA	5′-CGGCTTTATATTTACCCAGC-3′	5′-CGCAACAAACGCGGCTAC-3′	Outside
frdA	5′-ACGCTTCAACCTTCATACCG-3′	5′-GATGCCGTTTCGCTCATAGT-3′	Outside
poxB	5′-ATGGATATCGTCGGGTTTGA-3′	5′-AAGCAATAACGTTCCGGTTG-3′	Outside
pta	5′-GTTCGCCTGCTTCGTTAGTC-3′	5′-CTGCCGCTATGTTGAAGACA-3′	Outside
adhE	5′-AAAGCGATGCTGAAAGGTGT-3′	5′-AGAAAGCGTCAGGCAGTGTT-3′	Outside
pykF	5′-GCGCCAATTGACTCTTGAAT-3′	5′-CCTGCCAGCAGAGTAGAACC-3′	Outside
zwf	5′-CGCGTAACAATTGTGG-3′	5′-CTGGATTTTTTCCAGC-3′	Outside
maeB	5′-CTGTTTGATGCCGTCTAACTCGTTC-3′	5′-CTTTATCCATGAGTCGCCGCCTGTG-3′	Outside
ldhA	5′-TACCCAACGAACCAATTTTC-3′	5′-GCTGGAAGAGCTGAAAAAGC-3′	Inside
frdA	5′-TTACGTGCCATTGCGGAGTG-3′	5′-TCACGATACAGTAGCGGGTG-3′	Inside
poxB	5′-CCACCAGCTTTCATCTCCAT-3′	5′-TATTCCCTCCAGCGAAATTG-3′	Inside
pta	5′-TCAGATCCGGGAAGATGAAC-3′	5′-TGTGCTGATGGAAGAGATCG-3′	Inside
adhE	5′-TGAATGCAGTCTGCTTGGTC-3′	5′-AAAACGTTGGTCCGAAACAC-3′	Inside
pykF	5′-CACCGTACTGGTTGACGATG-3′	5′-CACAACGCCTTTGCTCAGTA-3′	Inside
zwf	5′-TCTACCCATTTCCAGGCTTC-3′	5′-TGGGACACCCTGAGTGCACG-3′	Inside
maeB	5′-GAGCTGTCCGGCATACGGTC-3′	5′-CGCTGGCAGGCAAACCGGTG-3′	Inside

[a] Inside means the primers specified to deleted part of the gene, while outside means the primers specified to undeleted portions of the gene.

after centrifugation and transferred to a pre-weighed tube. The tube containing cell pellet was dried at 100 °C for 24 h, then weighed and the mass of the empty tube was subtracted to determine cell dry weight concentration in the culture. Metabolite concentrations including glucose and other secreted byproducts were determined using a HPLC system (Shimadz10A, Shimadzu, Columbia, MD, USA) equipped with an autosampler (SIL-10AF), a cation exchange column (HPX-87H, Biorad Labs, Hercules, CA, USA) and two detectors in series including a UV–vis detector (SPD-10A) and a refractive index detector (RID-10A). The column was run in an isocratic mode at 50 °C at a flow rate of 0.6 ml/min with 0.01 M H_2SO_4 as a mobile phase. A standard curve correlating area to concentration of metabolites was used to determine the quantity of metabolites in the sample.

2.5. Carotenoid analysis

Intracellular carotenoids were extracted from 5 ml cell culture. The centrifuged cell pellet was first washed with 1 ml of cold 0.9% NaCl solution. The washed cell pellet was then extracted with 1 ml acetone with vigorous vortexing for 5 min. After centrifugation, the remaining cell pellet was resuspended in 10% (g/g) KOH and incubated at 30 °C for 24 h. The KOH extract solution was later purified by liquid–liquid extraction with chloroform. Carotenoid content in acetone and KOH extract solution was quantified by measuring the absorbance at 506 and 490 nm, respectively. Concentration of carotenoid was determined using extinction coefficients obtained from Carote-Nature (Lupsingen, Switzerland) and appropriate dilution factor. The extinction coefficient of diapolycopendial at 506 nm and diapolycopendioic acid at 490 nm are 4084 and 2784 $(g/100\ ml)^{-1}\ cm^{-1}$, respectively.

2.6. Elementary mode analysis

The metabolic network was based on the previously described *E. coli* network (Carlson and Srienc 2004a). Carotenoid synthesis through the non-mevalonate pathway was included in the model to account for the production of carotenoid in the recombinant cells. The pathway details were collected from the Ecocyc database that is available at http://www.ecocyc.org. The model was based on utilizing glucose as the carbon source. Cell growth in the metabolic network is described through the production of biomass which is formed from precursors and energy draining of the central metabolic pathway of *E. coli* (Ingraham et al., 1983). Fig. 1 shows the metabolic map of the carotenoid-producing *E. coli* used in this study. The model has been analyzed using METATOOL software version 3.9.2 (Pfeiffer et al., 1999). Elementary mode results were analyzed using Excel Microsoft Corp. for mode sorting and filtering.

3. Results

3.1. Identification of target gene knockouts

We first constructed a carotenoid expressing *E. coli* metabolic model by adding the biosynthesis pathway of carotenoids DPL and DPA into the central metabolic network of *E. coli* MG1655 which was used as the host for synthesizing carotenoids. Fig. 1 presents the metabolic network of carotenoid expressing *E. coli* considered in this study. It included a total of 58 metabolic reactions (22 reversible, 36 irreversible) and 57 metabolites. Elementary mode analysis on the network revealed a total of 29,532 elementary modes. Each mode represents a unique, possible pathway with balanced metabolites and cofactors. A total of 24,155 modes is aerobic, while 5,377 modes exist under anaerobic conditions. There are 5,923 total modes in which production of biomass and carotenoids diapolycopendial and diapolycopendioic acid are coupled. The large number of elementary modes illustrates the flexibility and robustness of the cell to adapt itself to particular conditions by using pathways that provides the optimal fitness. Comparison of all elementary modes revealed a maximum carotenoid yield of 0.83 carbon mole of carotenoid per carbon mole of glucose. The maximum possible biomass yield was 0.84 carbon mole of biomass per carbon mole of glucose, consistent with previous results (Carlson and Srienc, 2004a, 2004b).

The knowledge of all pathway options allowed the evaluation of gene knockout effects on cell phenotype. Gene knockouts were simulated by removing the enzymatic reaction corresponding to that gene from the stoichiometric matrix. The phenotype of that specific knockout mutant was then represented by a combination of remaining elementary modes when that reaction was deleted. Specifically, we focused on gene knockout effects on cell viability (biomass yield), maximal yield of carotenoid and fraction of remaining modes after each individual single gene knockout in which a specific gene and its corresponding reaction was eliminated from the metabolic network. The effects of the single

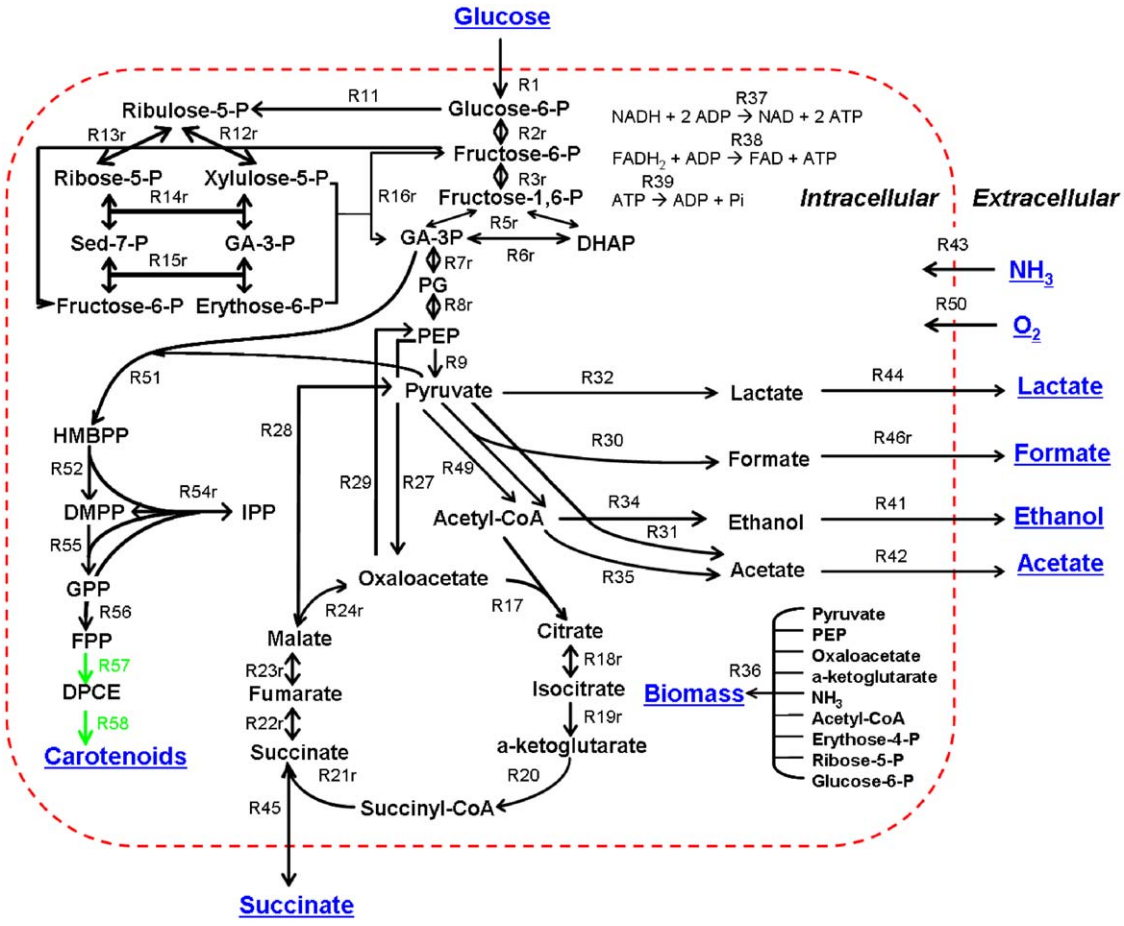

Fig. 1. Metabolic network of recombinant *E. coli* expressing the biosynthesis pathway of diapolycopendial/diapolycopendioic acid. Reaction designations in the network correspond to the reaction number in the elementary mode analysis. Reactions shown in green are introduced into the cell via genes on plasmid pACMNOx. External metabolites are underlined.

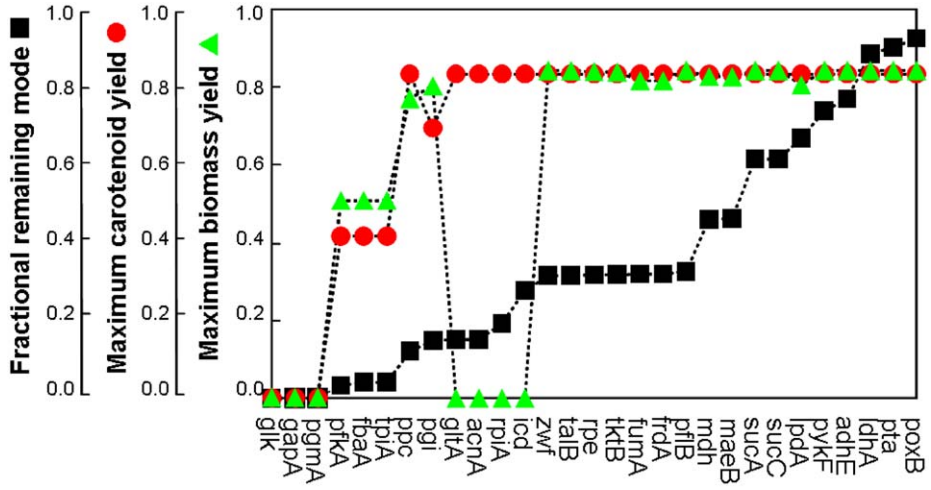

Fig. 2. Effect of single deleted genes on the synthesis of diapolycopendial/diapolycopendioic acid, biomass formation and the fraction of remaining modes. Maximum yield of carotenoids (●) and biomass (▲) as well as fraction of remaining modes (■) of single gene deletions are shown. Only the first of a linear sequence of reactions is shown. The yield is defined as ratio of carbon mole product per carbon mole glucose. The fraction of remaining modes is defined as the ratio of remaining elementary modes in the single gene knockouts to those in the wild-type. Potential knockout targets are genes whose deletion still support maximum yield of carotenoid and biomass formation while minimizing the fraction of remaining modes.

gene deletions are summarized in Fig. 2. All gene knockouts were sorted in increasing orders of the fraction of remaining elementary modes. Gene knockouts which result in zero yield of biomass are considered lethal and therefore, these genes cannot be deleted from the cell. The result in Fig. 2 assists the process of screening and selecting potential gene knockout targets. Among the set of all individual knockouts, the target genes are identified when elimination of that gene still maintains the maximum possible yield of carotenoid product and retains a reasonable yield of biomass while the largest possible number of elementary

modes is eliminated (minimum number of fractional remaining modes). For example, *zwf* is a target for deletion since deletion of this gene eliminates more than 60% of identified elementary modes and still supports high yield of carotenoid and biomass. This approach allows the determination of optimal gene knockout targets which eliminate the largest number of elementary modes without affecting the most efficient pathway for carotenoid and biomass synthesis. We applied this identification tool in a sequential manner for selecting an optimal combination of multiple-gene knockouts. That is, the identified gene knockout from previous steps was used as the genetic background in the next steps for determining additional gene knockouts.

Table 3
Total elementary modes remaining after sequential deletion of multiple genes. The table illustrates the reduction of elementary modes after each sequential gene deletion step. The progressively decreasing number of available elementary modes after multiple-gene deletions restricts the cell's pathway options, hence forcing the cell to operate according to efficient pathways. Elementary modes are categorized as aerobic modes (oxygen consuming modes) and anaerobic modes which do not use oxygen.

Strain	Total modes	Aerobic modes	Anaerobic modes	Predicted CRT yield[a]
Wild-type	29,532	24,155	5377	0.0-426
Δ*ldhA*	15,662	13,405	2257	0.0-426
Δ*ldhA*Δ*frdA*	8573	7810	763	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*	7541	6861	680	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*	6171	5600	571	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*	4099	4099	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*	2573	2573	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*	375	375	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*Δ*maeB*	5	5	0	0.4-426

[a] Yield is in mg-diapolycopendioic acid/g–glucose.

Table 4
Gene knockout targets for improving the production of diapolycopendial/diapolycopendioic acid in recombinant *E. coli*. Gene and enzyme annotation were obtained from the Ecocyc database at www.ecocyc.org.

Deleted Reaction	Corresponding gene	Enzyme	Pathway
R9	*pykF*	Pyruvate kinase	Glycolysis
R11	*zwf*	Glucose-6-phosphate-1-dehydrogenase	Pentose phosphate
R22	*frdA*	Fumarate reductase	Fermentation
R28	*maeB*	Malate dehydrogenase	Anapleurotic
R31	*poxB*	Pyruvate oxidase	Fermentation
R32	*ldhA*	Lactate dehydrogenase	Fermentation
R34	*adhE*	Alcohol dehydrogenase	Fermentation
R35	*pta*	Phosphate acetyltransferase	Fermentation

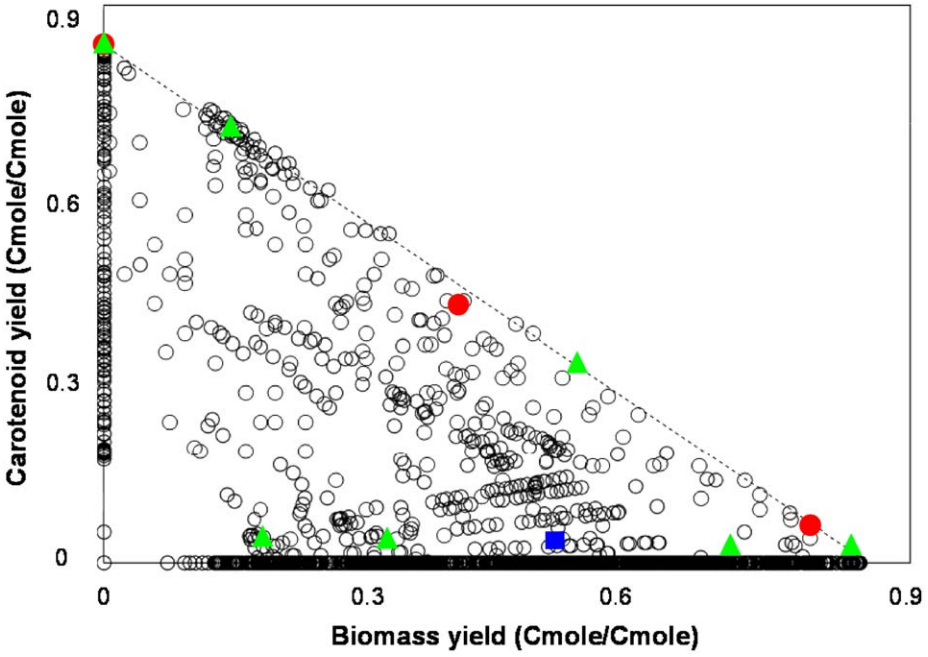

Fig. 3. Elementary modes of CRT028/pACMNOx as compared with total available elementary modes in the wild-type MG1655/pACMNOx. All elementary modes available in the wild-type (○). Elementary modes remaining after multiple-gene knockouts in the mutant assuming 4 moles NADP per mole DPA (●) and assuming 8 moles NADP per mole DPA (▲) and experimental carotenoid yield of CRT028/pACMNOx (■). The dashed line connecting the highest-yielding carotenoid mode with the highest-yielding biomass mode represent the yield of linear combinations of the two pathways. The negative slope indicates an inverse relationship between carotenoid and biomass formation.

This process was continued until no further elementary modes can be eliminated without reducing the maximum yield of carotenoid production. The combination of these gene knockout targets, therefore, forces the cell to function according to the remaining efficient carotenoid-producing elementary modes. Table 3 summarizes the number of remaining elementary modes when each gene knockout target was sequentially combined.

The identification approach resulted in a combination of eight gene knockouts which predicted a likely over-production of the carotenoids DPL and DPA in *E. coli*. The gene deletions included the removal of byproduct synthesis genes for lactate, succinate, acetate and ethanol which were experimentally accomplished by disrupting lactate dehydrogenase (*ldhA*; reaction R32), fumarate reductase (*frdA*; reaction R22), pyruvate oxidase (*poxB*; reaction R31), phosphate acetyltransferase (*pta*; reaction R35), alcohol

Table 5
Overall reaction stoichiometry and product yields of each elementary mode remaining in the multiple-gene knockout, CRT028/pACMNOx.

Ems	Overall reaction stoichiometry	Y_P[a]	Y_X[a]
1	Glucose+0.14 O_2→0.17 Carotenoids+1 CO_2	426	0
2	Glucose+0.14 O_2→0.17 Carotenoids+1 CO_2	426	0
3	Glucose+0.14 O_2→0.17 Carotenoids+1 CO_2	426	0
4	Glucose+0.11 O_2+0.53 NH_3→0.11 Carotenoid+0.91 CO_2+6.96×10^{-4} Biomass	270	0.26
5	Glucose+0.22 O_2+1.30 NH_3→1.56×10^{-4} Carotenoid+1 CO_2+1.73×10^{-3} Biomass	0.41	0.64

Carotenoid yield is presented in mg carotenoid/g glucose while biomass yield is shown in g biomass/g glucose.

[a] Y_P is carotenoid yield while biomass yield is shown in Y_X.

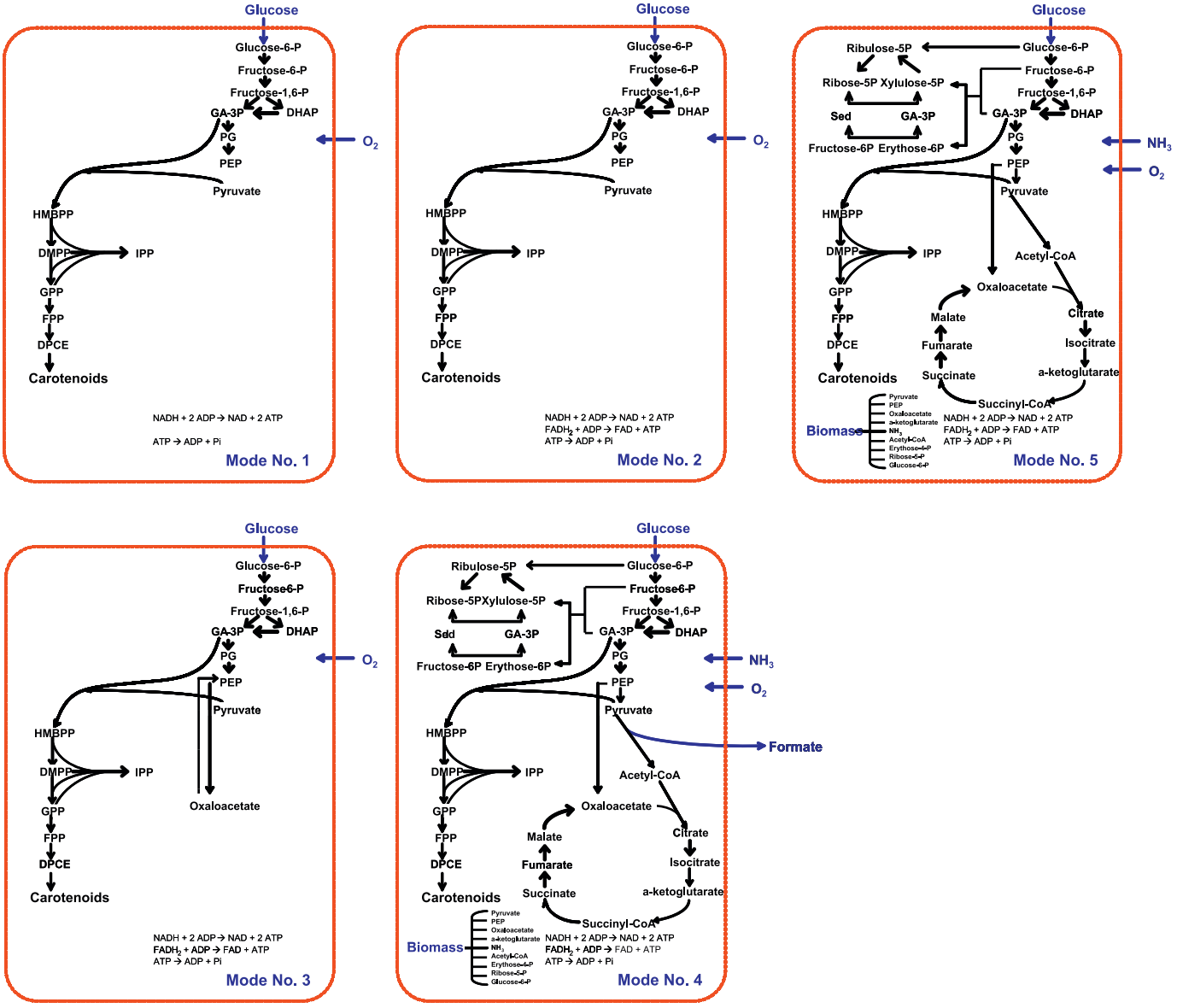

Fig. 4. The five available elementary modes in the multiple-gene knockout strain CRT028/pACMNOx.

dehydrogenase (*adhE*; reaction R34), respectively. Other target genes were pyruvate kinase (*pykF*; reaction R9), glucose-6-phosphate-1-dehydrogenase (*zwf*; reaction R11) and malate dehydrogenase (*maeB*; reaction R28). The reaction designation of these eight knockout targets is summarized in Table 4. The combination of these gene knockouts reduces the number of elementary modes to five remaining modes as shown in Fig. 3. The overall reaction stoichiometry and detailed pathways of these elementary modes are shown in Table 5 and Fig. 4, respectively. All of these remaining modes include carotenoid synthesis. Therefore, the knockout mutant is expected to produce the carotenoids DPL/DPA more efficiently than the wild-type which contains many modes that do not produce carotenoids at all. In two of the remaining modes, carotenoid synthesis is coupled with biomass formation.

3.2. Effect of NADP on production of diapolycopendioic acid

Since a detailed biosynthesis pathway of diapolycopendioic acid (DPA) has not yet been established, it is unclear how many NADP cofactors are required for the synthesis of one mole of DPA. Therefore, we evaluated the effect of the number of NADP cofactors on the number of available modes and predicted yields of DPA in the host background of CRT028/pACMNOx (Fig. 5) that was constructed with the assumption that 4 NADP are required per mole of DPA formed. This number was selected based on the assumption that for every atom of hydrogen lost in the oxygenation step, one mole of NADP cofactor is required. EMA showed that the change of the number of NADP has no effect on the predicted maximum yield of DPA. However, an increased number of NADP cofactors required for synthesis of one mole of DPA results in an increased number of available elementary modes and a decreased minimum predicted yield of DPA. Interestingly, all elementary modes available in CRT028/pACMNOx are all directed to the production of DPA regardless of the number of NADP chosen. This confirmed that for all numbers of NADP assumed, CRT028/pACMNOx is expected to produce DPA.

3.3. Quantification of diapolycopendial and diapolycopendioic acid

Three carotenogenic genes crtM, crtN and crtOx were responsible for the synthesis of the end product, diapolycopendioic acid.

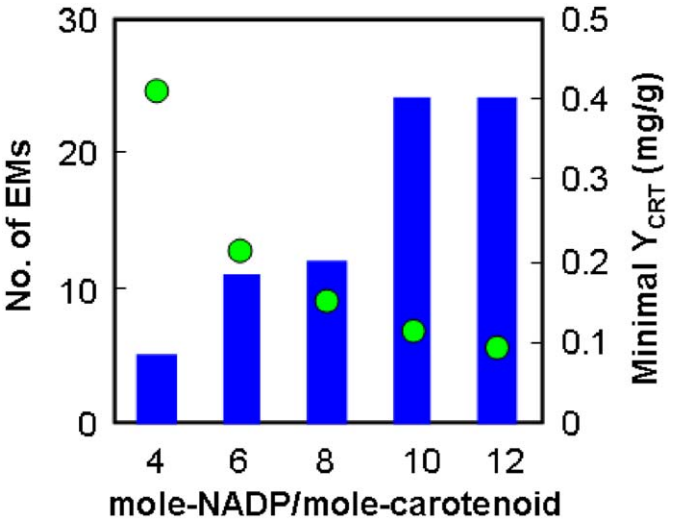

Fig. 5. Effect of the number of NADP cofactors on the number of available elementary modes (■) and minimum predicted yield of diapolycopendioic acid (●) in CRT028/pACMNOx.

A previous study has shown that *E. coli* expressing plasmid pACMNOx containing these genes stored the end product as well as its intermediates, diapolycopendial in the cell (Lee et al., 2004). These two products have distinct solubilities. Diapolycopendial is soluble in acetone while diapolycopendioic acid is soluble in an aqueous KOH solution. As a result, we developed two extraction processes to extract each of them from the cell membrane. We first extracted diapolycopendial from the cells using acetone. A 10% KOH solution was then used for the second extraction process to isolate the carotenoid diapolycopendioic acid. Absorption spectra of acetone and KOH extracts using a spectrophotometer are shown in Fig. 6. The distinct absorption spectra of the extract solutions confirmed that different carotenoids were being dissolved in each solution. In acetone supernatant, the maximum absorbance was found at a wavelength 506 nm, while maximum absorbance in KOH supernatant was at a wavelength of 490 nm. These observed wavelengths agree with the wavelength reported by in Tao et al. (2005), confirming the presence of carotenoid, diapolycopendial in acetone extract (reported λ_{max}, 507 nm) and diapolycopendioic acid in KOH extract (reported λ_{max}, 490 nm). Since both carotenoids, DPL and DPA, were found in the cells, carotenoid product measured in later experiments will be given as the sum of both carotenoids.

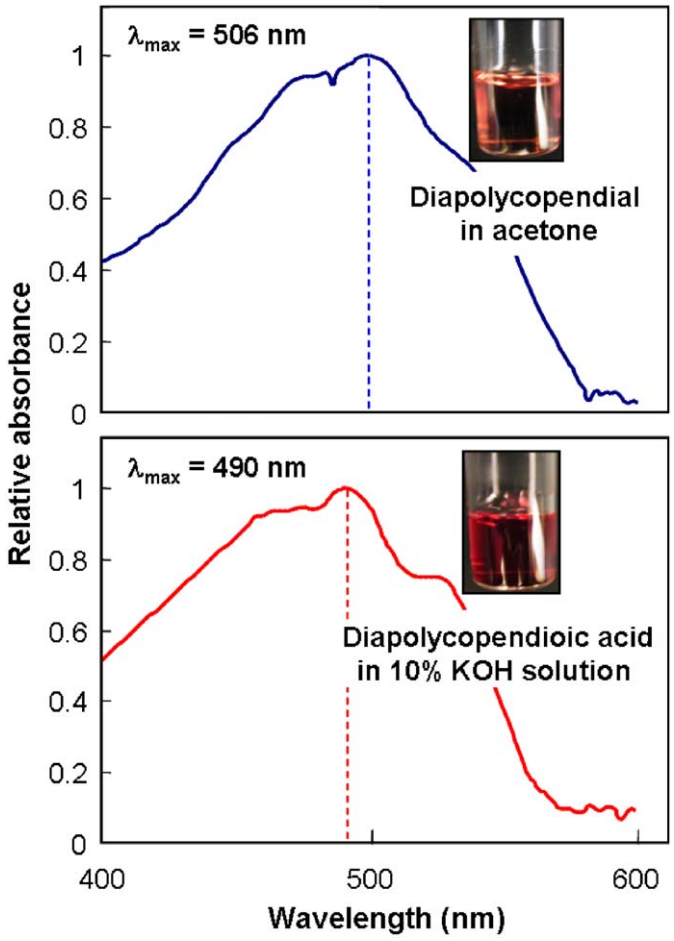

Fig. 6. Absorption spectra of extract solutions using acetone (top) and 10% KOH (bottom) from *E. coli* cells expressing plasmid pACMNOx. The maximum absorbance was at 506 nm in the acetone extract and at 490 nm in the KOH extract. The absorbance maxima in both extracts are in agreement with Tao et al. (2005) for diapolycopendial in acetone and for diapolycopendioic acid in 10% KOH solution, respectively.

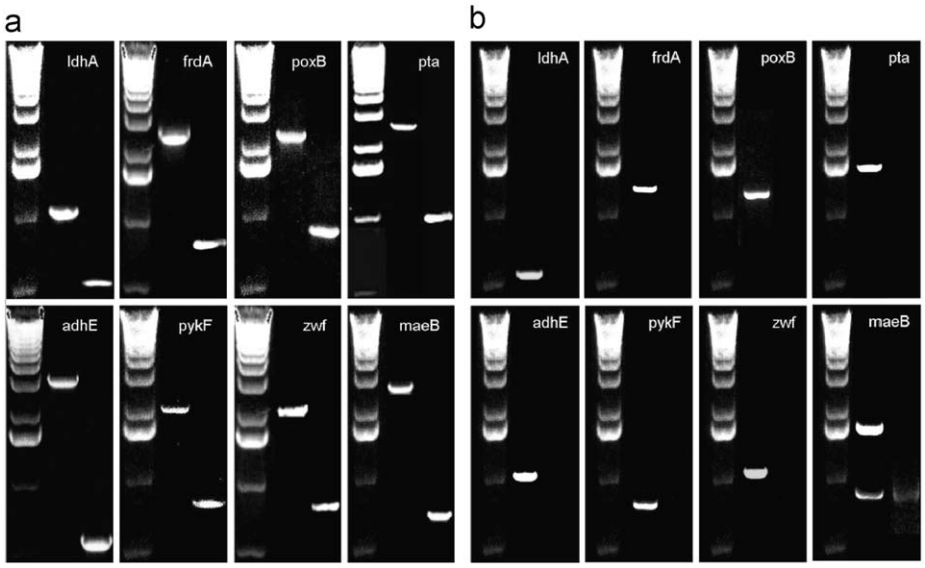

Fig. 7. Confirmation of gene deletions in mutant CRT028 by PCR and gel electrophoresis using outside (a) and inside (b) primers. For each gene, the first lane is a 1 kbp DNA ladder, the second is the PCR amplification product in the wild-type MG1655 and the third lane is the PCR amplification product in the mutant CRT028. The smaller size of the amplified gene product using outside primers and the absent amplification product using inside primers in the mutant confirm the partial deletion of the genes.

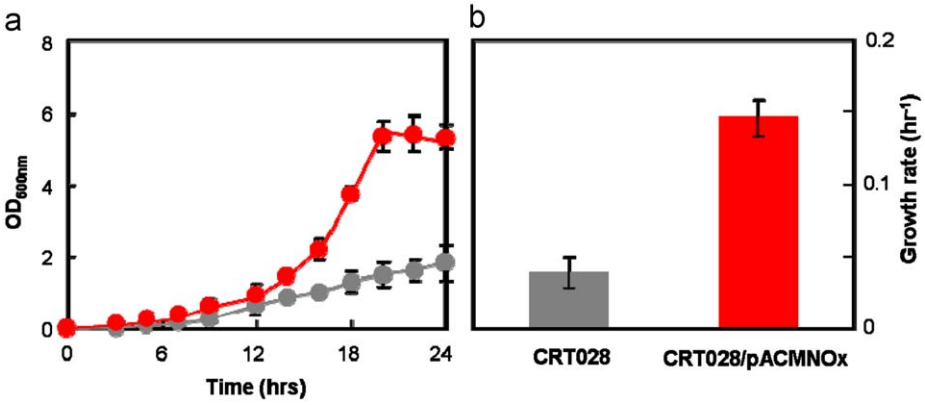

Fig. 8. Growth characteristic of mutant CRT028 and mutant CRT028 expressing plasmid pACMNOx: (a) cell growth ($OD_{600\,nm}$) vs. time and (b) specific growth rates. The experiments were conducted in baffled shake-flasks containing minimal medium supplied with glucose under aerobic conditions. Each value represents the mean of duplicate experiments.

3.4. Strain construction

To validate the knockout strain design based on EMA, the designed mutant strain CRT028, was constructed using established methods (Ausubel et al., 1995) and verified by PCR amplification and gel electrophoresis using both inside and outside primers specific to that deleted genes. Inside primers specified to the deleted part of the gene were used to ensure that a disrupted gene was not accidentally displaced somewhere else on the chromosome. Outside primers specified to undeleted portions of the gene were used to verify gene disruption at the known locations on the *E. coli* chromosome. The gel pictures in Fig. 7 confirmed of the gene disruptions in the mutant CRT028 as compared with the wild-type MG1655 containing undisrupted genes. Using the outside primers, the smaller size PCR product of the mutant as compared with that of the wild-type confirmed the gene disruption. In addition, the absence of gene product amplified using inside primers in the mutant confirmed the complete deletion of eight genes from its chromosome.

3.5. Strain characterization

According to the strain design by EMA, two of the remaining modes in the mutant are growth-associated carotenoid-producing modes (Table 5) which suggests a tight coupling between carotenoid synthesis and cell growth in the mutant. Therefore, we tested whether or not cell growth and carotenoid production are indeed coupled in the mutant strain. Growth experiments in aerobic shake-flasks showed that the mutant CRT028 grew very slowly without the plasmid pACMNOx and growth was restored when the plasmid was introduced into the mutant (Fig. 8). The growth rate of the mutant with plasmid was approximately 4-fold faster than that of the mutant without plasmid. Thus, carotenoid formation provided a strong selection pressure for plasmid even in the absence of antibiotics. Carotenoid yield (mg of carotenoid produced per g of glucose consumed) of both *E. coli* wild-type MG1655/pACMNOx and mutant strain CRT028/pACMNOx were measured. The experiments were conducted in both shake-flasks and in batch bioreactors. Fig. 9a presents the yield of DPL and DPA in wild-type and in the constructed knockout strain determined

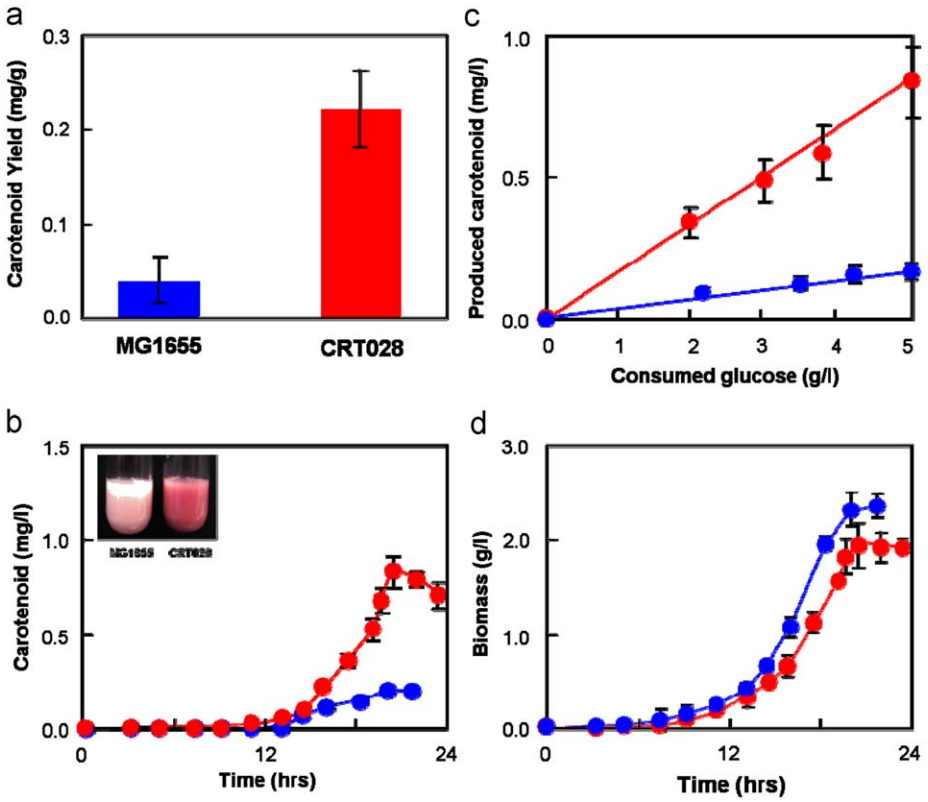

Fig. 9. Yield of diapolycopendial and diapolycopendioic acid of wild-type MG1655/pACMNOx (blue) and mutant CRT028/pACMNOx (red) in aerobic batch shake-flasks (a); carotenoid production profile of wild-type MG1655/pACMNOx (●) and mutant CRT028/pACMNOx (●) in aerobic batch bioreactors(b); a picture of the cell cultures for the two strains is shown in the upper left corner in (b); produced carotenoid vs. consumed glucose of wild-type MG1655/pACMNOx (●) and mutant CRT028/pACMNOx (●) in batch bioreactors (c); and time profiles of biomass production of wild-type MG1655/pACMNOx (●) and mutant CRT028/pACMNOx (●) in batch bioreactors (d). Overall carotenoid yield is indicated by the slope of the regression line. Yield on glucose of wild-type and mutant are 0.04 ± 0.00 and 0.17 ± 0.04 mg-carotenoids/g-glucose, respectively. The results are based on averages of three collected samples.

in aerobic shake flask experiments. Carotenoids production in the mutant CRT028/pACMNOx is significantly higher than in the wild-type MG1655/pACMNOx. In aerobic batch bioreactors, the mutant yielded 0.17 ± 0.04 mg-carotenoid/g-glucose compared to 0.04 ± 0.00 mg-carotenoid/g-glucose by the wild-type (Fig. 9c). The production profile of total DPL and DPA in batch bioreactors shown in Fig. 9b also revealed that the mutant produced carotenoid at faster production rate than the wild-type. In addition, comparison of the growth curves of MG1655/pACMNOx and CRT028/pACMNOx suggested that the mutant grew slower than the wild-type with the specific cell growth rate reduced by 23%. Growth and carotenoid phenotype of both wild-type and mutant are summarized in Table 6. The wild-type produced acetic acid closed to 1 g/l as a major byproduct. In contrast, no significant fermentation products were observed in culture of the mutant. The results from the online off-gas mass spectrometer also revealed in the case of the mutant a decrease in carbon dioxide production rate but a significantly increased oxygen uptake rate as compared to the wild-type.

Table 6
Comparison of wild-type MG1655/pACMNOx and mutant CRT028/pACMNOx synthesis of diapolycopendial and diapolycopendioic acid in aerobic batch bioreactors.

	MG1655/ pACMNOx	CRT028/ pACMNOx
Growth rate (/h)	0.17 ± 0.02	0.13 ± 0.01
Carotenoid production (mg/l)	0.19 ± 0.02	0.83 ± 0.20
Carotenoid yield (mg carotenoid/g glucose)	0.04 ± 0.00	0.17 ± 0.04
Specific production (mg carotenoid/g cell dry weight-h)	0.01 ± 0.00	0.10 ± 0.02

3.6. Strain comparison

Several *E. coli* strains have been previously developed to improve production of carotenoid through multiple-gene modifications. To compare the differences between our designed strain CRT028/pACMNOx and other mutant *E. coli* strains developed previously, we have applied elementary mode analysis to evaluate the effects of the gene knockouts on the production of DPA and on the reduction of inefficient pathways for carotenoid production (Table 7). The results show that the mutants containing gene knockouts previously identified for enhanced production of carotenoid still contain a significant number of inefficient carotenoid-producing pathways. These remaining inefficient pathways likely reduce the yield of carotneoid if the pathways are used. Unlike each of these mutants, CRT028/pACMNOx contains only carotenoid-producing elementary modes. Therefore, our designed strain is expected to be able to efficiently produce carotenoid diapolycopendioic acid.

3.7. Production of other carotenoids

Since all carotenoids are known to share a common biosynthesis pathway, we also tested the performance of the mutant for the production of other carotenoids including diapolycopene, tetradehydrolycopene, tetradehydrolycopendial and lycopene. The comparison of each carotenoid yield (mg-carotenoid/

Table 7
Elementary mode analysis of *E. coli* knockout strains using glucose as a carbon source for production of diapolycopendioic acid under aerobic conditions.

Knockout/pACMNOx	Total EMs	CRT yield of total EMs	CRT-producing EMs	CRT yield of CRT-producing EMs	References
ΔpykFΔpykA	4275	0.00–426	929	< 0.01–426	Farmer and Liao (2001)
ΔaceEΔfdhF	1049	0.00–426	242	0.10–426	Alper et al. (2005)
ΔaceEΔtalB	952	0.00–426	352	0.02–426	Alper et al. (2005)
ΔaceEΔtalBΔfdhF	303	0.00–426	114	0.10–426	Alper et al. (2005)
ΔldhAΔfrdAΔadhEΔpoxB ΔptaΔpykFΔzwfΔmaeB	5	0.41–426	5	0.41–426	This study

CRT, carotenoid diapolycopendioic acid. Yield is in mg-carotenoid/g-glucose.

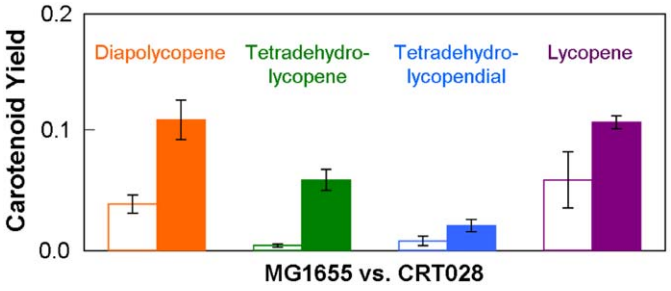

Fig. 10. Production of diapolycopene (■), tetradehydrolycopene (■), tetradehydrolycopendial (■) and lycopene (■) by wild-type MG1655 (empty bar) and mutant CRT028 (filled bar) in aerobic batch shake-flasks. Product yield is reported in mg-carotenoid/g-glucose. Carotenoid content was estimated using extinction coefficients and appropriate dilution factors. The extinction coefficient of diapolycopene at 470 nm is 4450 $(g/100\,ml)^{-1}\,cm^{-1}$ (Wieland et al., 1994). The extinction coefficient of tetradehydrolycopene at 508 nm and of tetradehydrolycopendial at 504 nm is 5933 $(g/100\,ml)^{-1}\,cm^{-1}$ and 5445 $(g/100\,ml)^{-1}\,cm^{-1}$, respectively. The values are estimated from extinction coefficient values of diapolycopene and diapolycopendial, respectively, by adjusting based on number of carbons since they have similar structure. The extinction coefficient of lycopene at 475 nm is 3410 $(g/100\,ml)^{-1}\,cm^{-1}$ (Britton et al., 1995).

g-glucose) between wild-type and the mutant shown in Fig. 10 demonstrates that the mutant outperforms the wild-type in the synthesis of other carotenoids as well. The yield improvement of all carotenoids tested in the mutant suggests that the designed mutant could be used as a platform host for the production of many other carotenoids in addition to DPL and DPA.

4. Discussion

Because carotenoid products are of considerable industrial and nutritional value, several groups have recently metabolically engineered recombinant cells for enhancing production of carotenoids. The strategies applied were based on over-expression and/or deletion of genes involved in the carotenoid biosynthesis pathway. Although significant improvement in carotenoid production was observed using this approach, the genetic manipulations in general have not been rationally suggested. Here, we have presented a systematic strategy by using elementary mode analysis (EMA) to design an optimized *E. coli* strain for the most efficient synthesis of carotenoids. In particular, we applied EMA to identify targets for gene deletions resulting in efficient production of diapolycopendial (DPL) and diapolycopendioic acid (DPA). The designed strain was obtained through a sequential implementation of eight multiple-gene deletions which narrowed the possible pathway space to the efficient carotenoid-producing options.

EMA decomposes a metabolic network into a set of unique and non-divisible pathways that represent all possible physiological states of the cells. Unlike other constraint-based models such as OptKnock or MOMA where only one optimal solution is identified, EMA identifies the complete set of possible optimal pathway solutions. As a result, the strain designed by EMA is guaranteed to function according to these optimal pathways. Based on elementary mode analysis, an inverse relationship between the fermentation product formation and carotenoid production due to shared precursors between these products can be observed. Therefore, removal of the identified fermentation product synthesis genes—*ldhA*, *frdA*, *poxB*, *pta*, *adhE*—is expected to direct more flux into the carotenoid synthesis pathway. A less intuitive target for disruption suggested by EMA is the pyruvate kinase; *pykF*. The *pykF* gene encodes for one of the two isoenzymes of pyruvate kinase which permits the inter-conversion between pyruvate and phosphoenolpyruvate. Deletion of *pykF* could block the further metabolism of the carotenoid precursors, pyruvate and glycerol-3-phosphate leading to an increased glycolytic supply for the carotenoid synthesis pathway. Likewise, the disruption of glucose 6-phosphate-1-dehydrogenase encoded by *zwf* diverts the carbon flux into the pool of carotenoid precursors and reduces carbon loss in form of carbon dioxide usually mediated by the pentose phosphate pathway (Sprenger, 1995). Deletion of malic enzymes *maeB* as suggested by EMA likely prevents carotenoid precursors from draining into the anaplerotic pathway.

Experimental results confirmed that the synthesis of carotenoid is tightly coupled with cell growth in the mutant CRT028 (Fig. 8). The coupling between carotenoid synthesis and cell growth can be used as a natural selection pressure for maintaining carotenoid synthesis plasmids in the cells. In the designed mutant, the production of diapolycopendial (DPL) and diapolycopendioic acid (DPA) was improved as compared with the wild-type under identical growth condition. It should be noted that a significant increase in the carotenoid production was observed only in the designed mutant where all inefficient carotenoid synthesis pathways were eliminated (Table 3). However, in the mutant, the specific growth rate is reduced. This is consistent with previous studies which showed a higher expression of carotenoid products in *E. coli* at reduced growth rates (Alper et al., 2006). The inverse relationship between production of carotenoid and biomass is also predicted by EMA. Experimental results also showed that the mutant requires a significantly higher amount of oxygen and reaches a lower cell density in comparison to the wild-type. The critical role of oxygen on carotenoid synthesis was also observed in the production of lycopene by Alper et al. (2006).

Testing the strain performance on the production of other carotenoids showed that the mutant CRT028 is better than the wild-type MG1655. This is anticipated due to the shared precursors of all carotenoid products. However, the mutant CRT028 performed at its best on the production of carotenoids DPL/DPA. This is also expected since the mutant was specifically optimized for the production of DPL/DPA. It is worth mentioning that although all carotenoids share the same precursors, each requires different types and numbers of cofactors. This suggested that different optimized strain may be required for an efficient production of different carotenoid product.

Based on the EMA prediction, mutant CRT028/pACMNOx can only function using a combination of the five remaining elementary modes with a predicted minimum carotenoid yield of 0.41 mg-carotenoid/g-glucose. However, experimental yield observed in the mutant was lower than the predicted value. It is likely that the inconsistency between observed yield and predicted possible yield may be a result of an incorrect number of NADP cofactor chosen. As shown in Fig. 5, the predicted yield of DPA and numbers of available elementary modes depend on number of NADP cofactor assumed for synthesis of one mole of DPA. If eight moles of NADP required per mole of DPA was chosen in the model instead, the predicted minimum yield of DPA in the mutant CRT028/pACMNOx would be 0.15 mg-carotenoid/g-glucose which is consistent with the experimental yield observed.

In addition, there were other factors that could possibly prevent the mutant from reaching a higher yield of carotenoid products. Based on overall reaction stoichiometry of elementary modes remaining in the mutant (Table 5), high oxygen level are required for product synthesis. It is likely that the supplied oxygen becomes limiting especially once the cell reaches a high cell density. Moreover, rate limiting steps in the downstream pathway of carotenoid biosynthesis could be responsible for preventing the cell from reaching higher product yields. There were several studies showing that over-expression of these downstream genes could lead to an increased yield. For example, production of carotenoids such as lycopene, torulene or β-carotene were found to be improved when the isoprenoid biosynthesis genes such as *dxs* or *idi* were overexpressed in *E. coli* (Lee et al., 2004). The product formation could also be inhibited by some unknown negative effects of the cell regulatory system. Future studies will be needed for investigating these factors to further improve the product yield and productivities. At this point, the improved performance of the mutant in this experimental study has confirmed that elementary mode analysis (EMA) is an efficient tool for rational strain development to enhance secondary metabolite production such as carotenoids.

Acknowledgments

This work has been supported in part by the Initiative for Renewable Energy and the Environment (IREE) and by the National Institutes of Health (NIH) Grant no. GM077529. The authors would like to thank Professor Claudia Schmidt-Dannert for kindly providing carotenoid plasmids and Erik Holtzapple for useful advice.

References

Albrecht, M., Takaichi, S., Misawa, N., Schnurr, G., Boger, P., Sandmann, G., 1997. Synthesis of atypical cyclic and acyclic hydroxyl carotenoids in *Escherichia coli* transformants. J. Biotechnol. 8, 177–185.

Albrecht, M., Misawa, N., Sandmann, G., 1999. Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids β-carotene and zeaxanthin. Biotechnol. Lett. 21, 791–795.

Alper, H., Yong-Su, J., Moxley, J.F., Stephanopoulos, G., 2005. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. Metab. Eng. 7, 155–164.

Alper, H., Miyaoku, K., Stephanopoulos, G., 2006. Characterization of lycopene-overproduing *E. coli* strains in high cell density fermentations. Appl. Microbiol. Biotechnol. 72, 968–974.

Ausubel, F., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1995. Short Protocols in Molecular Biology. Wiley & Sons, Inc., New York.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 2006-0008.

Bachmann, B.J., 1996. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In: F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, H.E. Umbarger, (Eds.), *Escherichia coli* and *Salmonella:* cellular and molecular biology. ASM Press, Washington, pp. 2460–2488.

Britton, G., S. Liaaen-Jensen, H. Pfander. Carotenoids. Spectroscopy. Birkauser Verlag, Basel, 1995, pp. 57–61.

Carlson, R., Srienc, F., 2004a. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: identification of reactions. Biotechnol. Bioeng. 85 (1), 1–19.

Carlson, R., Srienc, F., 2004b. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: creation of overall flux states. Biotechnol. Bioeng. 86 (2), 149–162.

Farmer, W.R., Liao, J.C., 2001. Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. Biotechnol. Prog. 17 (1), 57–61.

Ingraham, J.L., Maaloe, O., Neidhardt, F.C., 1983. In: Growth of the Bacterial Cell. Sinauer Associates Inc., Sunderland.

Lee, P.C., Mijts, B.N., Schmidt-Dannert, C., 2004. Investigation of factors influencing production of the monocyclic carotenoid torulene in metabolically engineered *Escherichia coli*. Appl. Microbiol. Biotechnol. 65 (5), 538–546.

Mijts, B.N., Lee, P.C., Schmidt-Dannert, C., 2005. Identification of a carotenoid oxygenase synthesizing acyclic xanthophylls combinatorial biosynthesis and directed evolution. Chem. Biol. 12, 453–460.

Pfeiffer, T., Sanchez-Valdenebro, I., Nuno, J.C., Montero, F., Schuster, S., 1999. Metatool: for studying metabolic networks. Bioinformatics 15 (3), 251–257.

Ruther, A., Misawa, N., Boger, P., Sandmann, G., 1997. Production of zeaxanthin in *Escherichia coli* transformed with different carotenogenic plasmids. Appl. Microbiol. Biotechnol. 48, 162–167.

Schuster, S., Fell, D., Dandekar, T., 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat. Biotechnol. 18, 326–332.

Sprenger, G., 1995. Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12. Arch. Microbiol. 164, 324–330.

Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E.D., 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420, 190–193.

Tao, L., Schenzle, A., Odom, J.M., Cheng, Q., 2005. Novel carotenoid oxidase involved in biosynthesis of 4,4′-diapolycopene dialdehyde. Appl. Environ. Microbiol. 71, 3294–3301.

Trinh, C.T., Carlson, R., Wlaschin, A., Srienc, F., 2006. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. Metab. Eng. 8 (6), 628–638.

Trinh, C.T., Unrean, P., Srienc, F., 2008. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. Appl. Environ. Microbiol. 74 (12), 3634–3643.

Vershinin, A., 1999. Biological functions of carotenoids—diversity and evolution. Biofactors 10, 99–104.

Wang, C.W., Oh, M.K., Liao, J.C., 1998. Engineered isoprenoid pathway enhances Astaxanthin production in *Escherichia coli*. Biotechnol. Bioeng. 62 (2), 235–241.

Wieland, B., Feil, C., Gloria-Maercker, E., Thumm, G., Lechner, M., Bravo, J.M., Poralla, K., Götz, F., 1994. Genetic and biochemical analyses of the biosynthesis of the yellow carotenoid 4,4'-diaponeurosporene of *Staphylococcus aureus*. J. Bacteriol. 176 (24), 7719–7726.