Bioresources and Bioprocessing

Metabolic pathway analysis for in silico design of efficient autotrophic production of advanced biofuels

Pornkamol Unrean[1]* , Kang Lan Tee[2] and Tuck Seng Wong[2]

Abstract

Herein, autotrophic metabolism of *Cupriavidus necator* H16 growing on CO_2, H_2 and O_2 gas mixture was analyzed by metabolic pathway analysis tools, specifically elementary mode analysis (EMA) and flux balance analysis (FBA). As case studies, recombinant strains of *C. necator* H16 for the production of short-chain (isobutanol) and long-chain (hexadecanol) alcohols were constructed and examined by a combined tools of EMA and FBA to comprehensively identify the cell's metabolic flux profiles and its phenotypic spaces for the autotrophic production of recombinant products. The effect of genetic perturbations via gene deletion and overexpression on phenotypic space of the organism was simulated to improve strain performance for efficient bioconversion of CO_2 to products at high yield and high productivity. EMA identified multiple gene deletion together with controlling gas input composition to limit phenotypic space and push metabolic fluxes towards high product yield, while FBA identified target gene overexpression to debottleneck rate-limiting fluxes, hence pulling more fluxes to enhance production rate of the products. A combination of gene deletion and overexpression resulted in designed mutant strains with a predicted yield of 0.21–0.42 g/g for isobutanol and 0.20–0.34 g/g for hexadecanol from CO_2. The in silico-designed mutants were also predicted to show high productivity of up to 38.4 mmol/cell-h for isobutanol and 9.1 mmol/cell-h for hexadecanol under autotrophic cultivation. The metabolic modeling and analysis presented in this study could potentially serve as a valuable guidance for future metabolic engineering of *C. necator* H16 for an efficient CO_2-to-biofuels conversion.

Keywords: Elementary mode analysis, Flux balance analysis, In silico efficient strain design, Genetic deletion simulation, Genetic overexpression simulation

Introduction

The concept of upgrading CO_2 waste into metabolites of interest via biological conversion routes is gaining increasing attention from both academia and industry to mitigate climate change as well as to promote the sustainable concepts of bioprocessing (Thakur et al. 2018; Psarras et al. 2017). *Cupriavidus necator* H16 (also known as *Ralstonia eutropha*), a lithoautotrophic bacterium, is one of the recommended microbial platforms for CO_2 bioconversion due to its high CO_2 fixation efficiency, versatile genetic manipulation, and ease of operation and

flexibility (Yu 2018; Tee et al. 2017). The organism is also considered as a valuable expression platform for many recombinant products. During autotrophy, *C. necator* H16 can fix CO_2 via the Calvin–Benson–Bassham cycle (CBB) and utilize it as carbon source for supporting growth and producing a number of bioproducts, e.g., polyhydroxybutyrate (Fukui et al. 2014), organic acids (Lee et al. 2016) and alcohols (Li et al. 2012; Lu et al. 2012), while H_2 is oxidized as energy source. Its ability to capture and utilize CO_2 for bioproducts makes the organism suitable for a number of biotechnological applications. Improving the efficiency of *C. necator* H16 for CO_2-to-bioproduct conversion is, therefore, necessary. One strategy for improving strain performance is systems metabolic engineering by utilizing metabolic modelling

*Correspondence: pornkamol.unr@biotec.or.th
[1] National Center for Genetic Engineering and Biotechnology, Pathum Thani 12120, Thailand
Full list of author information is available at the end of the article

Unrean *et al. Bioresour. Bioprocess.*   (2019) 6:49

Page 2 of 11

and pathway analysis tools for guiding in silico strain design (Chae et al. 2017; Blazeck and Alper 2010).

Analysis of metabolic characteristic and optimizing cellular phenotype at systems level could be achieved using pathway analysis tools. The metabolic pathway analysis techniques commonly used for characterizing cellular metabolism are flux balance analysis (FBA) and elementary mode analysis (EMA). Readers are referred to previous publications (Schuster et al. 2000; Orth et al. 2010) for detailed description of these computation tools. Both techniques are based on the balance of stoichiometric reactions within the reaction network of cellular metabolism and pseudo-steady-state assumption. FBA utilizes convex analysis and imposes an objective function to determine a pathway solution for cellular metabolism. EMA, on the other hand, identifies all possible pathway solutions existing in a metabolic network without imposing any objective function or any fixed fluxes. Metabolic modelling has been proven as an effective approach for characterizing cell's metabolic capabilities and predicting cellular phenotype in response to various environmental and genetic changes as continuously demonstrated in numerous studies (Curran et al. 2012; Zhao et al. 2013; Henson and Hanly 2014). The pathway topology and its metabolic capability identified by EMA and FBA have provided supporting information for systems metabolic engineering to enhance strain performance (Chae et al. 2017).
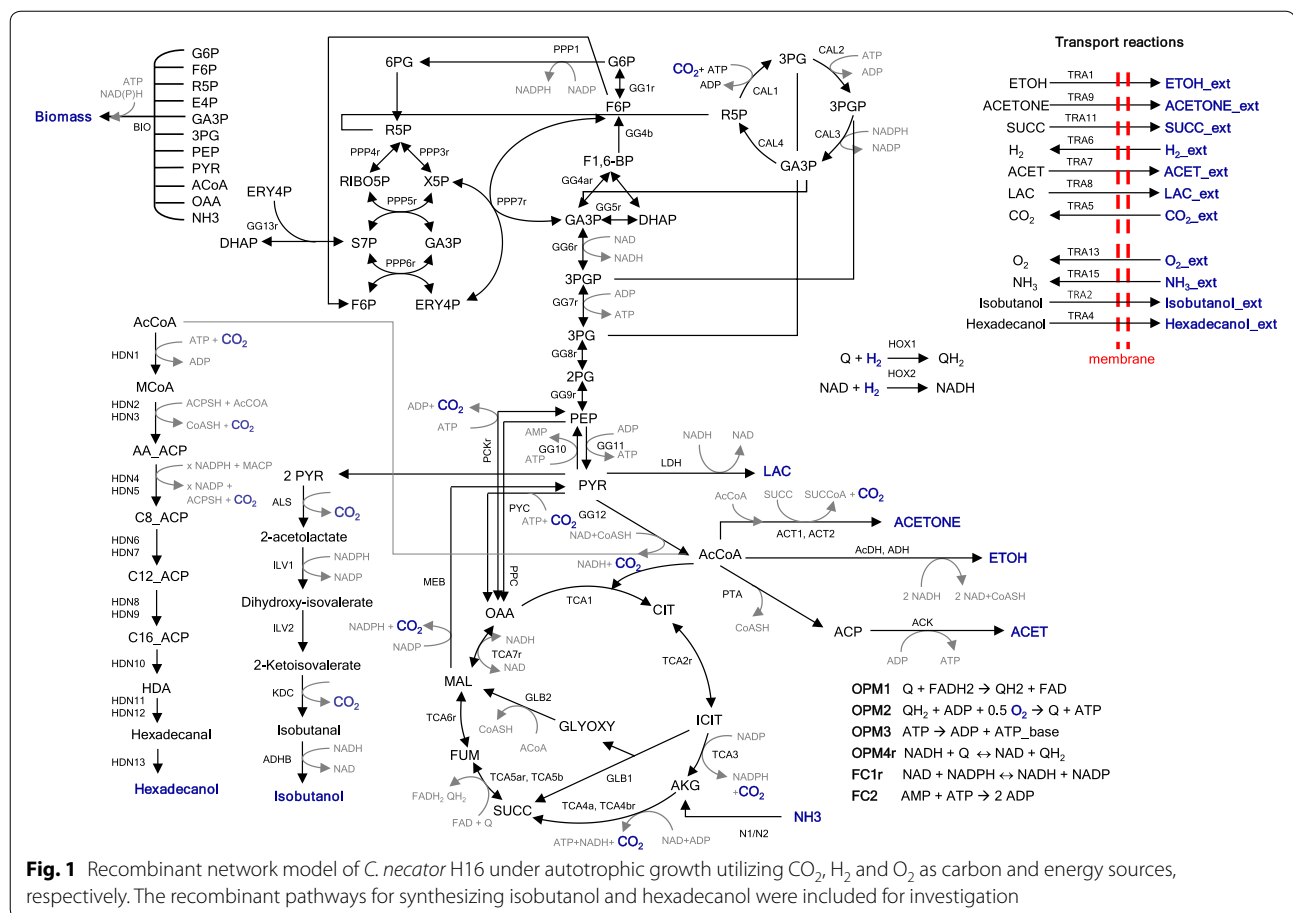
Although a number of metabolic pathway analysis for *C. necator* H16 have been reported, these analyses focused mainly on the metabolic network when the organism was cultivated in heterotrophic mode in which organic compounds were used as carbon source (Park et al. 2011a, b; Lopar et al. 2014; Ternon et al. 2014; Alagesan et al. 2018). On the contrary, this work aims at analyzing the organism's metabolism under autotrophic growth with CO_2, H_2 and O_2 gas input mixture as carbon and energy sources, respectively. The pathway structure of *C. necator* H16 growing on CO_2, H_2 and O_2 substrates is assessed by EMA and FBA tools for systematically designing efficient strain. As case studies, two mutants optimized for the production of short-chain (isobutanol) and long-chain (hexadecanol) alcohols, respectively, from CO_2 at high yield and high productivity are designed by the pathway analysis tools. These alcohols are chosen as products of interest since they can be used as drop-in biofuels to replace fossil fuel for ground and aviation transportation (Li et al. 2012; Guo et al. 2016). Specifically, EMA is exploited to identify target gene for deletion and cultivation condition, while FBA is utilized to determine target gene for overexpression for pushing and pulling metabolic fluxes of *C. necator* H16 towards target products.

Materials and methods

Metabolic model construction of *C. necator* H16 growing autotrophically

A recombinant metabolic network of *C. necator* H16 was constructed based on the previously published literatures on its metabolism (Schäferjohann et al. 1993; Pohlmann et al. 2006; Bruland et al. 2010; Kohlmann et al. 2011; Park et al. 2011a, b; Ternon et al. 2014), and available pathway databases (Kyoto Encyclopedia of Gene and Genomes, KEGG (http://www.genome.jp/kegg); Biocyc Database Collection project, BioCyc (http://www.biocyc.org)). The constructed model represented *C. necator* H16 growing autotrophically on CO_2, H_2 and O_2 gas mixture. Reaction describing PHB biosynthesis was excluded from the model presuming PHB-negative strain of *C. necator* H16 was used as host for synthesis of recombinant biofuels (Raberg et al. 2014). The recombinant production of two target biofuels, isobutanol and hexadecanol, served as the basis for the model simulation case studies. For all the simulations of isobutanol-producing *C. necator*, the recombinant pathway for hexadecanol was constrained to zero and vice versa. That is the isobutanol-producing *C. necator* contains only a recombinant plasmid expressing isobutanol biosynthesis pathway; whereas, the hexadecanol-producing *C. necator* contains only a recombinant plasmid expressing hexadecanol biosynthesis pathway. Figure 1 shows the recombinant metabolic map of the biofuels-producing *C. necator* analyzed in this study. Details on the constructed metabolic model are described in Additional file 1. For simplicity, linear series of reactions were lumped together as one. Thus, the reactions might not necessarily represent one gene or enzyme. Information on reaction reversibility was taken from the databases. External metabolites involved in the reaction network serving as source or sink were CO_2, H_2, O_2, ethanol, acetate, lactate, acetone, succinate, NH_4^+ and recombinant products (isobutanol, hexadecanol). These metabolites could be accumulated; hence, their net synthesis/consumption rates were not restricted to zero. All internal metabolites, however, were defined as metabolites with no accumulation; hence, their net synthesis/consumption rates were being constrained by steady-state assumption.

The biomass synthesis pathway, reconstructed by accounting for metabolite drain from the central metabolic pathways, was based on reaction stoichiometry reported in Ternon et al. 2014, representing the molar requirements of biosynthetic precursors and redox cofactors required for the production of cell components, e.g., amino acids, nucleotides, lipids, and carbohydrates (see Additional file 1 for details). Energy requirements for biomass production, including polymerization and synthesis of biosynthetic precursors, were also included. The maintenance energy represented by excess ATP to be

Fig. 1 Recombinant network model of *C. necator* H16 under autotrophic growth utilizing CO_2, H_2 and O_2 as carbon and energy sources, respectively. The recombinant pathways for synthesizing isobutanol and hexadecanol were included for investigation

consumed for maintenance processes was treated separately. The synthesis pathways for isobutanol and hexadecanol were based on the recombinant pathways reported in Li et al. 2012 and Guo et al. 2016, respectively. The synthesis pathway for isobutanol was a combination of valine biosynthesis and recombinant expression of ketoacid decarboxylases and alcohol dehydrogenases. The hexadecanol biosynthesis was diverted from fatty acid biosynthesis via recombinant expression of fatty acyl-CoA reductase and hexadecanol dehydrogenase.

Metabolic network analysis

The recombinant network was analyzed by elementary mode analysis (EMA) and flux balance analysis (FBA) based on stoichiometric and thermodynamic constraints.

Elementary mode analysis

EMA decomposed the recombinant network into all possible, unique and balanced metabolic pathways, allowing for the calculation of a phenotypic space from all the pathways. EMA was simulated using publicly available program METATOOL (http://pinguin.biologie.uni-jena.de/bioinformatik/networks/metatool/) and network

characteristics were examined by sorting and filtering through all identified paths based on relative flux inputs and/or product yields using Microsoft Excel spreadsheet (Microsoft). Pathway filtering was performed (1) during genetic perturbation by selecting for remaining pathways after each gene deletion, and (2) during different gas input by selecting for remaining pathways after restricting relative gas input. Yield was defined as the mass flux ratio of the amount of product produced per CO_2 substrate consumed. The predicted yields were sorted to identify maximum metabolic capacity of the studied network and the efficiency of each identified metabolic pathway. Phenotypic plot was made on the basis of relative gas uptake, computed as the value of the gas input rate per the value of total gas input, sum of CO_2, H_2 and O_2 uptakes. Reaction requirements of efficient pathways, maximum yielding paths, were categorized as either (1) core reactions, ones necessary for product synthesis during growth and/or non-growth autotrophic conditions, or (2) inactive reactions, ones not requiring for product synthesis. The core reactions were analyzed in gene overexpression simulation, while the inactive reactions were analyzed in gene deletion simulation.

Flux balance analysis

FBA of the recombinant strains of *C. necator* H16 was performed by solving pseudo-steady-state stoichiometric balance equations (S*v = 0, where S is a stoichiometric coefficient matrix of metabolites in each reaction and v is the flux of each reaction within the network.) using linear programming (LP) optimization based on MATLAB (Mathworks, Natick, MA, USA) and Mosek optimization toolbox (Mosek ApS, Copenhagen, Denmark). A flux solution of intracellular and extracellular metabolites was obtained based on maximization of growth constrain which was used as the objective based on the evolutionary pressure on the cell growing under autotrophy. The changes of metabolic fluxes under various genetic perturbations were examined by flux correlation (FC), to identify targets for overexpression. The FC was calculated as

$$FC = \frac{dq_P}{dR_i}, \tag{1}$$

which is the relative effect between increasing flux (R_i) through gene overexpression simulation and observed phenotype through predicted product synthesis rate (q_P) by FBA. The reaction with maximum FC value would be considered as bottleneck within the recombinant network, hence being targeted for overexpression to enhance product synthesis.

Results and discussion

Construction of *C. necator* metabolic network model

The metabolic network model of *C. necator* H16 under autotrophic growth was based on PHB-negative strain used as host organism for recombinant production of isobutanol and hexadecanol; therefore, native PHB biosynthesis pathway was excluded. Figure 1 presents the constructed model, composed of 21 reversible reactions, 57 irreversible reactions, 75 intracellular metabolites and 11 extracellular metabolites exchanged through cell membrane. To simulate autotrophic growth, the metabolic network model includes CO_2 fixation via Calvin–Benson–Bassham cycle, H_2 and O_2 oxidation via membrane-bound soluble hydrogenases and oxidative phosphorylation, respectively (Schwartz et al. 2003; Kohlmann et al. 2011). The metabolic model for supporting growth and energy synthesis also consists of known central pathways of the organism including glycolysis, gluconeogenesis, pentose phosphate pathway, anaplerotic, tricarboxylic acid cycle, glyoxylate shunt, energy generating reactions accounting for maintenance energy and ATP hydrolysis, mixed acid fermentation for ethanol, acetone, acetate, succinate and lactate, NH_4 assimilation and exchanged reactions. Biosynthesis pathways for amino acids, nucleotides, phospholipid, carbohydrates, cofactors and vitamins were lumped together based on

reaction stoichiometry reported in Ternon et al. 2014. To construct isobutanol/hexadecanol-producing *C. necator* metabolic model, the biosynthesis pathways of isobutanol through valine biosynthesis, and hexadecanol through fatty acid biosynthesis were incorporated into the autotrophic model of *C. necator*, which was used as the host for synthesizing the short-chain and long-chain alcohols (Fig. 1). Detailed synthesis pathways for isobutanol via recombinant expression of keto-acid decarboxylases/ alcohol dehydrogenases and hexadecanol biosynthesis via recombinant expression of fatty acyl-CoA reductase/hexadecanol dehydrogenase were provided in Li et al. 2012 and Guo et al. 2016, respectively. Metabolic network model with corresponding enzyme and gene for each reaction is given in Additional file 1.
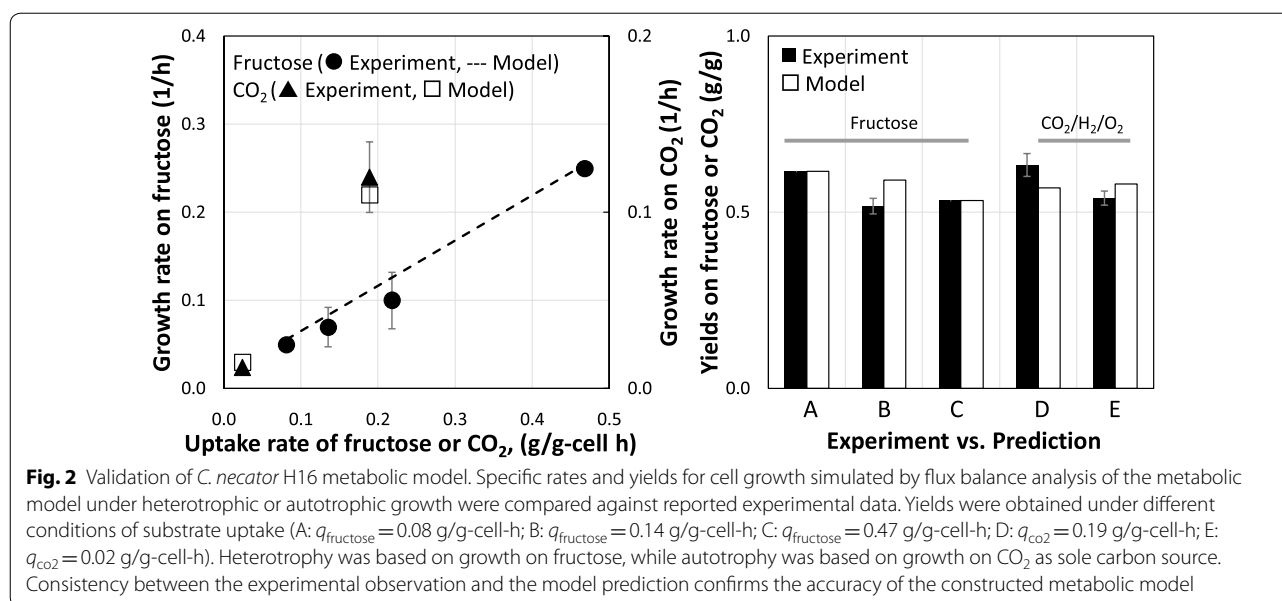
Growth phenotypes comparison: model prediction vs. experiment

Flux balance analysis under growth constraint was first performed to validate the metabolic model by comparing predicted value of cell growth with experimental observations reported in previous literatures (Fig. 2). The constructed model in Fig. 1 was simulated for both autotrophic and heterotrophic growths. Under heterotrophy, fructose was used as sole carbon source. Fructose transport and conversion reactions were incorporated into the model to simulate heterotrophic growth as follows,

(F1) fructose transport: fructose (extracellular)
→ fructose (intracellular)

(F2) fructose conversion: fructose
→ fructose 6 − phosphate.

For simulation of heterotrophic growth, the uptake rates of CO_2 and H_2 were constrained to zero and reactions related with CO_2 fixation cycle and H_2 oxidation were inactivated. In the simulation, the uptake rate of fructose was set at different values; then specific growth rates and cell yields under different fructose uptake rates were predicted by the constraint-based FBA. The in silico results were then compared with experimental data of chemostat culture under varying dilution rates (Park et al. 2011a, b). Likewise, for autotrophic growth on CO_2/ H_2/O_2 gas mixture, specific growth rates and yields were compared between in silico analysis and experimental data reported in Ishizaki and Tananka 1990, 1991. The simulation results could successfully represent growth phenotypes of *C. necator* under both heterotrophy and autotrophy, hence validating the constructed metabolic model in this study.

Unrean *et al. Bioresour. Bioprocess.* (2019) 6:49

Page 5 of 11

Fig. 2 Validation of *C. necator* H16 metabolic model. Specific rates and yields for cell growth simulated by flux balance analysis of the metabolic model under heterotrophic or autotrophic growth were compared against reported experimental data. Yields were obtained under different conditions of substrate uptake (A: $q_{fructose} = 0.08$ g/g-cell-h; B: $q_{fructose} = 0.14$ g/g-cell-h; C: $q_{fructose} = 0.47$ g/g-cell-h; D: $q_{co2} = 0.19$ g/g-cell-h; E: $q_{co2} = 0.02$ g/g-cell-h). Heterotrophy was based on growth on fructose, while autotrophy was based on growth on CO_2 as sole carbon source. Consistency between the experimental observation and the model prediction confirms the accuracy of the constructed metabolic model

The model was also capable of accurately predicting cell vitality under genetic perturbation. For example, GG8r (phosphoglycerate mutase) was predicted by constrain-based FBA as an essential gene, with its deletion leading to cell death under autotrophic growth, which agrees well with previous experimental study by Reutz et al. 1982. Double deletion of GG4b and CAL4 (fructose-1,6-/sedo-heptulose-1,7-bisphosphatase) was also predicted to be lethal during autotrophy, consistent with experimental data (Schwartz et al. 2009). The validated stoichiometric model was then further analyzed in the next section for recombinant production of isobutanol and hexadecanol.

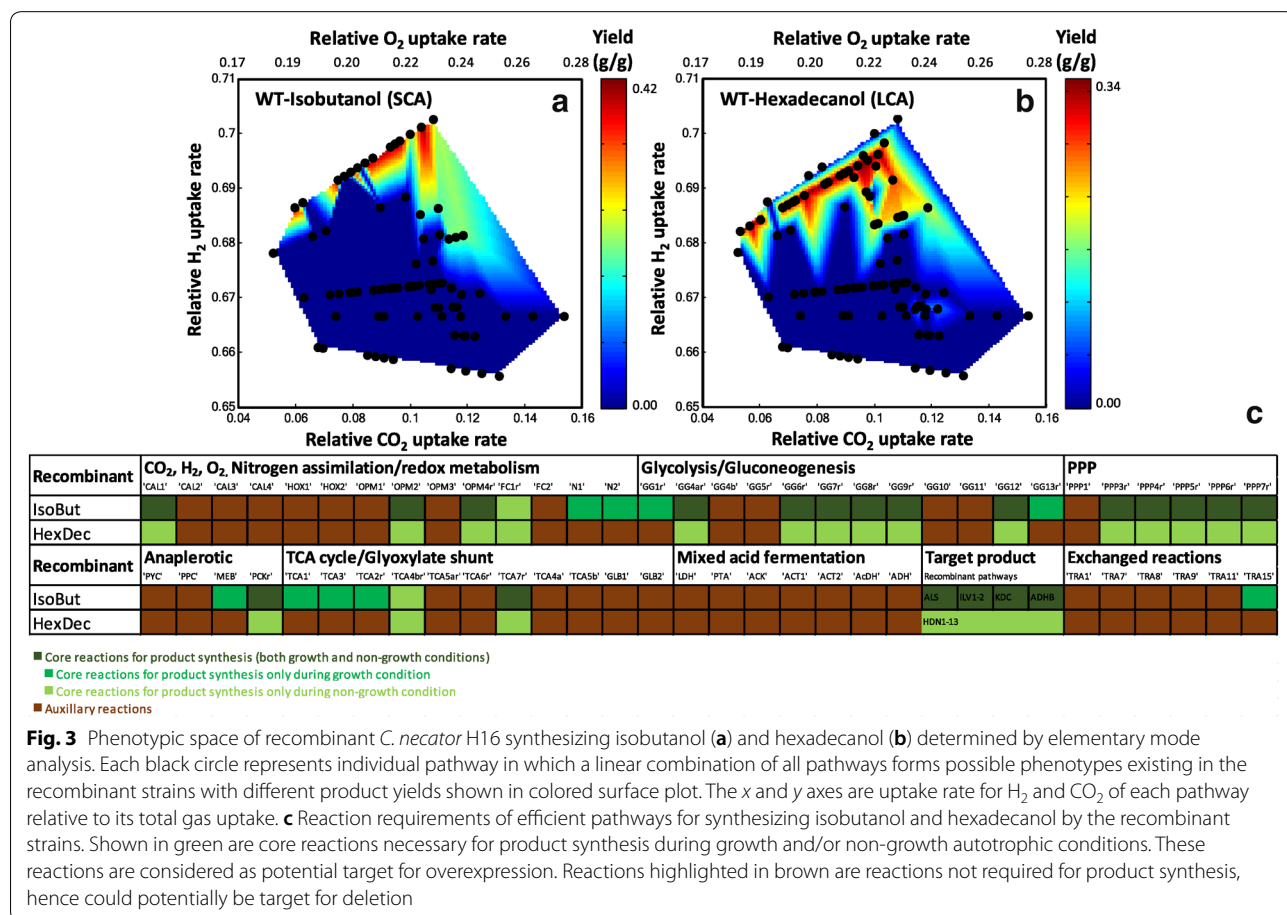**Determination of phenotypic space and pathway efficiency**

The network model of *C. necator* H16 grown on CO_2, H_2 and O_2 gas mixture was analyzed by EMA to determine all possible pathways existing in the cell metabolism and their metabolic capabilities. The analysis was based on predicted yield range for the production of isobutanol and hexadecanol as summarized in Table 1. Analysis of pathway topology inherent to the cell metabolic network also identified the most efficient paths for the product synthesis during autotrophy. EMA yielded a total of 8940 existing pathways in the recombinant network for isobutanol synthesis, and 9264 existing pathways in the recombinant network for hexadecanol synthesis. Each pathway represented a unique, possible path with balanced metabolites and cofactors and a non-negative linear combination of these identified pathways formed possible phenotypic space of a cell, known as functioning metabolism. Figure 3a, b show available phenotypes of the recombinant strains for isobutanol and hexadecanol production as a function of relative CO_2, O_2 and H_2 gas

input with the product yield ranges of 0–0.42 g/g and 0–0.34 g/g for produced isobutanol and hexadecanol per consumed CO_2, respectively. Cell yield under autotrophic growth was predicted at 0.22–0.57 g of biomass produced per gram of CO_2 consumed.

As summarized in Table 1, there existed 1068 paths supporting isobutanol synthesis equivalent to a 12% of all available pathways (*P*-frac = 12%), and 1392 paths supporting hexadecanol synthesis which was a 15% of all available pathways (*P*-frac = 15%). These pathways illustrated the flexibility of cellular metabolism to adapt itself to genetic changes using alternative operational pathways that provide optimal fitness. Interestingly, there was no existing pathway for hexadecanol synthesis during growth condition meaning that the product synthesis and cell growth could not be coupled. This was expected as hexadecanol was derived from fatty acid which has often been reported to accumulate during nitrogen-limiting, non-growth condition (Chen et al. 2015; Marella et al. 2018). Analyzing the topological structure of these metabolic pathways permitted the prediction of genetic alterations in the metabolism that would push and pull metabolic fluxes towards the production of desired products. Hence, the in silico strain design in the next section would focus on redirecting cell metabolism via combined gene deletion and overexpression simulations to enforce the cell into operating via the most efficient pathways resulting in increased percentage of existing pathways for product synthesis (*P*-frac = maximum) and higher product yield range (Y_P = maximum). Assessing the product synthesizing pathways for their yield efficiency revealed non-essential (inactive) and essential (core) reactions in the production of isobutanol and hexadecanol at maximum

Unrean *et al. Bioresour. Bioprocess.* (2019) 6:49

Page 6 of 11

Table 1 Existing pathways and predicted yields performance by recombinant strains and designed mutants with multiple gene deletions for maximizing isobutanol and hexadecanol synthesis

Strains	Total pathways	*P*-syn pathways[a]	*P*-frac (%)[b]	Y_X (g/g)[c]	Y_P (g/g)[d]
Isobutanol (SCA)					
WT-Isobutanol	8940	1068	12	0.22–0.57	0–0.42
ΔGG5rΔACT1ΔAcDH[e]	1476	252	17	0.22–0.57	0–0.42
ΔGG5rΔACT1ΔAcDH[e] under gas feeding control at $fq_{H_2} > 0.67$[f]	252	252	100	0.31	0.21–0.42
Hexadecanol (LCA)					
WT-Hexadecanol	9264	1392	15	0.22–0.57	0–0.34
ΔCAL4ΔACT1ΔAcDH[e]	3420	972	28	0.22–0.57	0–0.34
ΔCAL4ΔACT1ΔAcDH[e] under gas feeding control at $fq_{H_2} > 0.67$[f]	864	864	100	0	0.20–0.34

[a] Available metabolic pathways for product synthesis

[b] Percent of product synthesizing pathways from total available metabolic pathways

[c] Predicted cell yield during growth condition

[d] Predicted product yield during non-growth condition

[e] Genes and enzymes corresponding to deleted reactions are GG5r, triosephosphate isomerase (*tpiA*); ACT1, acetoacetyl-CoA thiolase/Acetyl-CoA acetyltransferase (*phaA*); AcDH, acetaldehyde dehydrogenase 1/2 (*mhpF/acoD*); CAL4, fructose/Sedoheptulose bisphosphatase (*cbbF*)

[f] fq_{H_2} is fraction of H_2 gas input relative to total gas input of CO_2, H_2 and O_2 mixture

Fig. 3 Phenotypic space of recombinant *C. necator* H16 synthesizing isobutanol (**a**) and hexadecanol (**b**) determined by elementary mode analysis. Each black circle represents individual pathway in which a linear combination of all pathways forms possible phenotypes existing in the recombinant strains with different product yields shown in colored surface plot. The *x* and *y* axes are uptake rate for H_2 and CO_2 of each pathway relative to its total gas uptake. **c** Reaction requirements of efficient pathways for synthesizing isobutanol and hexadecanol by the recombinant strains. Shown in green are core reactions necessary for product synthesis during growth and/or non-growth autotrophic conditions. These reactions are considered as potential target for overexpression. Reactions highlighted in brown are reactions not required for product synthesis, hence could potentially be target for deletion

Unrean *et al. Bioresour. Bioprocess.*   (2019) 6:49

Page 7 of 11

yield under growth and non-growth autotrophic cultivations (Fig. 3c). These nonessential reactions that were inactive during the product synthesis were examined in term of their potential for gene deletion in the in silico strain design. Likewise, to identify potential genetic overexpression, core reactions necessary for efficient production of the two products were analyzed.

In silico genetic perturbation simulation

The validated recombinant network was finally used to design the strategies via genetic perturbations to identify targets for metabolic engineering essential for the most efficient production of isobutanol, short-chain alcohol (SCA), and hexadecanol, long-chain alcohol (LCA), in *C. necator* H16 growing autotrophically.

Gene deletion target enforcing towards high yield

The pathway information provided by EMA allowed the evaluation of gene deletion effects on cell phenotype. In silico gene deletion simulations were performed by restricting reactions and its corresponding genes for deletion to zero and the phenotype of that specific knockout mutant, which was represented by a non-negative linear combination of remaining paths when that reaction was deleted, was predicted. Figure 4 shows gene deletion effects on cell viability (Y_X, biomass yield), products yield (Y_p, isobutanol and hexadecanol yield) and fraction of total remaining pathways, after each deletion of reaction and its corresponding genes, that produced products, isobutanol and hexadecanol (*P*-frac, % of total pathways). The effects of multiple gene deletions in sequential manner were also reported. In Fig. 4, yield was presented either in maximal value or in mean value, an average of minimal and maximal yields of all remaining pathways. Gene deletion which resulted in zero cell yield (mean $Y_X = 0$) was considered lethal and therefore, these reactions and their corresponding genes could not be deleted from the cell. The process of selecting combination of multiple gene deletions was started with analyzing single gene deletion. Among the set of all individual deletions,
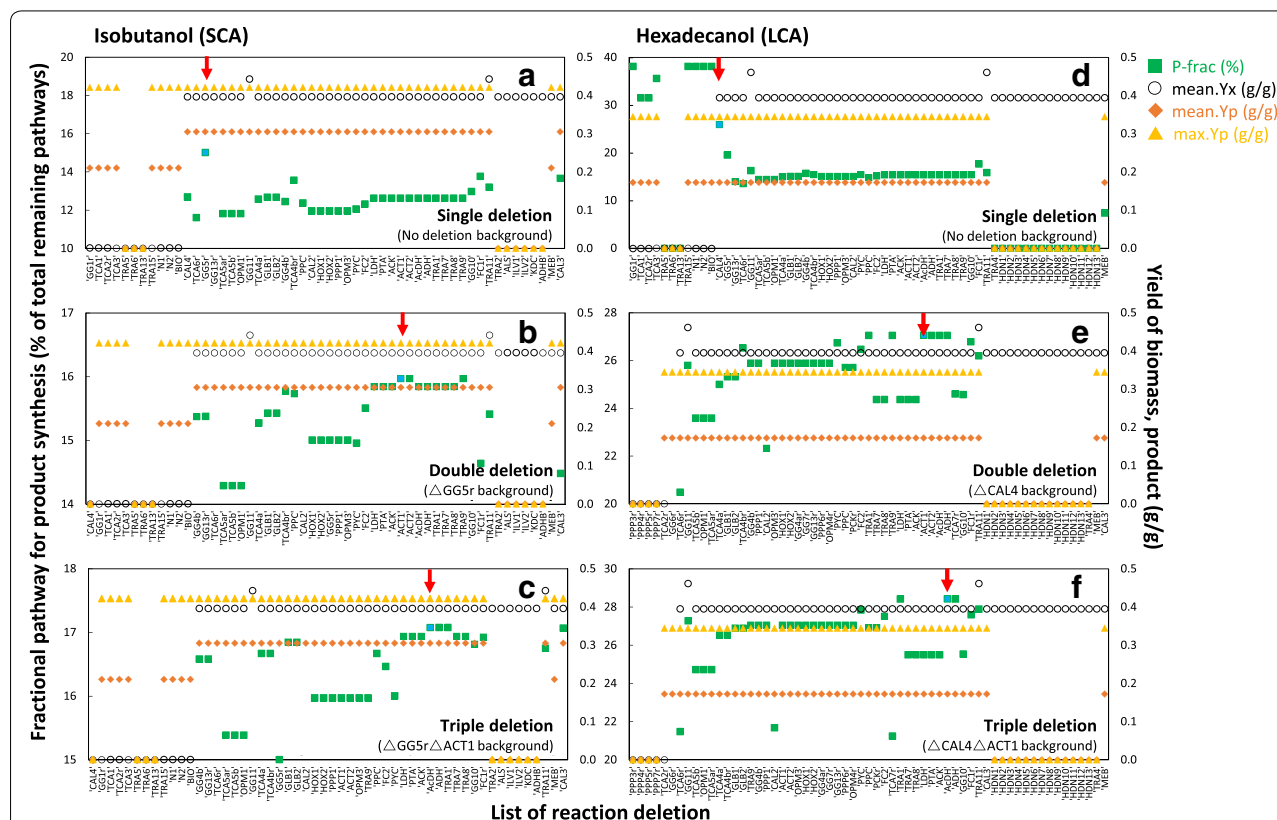

Fig. 4 Multiple gene deletion simulation by elementary mode analysis. Single (**a**, **d**), double (**b**, **e**) and triple (**c**, **f**) gene deletion simulation on the recombinant strains for isobutanol and hexadecanol synthesis were implemented to identify combination of deleted genes that eliminate inefficient pathways for product synthesis, while keeping efficient ones hence redirecting the cell's phenotypic space towards maximal product yield. Three criteria for identifying target genes are (1) product yields (maximal and mean Y_P g/g, yellow triangle, orange diamond), (2) fractional pathways for product synthesis (P-frac %, green square) and (3) biomass yield (mean Y_X g/g, empty circle) after each gene deletion simulation. Mean value is an average of minimal and maximal values. Target genes for deletion (indicated by red arrow) would reshape phenotypic space by pushing towards high product yield (max *P*-frac, max Y_P), while retain cell viability ($Y_X > 0$)

Unrean *et al. Bioresour. Bioprocess.* (2019) 6:49

Page 8 of 11

the target genes were identified when elimination of that gene led to

1. Product yield (Y_P) = maximum, to push towards high yield by keeping the most efficient paths for product synthesis and eliminating inefficient ones,
2. Percent of fractional remaining paths for product (*P*-frac) = maximum, to redirect phenotypic space by retaining the largest possible remaining paths towards product synthesis,
3. Biomass yield (Y_X) > 0, to maintain cell viability.

This approach allowed the determination of gene deletion targets which pushed the largest number of pathways towards products synthesis with the highest product yield and reasonable cell viability. For example, GG5r corresponding to triosephosphate isomerase (*tpiA*) was a target for deletion since deletion of this gene increased fractional remaining pathways for isobutanol synthesis (*P*-frac) to 15%, while still supported cell growth and the highest yield of product (Fig. 4a). Likewise, for hexadecanol, CAL4 corresponding to fructose/sedoheptulose bisphosphatase (*cbbF*) was selected as deletion target with an increased *P*-frac to 25% in the knockout mutant, while retaining the highest product yield and cell viability (Fig. 4d). We applied the screening process in a sequential manner for selecting an optimal combination of multiple gene deletion. The identified target from single gene deletion was used as the genetic background in the next steps for double gene deletion and continued on (Fig. 4b–c, e–f).
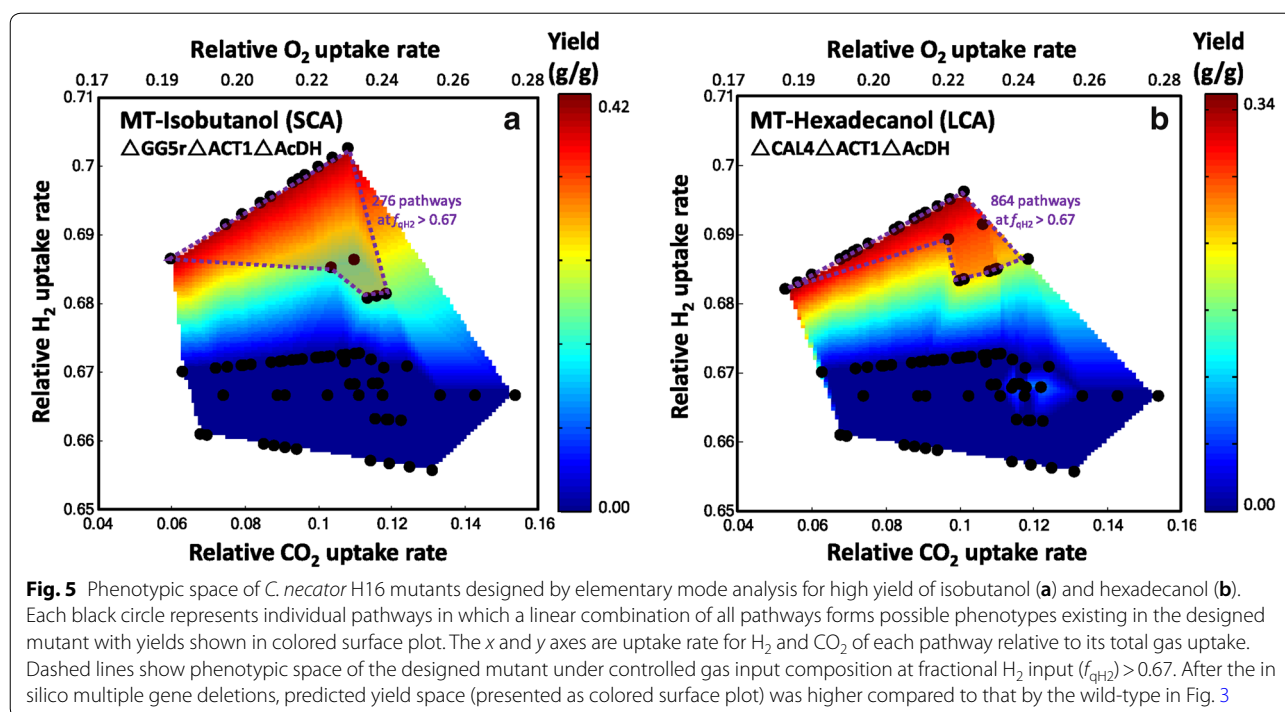
Table 1 summarizes the number of remaining pathways and predicted yields for product and cell growth after sequential multiple gene deletions. The gene deletion simulation resulted in targeted triple gene knockout for maximizing isobutanol and hexadecanol in *C. necator* under autotrophic growth. The multiple gene deletions for efficient production of isobutanol included disrupting GG5r (triosephosphate isomerase, *tpiA*), ACT1 (acetoacetyl-CoA thiolase/acetyl-CoA acetyltransferase, *phaA*) and AcDH (acetaldehyde dehydrogenase, *mhpF/acoD*). For efficient production of hexadecanol, removal of CAL4 (fructose/sedoheptulose bisphosphatase, *cbbF*), ACT1 and AcDH was predicted. The combination of these gene deletions in the designed mutants as summarized in Table 1 reduced the number of existing paths to 1476 with 17% of total paths directed to isobutanol, and 3420 with 28% of total paths directed to hexadecanol. The deletion of GG5r (*tpiA*) was likely to reduce NADH production via glycolysis which then leads to less production of byproducts associated with NADH oxidation such as succinate, lactate and ethanol as reported previously by Tokuyama et al. (2014). Likewise, deletion of ACT1

(phaA) and AcDH (mhpF/acoD) resulted in lowered fermentative products, acetone and ethanol, respectively, which was consistent with the previous reports (Singh et al. 2011; Zhou et al. 2011). Blocking CAL4 (cbbF) was likely to divert CO_2 fixation through glycolysis instead of pentose phosphate pathway, hence increasing availability of the precursors derived from the pathway including pyruvate and acetyl-CoA.

Phenotypic space of the design mutants for isobutanol and hexadecanol in Fig. 5a, b revealed that higher product yield could be achieved by the designed mutants when its fractional H_2 input (fq_{H2}) relative to total gas input was controlled at higher than 0.67. This restriction could be simply implemented by manipulating composition of gas feeding during the autotrophic cultivation of the mutants. Combining in silico multiple gene deletions and careful control of gas input composition at its optimal level, the predicted phenotypic space was redirected, as depicted by dashed lines in Fig. 5, towards efficient pathway options for maximized products at relatively high yield. Under the control of gas input (fq_{H2} > 0.67), the designed mutants were able to efficiently produce products with yield range for isobutanol at 0.21–0.42 g-isobutanol/g-CO_2, and for hexadecanol at 0.20–0.34 g-hexadecanol/g-CO_2, higher values compared to that by the wild-type (Table 1). It should be emphasized that the minimal yields achieved by the designed mutants under controlled gas feeding was non-zero, meaning that multiple gene deletion targets in the mutants pushed the cell's phenotypic space towards product synthesis with all existing pathways directed to products. Hence, their functioning phenotypes were more efficient compared to that of the wild-type which contained many pathway options that did not produce products, a zero minimal yield of product predicted by EMA.

Gene overexpression target debottlenecking for high productivity

Metabolic flux profiles of the two mutants (ΔGG5rΔACT1ΔAcDH for optimized isobutanol production; ΔCAL4ΔACT1ΔAcDH for optimized hexadecanol production) designed by EMA gene deletion simulation were then analyzed through FBA. The analysis aimed at identifying target overexpressed genes involved in enhanced productivity of the desired products. In silico gene overexpression simulations by FBA were performed by increasing each reaction flux at 1–20-fold(s) of the predicted flux values. Then, the predicted phenotype in terms of production rate of product synthesis corresponding to the flux increase was examined. To compare the effect of gene overexpression, flux correlation (FC) for each gene overexpression simulation as defined in Eq. (1) was calculated. The FC of each metabolic flux, that was
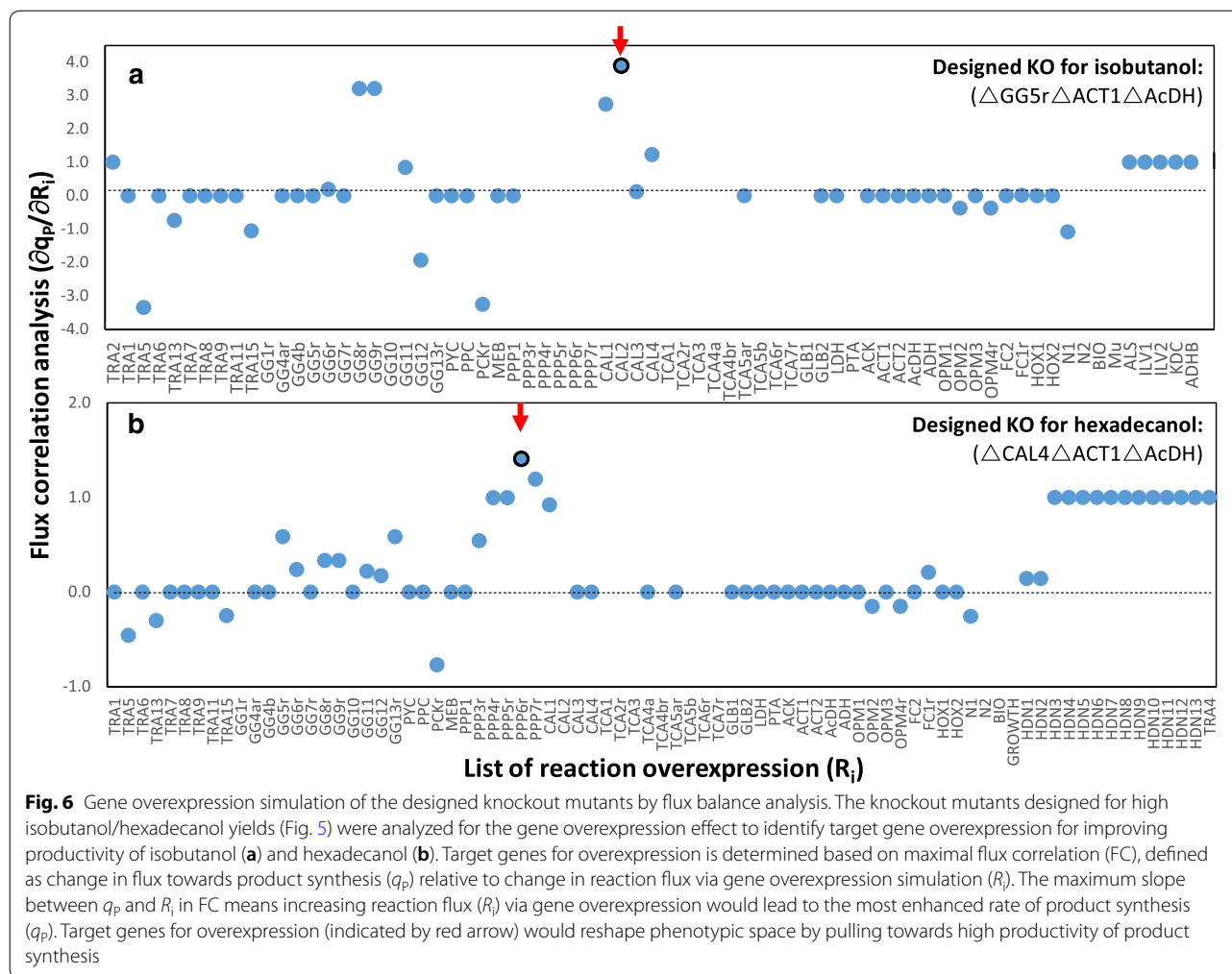
Fig. 5 Phenotypic space of *C. necator* H16 mutants designed by elementary mode analysis for high yield of isobutanol (**a**) and hexadecanol (**b**). Each black circle represents individual pathways in which a linear combination of all pathways forms possible phenotypes existing in the designed mutant with yields shown in colored surface plot. The *x* and *y* axes are uptake rate for H_2 and CO_2 of each pathway relative to its total gas uptake. Dashed lines show phenotypic space of the designed mutant under controlled gas input composition at fractional H_2 input (f_{qH2}) > 0.67. After the in silico multiple gene deletions, predicted yield space (presented as colored surface plot) was higher compared to that by the wild-type in Fig. 3

overexpressed, was computed as change in flux towards product synthesis (q_p) relative to change in reaction flux of gene overexpression (R_i). The rate-controlling flux with maximal FC, meaning the increased reaction flux (R_i) via gene overexpression leading to the most enhanced rate of product synthesis (q_p), was identified as target overexpressed genes. As shown in Fig. 6, FC comparison of all overexpressed reactions within the recombinant network identified CAL2 (phosphoglycerate kinase, *cbbK*) in the Calvin–Benson–Bassham cycle and PPP6r (transaldolase, *talB*) in the pentose phosphate pathway as the rate-limiting fluxes, the enzyme with the greatest flux influence for high productivity of isobutanol and hexadecanol, respectively. These targeted manipulations would reshape phenotypic space of the designed mutants by overcoming bottleneck, hence pulling more fluxes to enhance the products synthesis. With overexpressing CAL2 reaction, FBA predicted metabolic characteristics of the designed mutants capable of producing isobutanol at productivity up to 38.4 mmol/cell-h during autotrophy which was a maximum rate based on stoichiometric balance. The productivity of hexadecanol was predicted at 9.1 mmol/cell-h, a maximum rate based on stoichiometric balance, when PPP6r was overexpressed in the mutant designed for hexadecanol growing autotrophically. These rates were higher when compared to those achieved by

the wild-type which contained many pathways that did not produce products, zero productivity of the products.

Conclusion

Metabolic pathway analysis of the recombinant *C. necator* H16 has provided insight into pathway topology of the organism under autotrophic condition with mixed CO_2, H_2 and O_2 gas feeding for the production of advanced biofuels, isobutanol and hexadecanol, as case studies. Two pathway analysis tools, EMA and FBA, were implemented for assessing the cell's phenotypic capabilities under genetic deletion and overexpression variables, respectively. Analysis of these pathway possibilities provides an insight into the structure of the *C. necator* recombinant network that can guide possible genetic modification and optimization strategies for improving the product synthesis. A number of available pathways identified by pathway analysis tools revealed the complexity and cellular flexibility under genetic perturbations. Ultimately, through in silico design of efficient *C. necator* strains by pushing, via targeted gene deletion, and pulling, via target gene overexpression, more metabolic fluxes towards isobutanol and hexadecanol synthesis have been accomplished. Combination of these genetic manipulations together with controlling gas input composition eliminated

Fig. 6 Gene overexpression simulation of the designed knockout mutants by flux balance analysis. The knockout mutants designed for high isobutanol/hexadecanol yields (Fig. 5) were analyzed for the gene overexpression effect to identify target gene overexpression for improving productivity of isobutanol (**a**) and hexadecanol (**b**). Target genes for overexpression is determined based on maximal flux correlation (FC), defined as change in flux towards product synthesis (q_p) relative to change in reaction flux via gene overexpression simulation (R_i). The maximum slope between q_p and R_i in FC means increasing reaction flux (R_i) via gene overexpression would lead to the most enhanced rate of product synthesis (q_p). Target genes for overexpression (indicated by red arrow) would reshape phenotypic space by pulling towards high productivity of product synthesis

inefficient phenotypic space enforcing the cell to operate efficiently leading to strains with high yield and high productivity of the desired products. The knowledge of metabolic pathway analysis presented here could serve as basis for future systems metabolic engineering *C. necator* H16 to enhance CO_2-to-biofuels conversion efficiency.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s40643-019-0282-4.

Additional file 1. Metabolic network model of recombinant *C. necator* producing isobutanol and hexadecanol under autotrophic condition.

Abbreviations
EMA: elementary mode analysis; FBA: flux balance analysis; SCA: short-chain alcohol; LCA: long-chain alcohol.

Authors' contributions
PU performed the modelling analysis and wrote the manuscript. KLT and TSW provided relevant data for model validation and revised the manuscript. All authors read and approved the final manuscript.

Availability of data and materials
All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate
Not applicable.

Author details
[1] National Center for Genetic Engineering and Biotechnology, Pathum Thani 12120, Thailand. [2] Department of Chemical and Biological Engineering and Advanced Biomanufacturing Centre, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK.

References

Alagesan S, Minton Nigel P, Malys N (2018) 13C-assisted metabolic flux analysis to investigate heterotrophic and mixotrophic metabolism in *Cupriavidus necator* H16. Metabolomics 14(1):9

Blazeck J, Alper H (2010) Systems metabolic engineering: genome-scale models and beyond. Biotechnol J 5(7):647–659

Bruland N, Voß I, Brämer C, Steinbüchel A (2010) Unravelling the C3/C4 carbon metabolism in *Ralstonia eutropha* H16. J Appl Microbiol 109(1):79–90

Chae TU, Choi SY, Kim JW, Ko YS, Lee SY (2017) Recent advances in systems metabolic engineering tools and strategies. Curr Opin Biotechnol 47:67–82

Chen JS, Colón B, Dusel B, Ziesack M, Way JC, Torella JP (2015) Production of fatty acids in *Ralstonia eutropha* H16 by engineering β-oxidation and carbon storage. PeerJ. 3:e1468

Curran KA, Crook NC, Alper HS (2012) Using flux balance analysis to guide microbial metabolic engineering. Methods Mol Biol 834:197–216

Fukui T, Chou K, Harada K, Orita I, Nakayama Y, Bamba T, Nakamura S, Fukusaki E (2014) Metabolite profiles of polyhydroxyalkanoate-producing *Ralstonia eutropha* H16. Metabolomics 10(2):190–202

Guo W, Sheng J, Zhao H, Feng X (2016) Metabolic engineering of *Saccharomyces cerevisiae* to produce 1-hexadecanol from xylose. Microb Cell Fact 15:24

Henson MA, Hanly TJ (2014) Dynamic flux balance analysis for synthetic microbial communities. IET Syst Biol 8(5):214–229

Ishizaki A, Tanaka K (1990) Batch culture of *Alcaligenes eutrophus* ATCC 17697T using recycled gas closed circuit culture system. J Ferment Bioeng 69:170–174

Ishizaki A, Tanaka K (1991) Production of poly-β-hydroxybutyric acid from carbon dioxide by *Alcaligenes eutrophus* ATCC 17697T. J Ferment Bioeng 70:254–257

Kohlmann Y, Pohlmann A, Otto A, Becher D, Cramm R, Lultte S (2011) Analyses of soluble and membrane proteomes of *Ralstonia eutropha* H16 reveal major changes in the protein complement in adaptation to lithoautotrophy. J Proteome Res 10(6):2767–2776

Lee HM, Jeon BY, Oh MK (2016) Microbial production of ethanol from acetate by engineered *Ralstonia eutropha*. Biotechnol Bioprocess Eng 21(3):402–407

Li H, Opgenorth PH, Wernick DG, Rogers S, Wu TY, Higashide W, Malati P, Huo YX, Cho KM, Liao JC (2012) Integrated electromicrobial conversion of CO_2 to higher alcohols. Science 335(6076):1596

Lopar M, Špoljarić IV, Cepanec N, Koller M, Braunegg G, Horvat P (2014) Study of metabolic network of Cupriavidus necator DSM 545 growing on glycerol by applying elementary flux modes and yield space analysis. J Ind Microbiol Biotechnol. 41(6):913–930

Lu J, Brigham CJ, Gai CS, Sinskey AJ (2012) Studies on the production of branched-chain alcohols in engineered *Ralstonia eutropha*. Appl Microbiol Biotechnol 96(1):283–297

Marella ER, Holkenbrink C, Siewers V, Borodina I (2018) Engineering microbial fatty acid metabolism for biofuels and biochemicals. Curr Opin Biotechnol. 50:39–46

Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? Nat Biotechnol 28(3):245–248

Park JM, Kim TY, Lee SY (2011) Genome-scale reconstruction and in silico analysis of the *Ralstonia eutropha* H16 for polyhydroxyalkanoate synthesis, lithoautotrophic growth, and 2-methyl citric acid production. BMC Syst Biol 5:101

Pohlmann A, Fricke WF, Reinecke F, Kusian B, Liesegang H, Cramm R (2006) Genome sequence of the bioplastic-producing knallgas bacterium *Ralstonia eutropha* H16. Nat Biotech. 24(10):1257–1262

PsarrasP Comello S, Bains P, Charoensawadpong P, Reichelstein S, Wilcox J (2017) Carbon capture and utilization in the industrial sector. Environ Sci Technol 51(19):11440–11449

Raberg M, Voigt B, Hecker M, Steinbüchel A (2014) A closer look on the polyhydroxybutyrate— (PHB—) negative phenotype of *Ralstonia eutropha* PHB-4. PLoS ONE 9(5):e95907

Reutz I, Schobert P, Bowien B (1982) Effect of phosphoglycerate mutase deficiency on heterotrophic and autotrophic carbon metabolism of *Alcaligenes eutrophus*. J Bacteriol 151(1):8–14

Schäferjohann J, Yoo JG, Kusian B, Bowien B (1993) The*cbb*operons of the facultative chemoautotroph *Alcaligenes eutrophus* encode phosphoglycolate phosphatase. J Bacteriol 175(22):7329–7340

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

Schwartz E, Henne A, Cramm R, Eitinger T, Friedrich B, Gottschalk G (2003) Complete nucleotide sequence of pHG1: a *Ralstonia eutropha* H16 megaplasmid encoding key enzymes of H_2-based lithoautotrophy and anaerobiosis. J Mol Biol 332(2):369–383

Schwartz E, Voigt B, Zühlke D, Pohlmann A, Lenz O, Albrecht D (2009) A proteomic view of the facultatively chemolithoautotrophic lifestyle of *Ralstonia eutropha* H16. Proteomics 9(22):5132–5142

Singh A, Cher Soh K, Hatzimanikatis V, Gill RT (2011) Manipulating redox and ATP balancing for improved production of succinate in *E. coli*. Metab Eng 13(1):76–81

Tee KL, Grinham J, Othusitse AM, González-Villanueva M, Johnson AO, Wong TS (2017) An Efficient Transformation method for the bioplastic-producing "Knallgas" bacterium *Ralstonia eutropha* H16. Biotechnol J 12:11

Ternon C, Groussean E, Gunther J, Gorret N, Guillouet S, Sinskey J, Aceves-Lara A, Roux G (2014) Dynamic model for isopropanol production by *Cupriavidus necator*. IFAC Proc Vol 47(3):4388–4393

Thakur IS, Kumar M, Varjani SJ, Wu Y, Gnansounou E, Ravindran S (2018) Sequestration and utilization of carbon dioxide by chemical and biological methods for biofuels and biomaterials by chemoautotrophs: opportunities and challenges. Bioresour Technol 256:478–490

Tokuyama K, Ohno S, Yoshikawa K, Hirasawa T, Tanaka S, Furusawa C, Shimizu H (2014) Increased 3-hydroxypropionic acid production from glycerol, by modification of central metabolism in *Escherichia coli*. Microb Cell Fact 13:64

Yu J (2018) Fixation of carbon dioxide by a hydrogen-oxidizing bacterium for value-added products. World J Microbiol Biotechnol 34(7):89

Zhao Q, Yu S, Shi J (2013) Applications of elementary mode analysis in biological network and pathway analysis. Chin J Biotechnol 29(6):701–715

Zhou L, Zuo ZR, Chen XZ, Niu DD, Tian KM, Prior BA, Shen W, Shi GY, Singh S, Wang ZX (2011) Evaluation of genetic manipulation strategies on D-lactate production by *Escherichia coli*. Curr Microbiol 62(3):981–989