Alternative isoleucine synthesis pathway in cyanobacterial species

Bing Wu,^{1†} Baichen Zhang,^{2†} Xueyang Feng,^{1†} Jacob R. Rubens,³ Rick Huang,¹ Leslie M. Hicks,² Himadri B. Pakrasi^{1,3} and Yinjie J. Tang¹

Correspondence

Yinjie J. Tang
yinjie.tang@seas.wustl.edu

¹Department of Energy, Environmental and Chemical Engineering, Washington University, St Louis, MO 63130, USA

²Donald Danforth Plant Science Center, 975 N. Warson Rd, St Louis, MO 63132, USA

³Department of Biology, Washington University, St Louis, MO 63130, USA

Cyanothece sp. ATCC 51142 is an aerobic N₂-fixing and hydrogen-producing cyanobacterium. Isotopomer analysis of its amino acids revealed an identical labelling profile for leucine and isoleucine when *Cyanothece* 51142 was grown mixotrophically using 2-¹³C-labelled glycerol as the main carbon source. This indicated that *Cyanothece* 51142 employs the atypical alternative citramalate pathway for isoleucine synthesis, with pyruvate and acetyl-CoA as precursors. Utilization of the citramalate pathway was confirmed by an enzyme assay and LC-MS/MS analysis. Furthermore, the genome sequence of *Cyanothece* 51142 shows that the gene encoding the key enzyme (threonine ammonia-lyase) in the normal isoleucine pathway is missing. Instead, the cce_0248 gene in *Cyanothece* 51142 exhibits 53% identity to the gene encoding citramalate synthase (CimA, GSU1798) from *Geobacter sulfurreducens*. Reverse-transcription PCR indicated that the cce_0248 gene is expressed and its transcriptional level is lower in medium with isoleucine than in isoleucine-free medium. Additionally, a BLAST search for citramalate synthase and threonine ammonia-lyase implies that this alternative isoleucine synthesis pathway may be present in other cyanobacteria, such as *Cyanothece* and *Synechococcus*. This suggests that the pathway is more widespread than originally thought, as previous identifications of the citramalate pathway are limited to mostly anaerobic bacteria or archaea. Furthermore, this discovery opens the possibility that such autotrophic micro-organisms may be engineered for robust butanol and propanol production from 2-ketobutyrate, which is an intermediate in the isoleucine biosynthesis pathway.

Received 16 June 2009
Revised 23 October 2009
Accepted 28 October 2009

INTRODUCTION

Cyanothece sp. ATCC 51142 is an aerobic unicellular marine cyanobacterium that has a robust diurnal cycle for photosynthesis and nitrogen fixation and is also capable of evolving hydrogen (Colón-López & Sherman, 1998; Stöckel *et al.*, 2008; Toepel *et al.*, 2008). Previous transcriptional network studies on *Cyanothece* sp. have shown unique periodic changes in its central metabolism throughout the day–night cycle (Stöckel *et al.*, 2008). The genome of this strain has recently been sequenced (Welsh *et al.*, 2008) and the enzyme (threonine ammonia-lyase, EC 4.3.1.19) catalysing the conversion of threonine to 2-ketobutyrate in the normal isoleucine synthesis pathway is not

annotated based on KEGG pathway maps (<http://www.genome.jp/kegg/>). However, the key intermediate in isoleucine biosynthesis pathways, 2-ketobutyrate, can also be synthesized from citramalate. While most bacteria employ the threonine pathway to form isoleucine, some anaerobic bacteria and archaea, such as *Methanococcus jannaschii* and *Geobacter sulfurreducens*, can synthesize isoleucine from citramalate via condensation of acetyl-CoA and pyruvate catalysed by citramalate synthase (CimA) (Howell *et al.*, 1999; Risso *et al.*, 2008; Tang *et al.*, 2009b; Xu *et al.*, 2004).

In this study, we investigated the actual route for isoleucine synthesis in *Cyanothece* 51142 via ¹³C-based isotopomer analysis. These experiments consisted of feeding the culture a defined ¹³C-labelled substrate and measuring the isotopic enrichment in metabolites once the labelled carbon had percolated through the metabolic network. The isotopomer labelling information allowed us to decipher carbon transitions in the isoleucine synthesis pathways (Tang *et al.*, 2009a). The identification of such enzymes can

†These authors contributed equally to this work.

Abbreviations: CID, collision-induced dissociation; MRM, multiple reaction monitoring.

Two supplementary figures are available with the online version of this paper.

improve our understanding of carbon metabolism in *Cyanothece* 51142 as well as in other recently sequenced cyanobacterial strains. Furthermore, because an intermediate in the citramalate pathway is 2-ketobutyrate (the precursor of 1-propanol and 1-butanol: Atsumi & Liao, 2008), this discovery may have potential applications to CO₂-based biofuel research by making cyanobacteria a viable vehicle for the autotrophic production of butanol and propanol.

METHODS

Culture conditions. *Cyanothece* 51142 was grown in 150 ml Erlenmeyer flasks fed with ASP2 medium (Toepel *et al.*, 2008). The medium composition for 1 litre was as follows: Na₂EDTA, 0.03 g; K₂HPO₄, 0.05 g; KCl, 0.6 g; NaCl, 18 g; NaNO₃, 1.5 g; MgSO₄, 2.24 g; CaCl₂·2H₂O, 0.37 mg; FeCl₃·6H₂O, 3.9 mg; HCl, 0.1 µmol; TAPS, 1 g; TAPSO, 1 g; H₃BO₃, 34.3 mg; MnCl₂·4H₂O, 4.32 mg; ZnSO₄·7H₂O, 0.668 mg; CuSO₄·5H₂O, 3 µg; CoCl₂·6H₂O, 12 µg; Na₂MoO₄·2H₂O, 44 µg. The carbon source was [2-¹³C]glycerol (99 %, Cambridge Isotope Laboratories), at an initial concentration of 7 g l⁻¹. The strain was initially grown in 20 ml labelled culture medium (with 3 % inoculum volume). At the mid-exponential phase of growth, a 3 % inoculum was added to a 50 ml subculture containing the same labelled medium, which reduced the effect of unlabelled carbon from the initial stock. All cultures were grown aerobically at 30 °C under continuous light (50 µmol photons m⁻² s⁻¹) in 125 ml flasks shaking at 150 r.p.m.

Analytical measurements. During the growth of *Cyanothece* 51142, cell density was monitored on a UV-Vis spectrophotometer (Genesys, Thermo Scientific) at 730 nm. At each time point, samples were collected and glycerol concentrations were measured using a glycerol enzymic assay kit (R-Biopharm). When the growth of the strain approached the mid-exponential phase, the biomass was harvested by centrifugation (10 000 g, 10 min), and the cell pellet was hydrolysed in 6 M HCl at 100 °C for 24 h. After air-drying overnight, the dried samples (containing free amino acids) were derivatized in tetrahydrofuran and *N*-(tert-butyltrimethylsilyl)-*N*-methyltrifluoroacetamide (Sigma-Aldrich) at 70 °C for 1 h. Isotopomer measurements were made on a gas chromatograph (Hewlett Packard, model 6890, Agilent Technologies) equipped with a DB5-MS column (J&W Scientific) and a mass spectrometer (5975, Agilent Technologies). Four types of charged fragments were detected by GC-MS for the amino acids: the [M-57]⁺ or [M-15]⁺ group, which contains unfragmented amino acids; and the [M-159]⁺ or [M-85]⁺ group, which contains amino acids having lost α-carboxyl groups. Because the [M-57]⁺ peaks in leucine and isoleucine overlap with other peaks, the [M-15]⁺ group was used to obtain the isotopomer labelling information of the unfragmented leucine and isoleucine. When amino acids are derivatized for GC separation, derivatization groups introduce the natural isotopes, including ¹³C (1.13 %), ¹⁸O (0.20 %), ²⁹Si (4.70 %) and ³⁰Si (3.09 %). Published algorithms were used to correct for the effects of those natural isotopes on the mass distributions of amino acids (Wahl *et al.*, 2004).

LC-MS/MS (composed of a Shimadzu HPLC system with a LEAP PAL autosampler and a 4000 QTRAP with TurboIonSpray ion source, Applied Biosystems) was used to confirm the production of citramalate by *Cyanothece* 51142. A 5 ml sample of culture taken during the exponential growth phase was centrifuged for 10 min at 8000 g. The intracellular metabolites were extracted from the cell pellet using 2 ml cold methanol and 2 ml chloroform. Then 1 ml distilled water (0 °C) was added to partition the solution. The hydrophobic layer at the bottom was removed, and extracted

metabolites in the aqueous solution were lyophilized and reconstituted in 1 % formic acid. After vortexing and sonication, the samples were centrifuged at room temperature for 4 min at 13 200 g to remove water-insoluble components. The LC gradient was optimized by injecting a citramalate standard that was detected by MS in multiple reaction monitoring (MRM) mode (Heinig & Henion, 1999). Under the optimized LC-MS/MS conditions, 50 µl of test sample was injected into the LC-MS/MS. Polar acidic components in the sample were resolved by an Onyx Monolith column (300 mm × 4.6 mm, Phenomenex) in gradient elution mode. Negative electrospray ionization (ESI) mode was used for detecting citramalate in both MRM and MS/MS scan modes. Collision-induced dissociation (CID) spectra of the citramalate parent ion *m/z* 147 (M-H)⁻ and MRM parameters were optimized for sensitive detection.

An identical LC-MS setup was used for measurement of citramalate's labelling pattern with the following modifications: (1) the HPLC system was changed to the Agilent 1200 HPLC system; (2) the injection volume was set to 40 µl. After obtaining citramalate mass patterns, ¹³C-labelled citramalate was subjected to MS/MS fragment analysis using the QTRAP 4000 system. If citramalate was labelled, the pseudo-molecular parent ion of citramalate (*m/z* 147, no loss of carbons) and its fragment ions (*m/z* 57, 85 and 87) would have a higher mass-to-charge ratio. Such information could be used to determine the isotopic labelling pattern of citramalate.

Enzymic activity assay. The methods of Howell *et al.* (1999) and Risso *et al.* (2008) were used. A wet weight of 100–200 mg of cell pellet was suspended in 0.1 M TES buffer, pH 7.5. The cell extract was prepared by sonication (MISONIX) of the cell pellet for 3 min with a 30 s on/20 s off cycle. A 100 µl sample of cell extract was mixed with TES buffer (0.2 M, pH 7.5, 500 µl), pyruvate (10 mM, 100 µl) and acetyl-CoA (50 mM, 20 µl), then topped up with distilled water to a final volume of 1000 µl. Blank samples were prepared as for the test samples but without pyruvate. The resulting test sample and blank sample solutions were then incubated at 30, 45 or 60 °C in an oven for 2 h. At intervals of 20 min, a 100 µl test sample and a blank sample were taken from the incubator and added to a solution containing 50 µl 10 mM DTNB [5,5'-dithiobis(2-nitrobenzoic acid)] in 0.1 M Tris/HCl (pH 8.0), 70 µl 1 M Tris/HCl (pH 8.0), and distilled water to a final volume of 0.9 ml. The A₄₁₂ values for the test sample and blank sample were recorded. The production of SH-CoA was determined based on a standard curve generated with 2-mercaptoethanol. The enzyme activity was assayed by detecting the production of SH-CoA over 2 h. All the chemicals employed in this assay were from Sigma-Aldrich.

RNA extraction and reverse transcription PCR (RT-PCR). In order to confirm the expression of *cce_0248* (encoding CimA), *Cyanothece* 51142 was cultured in ASP2 medium with glycerol (7 g l⁻¹). In the exponential growth phase (OD₇₃₀ ~0.6), isoleucine was added to the culture to a final concentration of 10 mM. The control experiment (without addition of isoleucine) was performed in parallel and the both culture samples were collected after 30 min. The RT-PCR protocol was based on a previous report (Johnson *et al.*, 2005). In brief, total RNA was extracted from the frozen cell pellets using RNAwiz reagent (Ambion) according to the product manual. Contaminant DNA was removed by DNase treatment using an RQ1 RNase-Free DNase kit (Promega) in accordance with the manufacturer's protocol. The extracted RNA was transcribed to cDNA by a SuperScript II Reverse Transcriptase kit (Invitrogen), also according to the manufacturer's protocol. The forward primer TTAGCTGCAGGAACACGATG and reverse primer TCTCGAACAATGCGACTGAC were employed to amplify the *cce_0248* gene, the forward primer AGAGGATGAGCAGCCACACT and reverse primer TAATTCGGATAACGCTTGC to amplify the *rrn* 16S gene (i.e. 16S

rRNA) as a positive control gene, and the forward primer GACC-CCCATTAAGCGAGAA and reverse primer TTAACCAAGGAGGCGGATTT to amplify the *nifX* gene (as a negative control gene) in the cDNA by PCR using the Platinum Pfx DNA Polymerase kit (Invitrogen) (Welsh *et al.*, 2008). PCRs were conducted with the following cycle conditions: 2 min of activation of the polymerase at 94 °C followed by 30 cycles consisting of 1 min at 94 °C, 30 s at 58 °C and 2 min at 72 °C; finally, a 10 min extension process was performed at 72 °C. The final PCR product was observed directly on agarose gels after electrophoresis.

RESULTS AND DISCUSSION

Cyanothece 51142 exhibited a maximum specific growth rate of $\sim 0.9 \text{ day}^{-1}$ in ASP2 medium in the presence of glycerol during the exponential phase (Fig. 1). The glycerol was quickly utilized for biomass production during the exponential growth phase and most key amino acids were highly labelled (Fig. 2). For example, >90 % of alanine was labelled with one or more carbons, indicating that a significant amount of labelled glycerol was directed into the central metabolism (glycolysis, TCA and pentose phosphate pathway), which produced the building blocks for amino acids and other metabolites. In comparison, CO_2 (unlabelled carbon) fixation via the Calvin cycle was minimal when glycerol was present in the medium (only 7 % of alanine was unlabelled).

The isotopomer data from $[\text{M}-159]^+$ (loss of the α -carboxyl group) showed that the labelling profiles for leucine ($\text{M}_0=0.01$, $\text{M}_1=0.14$, $\text{M}_2=0.70$, $\text{M}_3=0.13$; M_0 , M_1 , M_2 ... are fractions of unlabelled, singly labelled and doubly labelled amino acids, respectively) and isoleucine ($\text{M}_0=0.01$, $\text{M}_1=0.14$, $\text{M}_2=0.71$, $\text{M}_3=0.13$) were similar (the standard error for isotopomer analysis is <0.05). Furthermore, the isotopomer data from $[\text{M}-15]^+$ (containing unfragmented amino acids) also showed identical labelling for leucine ($\text{M}_0=0.01$, $\text{M}_1=0.03$, $\text{M}_2=0.21$, $\text{M}_3=0.69$) and isoleucine ($\text{M}_0=0.01$, $\text{M}_1=0.03$, $\text{M}_2=0.24$, $\text{M}_3=0.67$). These data suggest that leucine and isoleucine are synthesized from the same precursors (i.e. pyruvate and acetyl-CoA), and thus that isoleucine

synthesis in *Cyanothece* 51142 takes place via an alternative pathway to the threonine-pyruvate pathway employed by most organisms. Further supporting this conclusion, isoleucine could not be labelled with three carbons ($[\text{M}-15]^+$; $\text{M}_3=0.67$) if threonine and pyruvate were its precursors: threonine was mostly labelled with one carbon ($[\text{M}-57]^+$, $\text{M}_1=0.64$) and alanine (pyruvate as its precursor) was also mostly labelled with one carbon (alanine $[\text{M}-57]^+$, $\text{M}_1=0.84$) (Fig. 2).

The citramalate pathway via citramalate synthase is an alternative to the common threonine pathway for the synthesis of isoleucine. Citramalate synthase (CimA), which catalyses the synthesis of the isoleucine precursor 2-ketobutyrate, has been discovered in other organisms via isotopomer analysis, enzyme chemistry or cloning methods (Table 1). Two types of genes encoding CimA have been identified and both genes are homologous to homocitrate synthase. CimA was previously reported in an archaeal species (*Methanococcus jannaschii*, gene MJ1392) (Howell *et al.*, 1999) and recently a new CimA gene homologue (*Geobacter sulfurreducens*, gene GSU1798) has been discovered (Risso *et al.*, 2008); further CimA identifications are listed in Table 1. According to the genome sequence of *Cyanothece* 51142 (from the Joint Genome Institute: <http://www.jgi.doe.gov>), the gene for threonine-ammonia lyase (EC 4.3.1.19) catalysing the conversion of threonine to 2-ketobutyrate is missing. We also confirmed the absence of this gene from the *Cyanothece* 51142 genome by BLAST analysis (Table 2). In contrast, BLAST search results indicate the presence of a complete citramalate pathway for isoleucine synthesis and the gene for CimA is identified as *cce_0248*, which has 53 % identity to the reported citramalate synthase gene (GSU1798) in *G. sulfurreducens*.

To confirm citramalate synthase activity in *Cyanothece* 51142, the specific activities of this enzyme were measured. The enzyme activities in crude cell extracts at elevated temperatures (30, 45 and 60 °C) were 8.54 ± 2.24 , 5.69 ± 1.79 and $2.79 \pm 0.07 \text{ nmol min}^{-1} (\text{mg dry biomass})^{-1}$, respectively. In comparison, the citramalate synthase from the thermophilic *M. jannaschii* shows higher thermostability (Howell *et al.*, 1999). However, this enzyme assay does not conclusively prove the presence of citramalate synthase (Risso *et al.*, 2008) and therefore high-sensitivity and selectivity LC-MS/MS was used to detect intracellular citramalate concentrations in *Cyanothece* 51142. Initially, the LC-MS/MS conditions were optimized using the signal of an authentic citramalate standard. MRM-based methods then confidently identified citramalate from cell extracts in our unlabelled cultures (Fig. 3). We spiked the cell extracts with an authentic citramalate standard and the spiked standard co-eluted with endogenous citramalate as a single peak. The measured intracellular concentration of citramalate [$0.12 \pm 0.03 \mu\text{mol} (\text{g dry biomass})^{-1}$] was smaller than the reported pools of central metabolites [such as α -ketoglutarate and malate, $\sim 2\text{--}4 \mu\text{mol} (\text{g dry biomass})^{-1}$] in cyanobacteria (such as *Synechocystis*) (Shastri & Morgan,

Fig. 1. Growth of *Cyanothece* 51142 and utilization of glycerol. ○, OD₇₃₀; □, glycerol concentration.

Fig. 2. An alternative isoleucine pathway in *Cyanotheca* 51142. The standard errors of the GC-MS data from labelling experiments are below 0.05 ($n=2$). * represents the labelled carbon. Kdc, 2-keto acid decarboxylase; Adh, alcohol dehydrogenase; LeuABCD, isopropylmalate-related enzymes. The detected labelling patterns in pyruvate, threonine, isoleucine and leucine exclude a biosynthetic pathway via threonine as an intermediate but are in line with the citramalate pathway.

2007). This lower concentration is consistent with the proposed role for citramalate as a secondary metabolite in *Cyanotheca* 51142. Furthermore, the citramalate labelling

pattern resulting from the tracer experiment (using $[2-^{13}\text{C}]\text{glycerol}$ as the carbon source) was also investigated. The spectra of the parent ion $(M-H)^+$ had a mass-to-

Table 1. Reported threonine-independent isoleucine synthesis pathways

Species	Reference(s)	Main methods	<i>cimA</i> similarity (GSU 1798)	<i>cimA</i> similarity (MJ1392)	Growth conditions
<i>Methanobacterium thermoautotrophicum</i>	Eikmanns <i>et al.</i> (1983)	¹⁴ C-labelling experiments	MTH1481 28 %	MTH723 58 %	Anaerobic
<i>Methanococcus jannaschii</i>	Howell <i>et al.</i> (1999)	<i>cimA</i> gene cloned and expressed in <i>E. coli</i>	ND	MJ1392 100 %	Aerobic (<i>E. coli</i>)
<i>Leptospira interrogans</i>	Charon <i>et al.</i> (1974); Westfall <i>et al.</i> (1983) Xu <i>et al.</i> (2004)	¹⁴ C-labelling experiments <i>cimA</i> gene cloned and expressed in <i>E. coli</i>	LIC11726 26 %	LIC11726 41 %	Aerobic (<i>E. coli</i>)
<i>Thermoproteus neutrophilus</i>	Schäfer <i>et al.</i> (1989)	¹³ C NMR	Tneu_0320 45 %	Tneu_0832 55 %	Anaerobic, thermophilic
<i>Ignicoccus hospitalis</i>	Jahn <i>et al.</i> (2007)	Enzyme assays; ¹³ C-labelling experiments	Igni_0645 45 %	Igni_0983 52 %	Anaerobic, thermophilic
<i>Geobacter sulfurreducens</i>	Risso <i>et al.</i> (2008)	¹³ C GC-MS and gene/enzyme assays	GSU1798 100 %	GSU1906 41 %	Anaerobic
<i>Geobacter metallireducens</i>	Tang <i>et al.</i> (2007)	¹³ C GC-MS	Gmet_1879 92 %	Gmet_1265 42 %	Anaerobic
<i>Serratia marcescens</i>	Kisumi <i>et al.</i> (1977)	Enzyme assays	Spro_0745 (26 %)*	Spro_0745 (37 %)*	Anaerobic
<i>Thermoanaerobacter</i> sp. X514	Feng <i>et al.</i> (2009)	¹³ C GC-MS; enzyme assays	Teth514_1204 49 %	Teth514_0415 45 %	Anaerobic, thermophilic
<i>Dehalococcoides ethenogenes</i> 195	Tang <i>et al.</i> (2009c)	¹³ C GC-MS; qPCR	DET0825 53 %	DET0830 41 %	Anaerobic

ND, Not determined.

*Sequence similarity is based on *Serratia proteamaculans* 568 because the *Serratia marcescens* sequence is not available.

charge ratio of 149 (from the ¹³C-labelled sample) instead of 147 (from the unlabelled sample), which indicates that citramalate was mostly labelled with two carbons in this

experiment. The two labelled carbons in citramalate were derived from pyruvate and acetyl-CoA (Fig. 2). The carbon labelling positions in citramalate were determined by

Table 2. BLAST searches of amino acid sequences for key genes similar to the target genes (GSU1798 in *G. sulfurreducens* and MJ1392 in *M. jannaschii*, encoding *CimA*) and threonine ammonia-lyase (b3772 in *E. coli* K-12) in several sequenced cyanobacterial species

Cyanobacterium	Strain	GSU1798		MJ1392		b3772	
		Gene	Identity (%)	Gene	Identity (%)	Gene	Identity (%)
<i>Cyanothece</i> sp.	ATCC 51142	cce_0248	53	cce_4008	41	cce_1455	28
	CCY 0110	CY 0110_21607	53	CY0110_30221	41	CY0110_21225	29
	PCC 7424	PCC7424_3413	53	PCC7424_1126	41	PCC7424_0305	34
	PCC 7425	Cyan7425_0344	25	Cyan7425_0344	41	Cyan7425_3269	50
	PCC 7822	Cyan7822DRAFT_4501	53	Cyan7822DRAFT_3533	41	Cyan7822DRAFT_1627	29
	PCC 8801	PCC8801_0449	55	PCC8801_0832	41	PCC8801_1025	28
	PCC 8802	Cyan8802DRAFT_1361	54	Cyan8802DRAFT_0632	41	Cyan8802DRAFT_3061	28
	PCC 8802	Cyan8802DRAFT_1361	54	Cyan8802DRAFT_0632	41	Cyan8802DRAFT_3061	28
<i>Synechococcus</i> sp.	<i>elongatus</i> PCC 6301	syc 1089_d	53	syc0145_d	40	syc1498_c	33
	<i>elongatus</i> PCC 7942	Synpcc7942_0428	53	Synpcc7942_1410	40	Synpcc7942_2612	33
	JA-2-3 Ba (2-13)	CYB_1307	51	CYB_1442	43	CYB_2298	27
	JA-3-3 Ab	CYA_0566	51	CYA_2223	43	CYA_0143	51
	PCC 7002	SYNPCC7002_A0730	51	SYNPCC7002_A1356	41	SYNPCC7002_A1616	50
	PCC 7002	SYNPCC7002_A0730	51	SYNPCC7002_A1356	41	SYNPCC7002_A1616	50
<i>Synechocystis</i> sp.	PCC 6803	sll1564	52	slr0186	41	slr2072	49

Fig. 3. LC-MS/MS spectra for citramalate: (a) MRM signals of authentic citramalate standard; (b) MRM signals of cell extracts; (c) MRM signals of cell extracts spiked with authentic citramalate standard; (d) MS/MS CID spectra from authentic citramalate standard; (e) MS/MS CID spectra from extracted citramalate in *Cyanothece* 51142.

analysing the fragments of the parent ion (see Supplementary Fig. S1, available with the online version of this paper). It was concluded that citramalate was labelled predominantly at the C-2 and C-4 positions, based on MRM experiments and MS/MS fragmentation of ^{13}C -labelled citramalate (parent ion m/z 149) (Fig. 2).

Finally, expression of *cce_0248*, the key gene in *Cyanothece* 51142 that encodes the putative citramalate synthase, was confirmed by RT-PCR. The RT-PCR products from the amplification of *cce_0248* and the 16S rRNA gene (as a positive control) were distinct and as expected, although the expression of *cce_0248* was lower than that of the 16S rRNA gene (Supplementary Fig. S2). The *nifX* gene (a gene involved in the fixation of atmospheric nitrogen) was not expressed under either of the culture conditions (with or without isoleucine), because the ASP2 medium contained nitrate as the nitrogen source. So the *nifX* gene was treated as a negative control gene. RT-PCR products also showed that the presence of isoleucine (10 mM) in the medium reduced the expression of *cce_0248* relative to the control experiments. This result could be explained by a previous report that citramalate synthase was subject to feedback inhibition by 5 mM isoleucine (Xu *et al.*, 2004).

Our findings led us to search for further homologues of CimA of *G. sulfurreducens* in other species and we found that homologues are present in a wide diversity of cyanobacteria. BLAST searches for citramalate synthase (GSU1798) and

threonine ammonia-lyase (b3772) were conducted in *Synechocystis* sp. PCC 6803, seven strains of *Cyanothece* sp. and 16 strains of *Synechococcus* sp., based on amino acid sequence similarity. The BLAST results (Table 2) indicate that at least six strains of *Cyanothece* sp. and five strains of *Synechococcus* sp. have high probabilities of employing the citramalate pathway to synthesize isoleucine rather than the normal isoleucine pathway via threonine ammonia-lyase (CimA identity >50%). *Synechococcus* JA-3-3 Ab and PCC 7002, and *Synechocystis* PCC 6803, may be capable of employing two different pathways for isoleucine synthesis. Twelve strains of *Prochlorococcus* sp. have also been examined (data not shown), and all of them appeared not to have the genes encoding the identical CimA based on the KEGG pathway map (<http://www.genome.jp/kegg/>); these strains may only have the normal isoleucine pathway via threonine ammonia-lyase. The predictions in Table 2 require further experimental verification. For comparison, the two other genes closest to GSU1798 in sequence identity within the cyanobacteria have relatively low sequence identities (20–30%; data not shown) to GSU1798; these genes encode 2-isopropylmalate synthase (which participates in the biosynthesis of L-leucine) and homocitrate synthase.

Conclusions

Our isotopomer information, enzyme activity measurement, metabolite analysis and gene expression data unambiguously

indicate that the citramalate pathway is utilized for isoleucine synthesis in *Cyanothece* 51142. Furthermore, genome sequences suggest that quite a few other cyanobacterial strains may also employ the citramalate pathway for isoleucine synthesis. To our knowledge, citramalate pathways have previously been identified only in some anaerobic bacteria and archaea. The discovery of this alternative isoleucine pathway in cyanobacteria may help us understand their unique metabolic regulation related to photosynthesis and nitrogen fixation. Moreover, the citramalate pathway may be utilized to synthesize 2-ketobutyrate, a precursor of both butanol and propanol. As such, the discovery of citramalate synthase in cyanobacteria may provide a new biofuel synthesis route utilizing autotrophic micro-organisms.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Shan Yi (University of California, Berkeley), Bert Berla, Jana Stöckel and LeeAnn Perry (Washington University in St. Louis) for their help in RT-PCR experiments. This study was supported in part by a DOE bioenergy research grant (DEFG0208ER64694) and in part by a grant from I-CARES (International Center for Advanced Renewable Energy and Sustainability) at Washington University in St. Louis to the Tang lab. Funding for the 4000 QTRAP LC-MS/MS was provided by NSF-MRI grant no. DBI-0521250 to the Hicks lab, and that for the Agilent GC-MS was provided by the Missouri Life Sciences Trust Fund to the Pakrasi lab.

REFERENCES

- Atsumi, S. & Liao, J. C. (2008). Directed evolution of *Methanococcus jannaschii* citramalate synthase for biosynthesis of 1-propanol and 1-butanol by *Escherichia coli*. *Appl Environ Microbiol* **74**, 7802–7808.
- Charon, N. W., Johnson, R. C. & Peterson, D. (1974). Amino acid biosynthesis in the spirochete *Leptospira*: evidence for a novel pathway of isoleucine biosynthesis. *J Bacteriol* **117**, 203–211.
- Colón-López, M. S. & Sherman, L. A. (1998). Transcriptional and translational regulation of photosystem I and II genes in light-dark and continuous-light-grown cultures of the unicellular cyanobacterium *Cyanothece* sp. strain ATCC 51142. *J Bacteriol* **180**, 519–526.
- Eikmanns, B., Linder, D. & Thauer, R. K. (1983). Unusual pathway of isoleucine biosynthesis in *Methanobacterium thermoautotrophicum*. *Arch Microbiol* **136**, 111–113.
- Feng, X., Mouttaki, H., Lin, L., Huang, R., Wu, B., Hemme, C. L., He, Z., Zhang, B., Hicks, L. M. & other authors (2009). Characterization of the central metabolic pathways in *Thermoanaerobacter* sp. X514 via isotope-assisted metabolite analysis. *Appl Environ Microbiol* **75**, 5001–5008.
- Heinig, K. & Henion, J. (1999). Fast liquid chromatographic-mass spectrometric determination of pharmaceutical compounds. *J Chromatogr B Biomed Sci Appl* **732**, 445–458.
- Howell, D. M., Xu, H. M. & White, R. H. (1999). (R)-Citramalate synthase in methanogenic archaea. *J Bacteriol* **181**, 331–333.
- Jahn, U., Huber, H., Eisenreich, W., Hugler, M. & Fuchs, G. (2007). Insights into the autotrophic CO₂ fixation pathway of the archaeon *Ignicoccus hospitalis*: comprehensive analysis of the central carbon metabolism. *J Bacteriol* **189**, 4108–4119.
- Johnson, D. R., Lee, P. K. H., Holmes, V. F., Fortin, A. C. & Alvarez-Cohen, L. (2005). Transcriptional expression of the *tceA* gene in a *Dehalococcoides*-containing microbial enrichment. *Appl Environ Microbiol* **71**, 7145–7151.
- Kisumi, M., Komatsubara, S. & Chibata, I. (1977). Pathway for isoleucine formation from pyruvate by leucine biosynthetic enzymes in leucine-accumulating isoleucine revertants of *Serratia marcescens*. *J Biochem* **82**, 95–103.
- Risso, C., Van Dien, S. J., Orloff, A., Lovley, D. R. & Coppi, M. V. (2008). Elucidation of an alternate isoleucine biosynthesis pathway in *Geobacter sulfurreducens*. *J Bacteriol* **190**, 2266–2274.
- Schäfer, S., Paalme, T., Vilu, R. & Fuchs, G. (1989). ¹³C-NMR study of acetate assimilation in *Thermoproteus neutrophilus*. *Eur J Biochem* **186**, 695–700.
- Shastri, A. A. & Morgan, J. A. (2007). A transient isotopic labeling methodology for ¹³C metabolic flux analysis of photoautotrophic microorganisms. *Phytochemistry* **68**, 2302–2312.
- Stöckel, J., Welsh, E. A., Liberton, M., Kunnvakkam, R., Aurora, R. & Pakrasi, H. B. (2008). Global transcriptomic analysis of *Cyanothece* 51142 reveals robust diurnal oscillation of central metabolic processes. *Proc Natl Acad Sci U S A* **105**, 6156–6161.
- Tang, Y. J., Chakraborty, R., Martin, H. G., Chu, J., Hazen, T. C. & Keasling, J. D. (2007). Flux analysis of central metabolic pathways in *Geobacter metallireducens* during reduction of soluble Fe(III)-NTA. *Appl Environ Microbiol* **73**, 3859–3864.
- Tang, Y. J., Martin, H. G., Myers, S., Rodriguez, S., Baidoo, E. E. & Keasling, J. D. (2009a). Advances in analysis of microbial metabolic fluxes via ¹³C isotopic labeling. *Mass Spectrom Rev* **28**, 362–375.
- Tang, Y. J., Sapra, R., Joyner, D., Hazen, T. C., Myers, S., Reichmuth, D., Blanch, H. & Keasling, J. D. (2009b). Analysis of metabolic pathways and fluxes in a newly discovered thermophilic and ethanol-tolerant *Geobacillus* strain. *Biotechnol Bioeng* **102**, 1377–1386.
- Tang, Y. J., Yi, S., Zhuang, W.-Q., Zinder, S. H., Keasling, J. D. & Alvarez-Cohen, L. (2009c). Investigation of carbon metabolism in “*Dehalococcoides ethenogenes*” strain 195 via isotopic and transcriptomic analyses. *J Bacteriol* **191**, 5224–5231.
- Toepel, J., Welsh, E., Summerfield, T. C., Pakrasi, H. B. & Sherman, L. A. (2008). Differential transcriptional analysis of the cyanobacterium *Cyanothece* sp strain ATCC 51142 during light-dark and continuous-light growth. *J Bacteriol* **190**, 3904–3913.
- Wahl, S. A., Dauner, M. & Wiechert, W. (2004). New tools for mass isotopomer data evaluation in ¹³C flux analysis: mass isotope correction, data consistency checking, and precursor relationships. *Biotechnol Bioeng* **85**, 259–268.
- Welsh, E. A., Liberton, M., Stöckel, J., Loh, T., Elvitigala, T., Wang, C., Wollam, A., Fulton, R. S., Clifton, S. W. & other authors (2008). The genome of *Cyanothece* 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. *Proc Natl Acad Sci U S A* **105**, 15094–15099.
- Westfall, H. N., Charon, N. W. & Peterson, D. E. (1983). Multiple pathways for isoleucine biosynthesis in the spirochete *Leptospira*. *J Bacteriol* **154**, 846–853.
- Xu, H., Zhang, Y. Z., Guo, X. K., Ren, S., Staempfli, A. A., Chiao, J., Jiang, W. & Zhao, G. (2004). Isoleucine biosynthesis in *Leptospira interrogans* serotype lai strain 56601 proceeds via a threonine-independent pathway. *J Bacteriol* **186**, 5400–5409.

Edited by: C. C. Zhang