Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol

Harry Yim[1,3], Robert Haselbeck[1,3], Wei Niu[1,3], Catherine Pujol-Baxley[1,3], Anthony Burgard[1,3], Jeff Boldt[1], Julia Khandurina[1], John D Trawick[1], Robin E Osterhout[1], Rosary Stephen[1], Jazell Estadilla[1], Sy Teisan[1], H Brett Schreyer[1], Stefan Andrae[1], Tae Hoon Yang[1], Sang Yup Lee[2], Mark J Burk[1] & Stephen Van Dien[1]*

1,4-Butanediol (BDO) is an important commodity chemical used to manufacture over 2.5 million tons annually of valuable polymers, and it is currently produced exclusively through feedstocks derived from oil and natural gas. Herein we report what are to our knowledge the first direct biocatalytic routes to BDO from renewable carbohydrate feedstocks, leading to a strain of *Escherichia coli* capable of producing 18 g l⁻¹ of this highly reduced, non-natural chemical. A pathway-identification algorithm elucidated multiple pathways for the biosynthesis of BDO from common metabolic intermediates. Guided by a genome-scale metabolic model, we engineered the *E. coli* host to enhance anaerobic operation of the oxidative tricarboxylic acid cycle, thereby generating reducing power to drive the BDO pathway. The organism produced BDO from glucose, xylose, sucrose and biomass-derived mixed sugar streams. This work demonstrates a systems-based metabolic engineering approach to strain design and development that can enable new bioprocesses for commodity chemicals that are not naturally produced by living cells.

Oil and natural gas are used as the primary raw materials for manufacturing an array of large-volume chemicals, polymers and other products that improve our overall standard of living. Growing concerns over the environment and volatile fossil-energy costs have inspired a quest to develop more sustainable processes that afford these same products from renewable feedstocks with lower cost, lower energy consumption, and reduced waste and greenhouse gas emissions. Metabolic engineering of microorganisms is emerging as a powerful approach to address this need, and it entails the creation of new high-performance cellular systems that convert inexpensive plant-derived carbohydrates into bio-based fuels, chemicals and polymers[1]. Here we report the direct production of BDO, a major commodity chemical used to make over 2.5 million tonnes of plastics, polyesters and spandex fibers annually.

BDO currently is manufactured entirely from petroleum-based feedstocks such as acetylene, butane, propylene and butadiene[2]. Given the importance of BDO as a chemical intermediate and the issues associated with petroleum feedstocks, alternative low-cost renewable routes from sugars have been highly sought after. However, the highly reduced nature of BDO relative to carbohydrates has thwarted attempts thus far to develop effective pathways and organisms for direct production. Furthermore, like many commodity chemicals of interest, BDO is not a compound produced naturally in any known organism. The need for an efficient, sustainable process led us to a detailed assessment of biosynthetic pathways and host-strain designs for direct one-step production of BDO in a microbial fermentation process.

Engineering a microbe for the production of a heterologous compound requires establishment of a new biochemical pathway, in addition to a thorough knowledge of metabolic pathways and metabolism. A textbook example of a successful metabolic engineering project is the production of 1,3-propanediol (PDO) in *Escherichia coli* developed by Genencor and DuPont[3], which led to a commercial

process. By introducing a four-step pathway consisting of genes from PDO-synthesizing bacterial species, together with targeted changes to the host central metabolism, researchers at these companies were able to achieve PDO production with high rate and titer. Microbial processes have also been reported for a number of key chemical intermediates such as 1,2-propanediol[4], isobutanol[5], isoprene[6] and putrescine[7], in addition to proposed platform chemicals such as succinic acid[8], glucaric acid[9] and 3-hydroxypropionate[10]. Direct biological production of BDO introduces the additional challenges that this compound is highly reduced and is not produced naturally in any known organism. To facilitate and expedite our effort, we leveraged predictive computational modeling of metabolism and model-driven analysis of experimental data. Constraint-based metabolic modeling is a rapidly developing discipline, and researchers have reported the successful application of metabolic models of *E. coli* to engineer strains that produce high levels of threonine[11], valine[12] and succinic acid[13].

Here we report the use of an accurate genome-scale metabolic model of *E. coli*[14] and biopathway prediction algorithms[15] to broadly survey and prioritize specific BDO pathways that are predicted to lead to optimal performance. In addition, the model engendered metabolic engineering strategies for balancing energy and redox needs, and for elimination of potentially toxic by-products. Translating our computational design into the laboratory, we introduced and optimized two heterologous pathways for BDO production in *E. coli* and engineered the host metabolism to direct carbon and energy into the pathways. Finally, our production strains were capable of synthesizing BDO from a range of renewable feedstocks.

RESULTS

BDO is a non-natural compound not synthesized by any known organism, so there are no complete biosynthetic pathways we could harness for BDO production. We therefore used our in-house

[1]Genomatica, Inc., San Diego, California, USA. [2]Department of Chemical and Biomolecular Engineering (BK21 program), Center for Systems and Synthetic Biotechnology, Institute for the BioCentury, Korea Advanced Institute of Science and Technology, Daejeon, South Korea. [3]These authors contributed equally to this work. *e-mail: svandien@genomatica.com
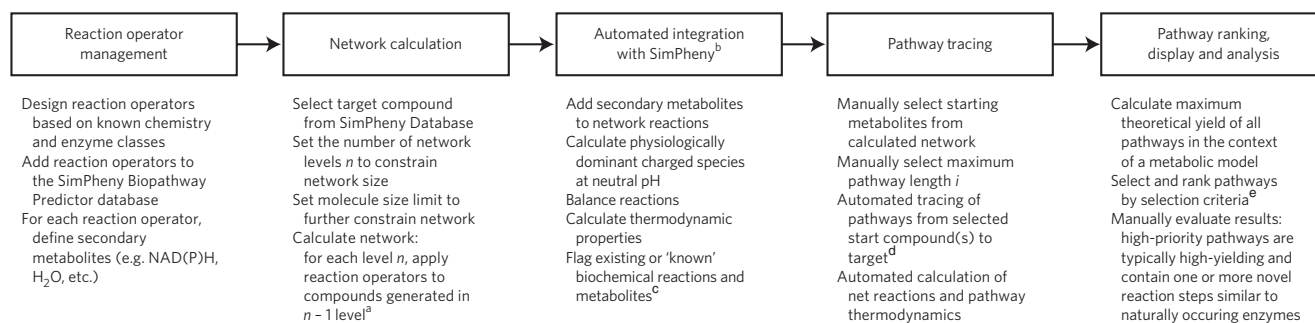
Reaction operator management	Network calculation	Automated integration with SimPheny[b]	Pathway tracing	Pathway ranking, display and analysis
Design reaction operators based on known chemistry and enzyme classes Add reaction operators to the SimPheny Biopathway Predictor database For each reaction operator, define secondary metabolites (e.g. NAD(P)H, H_2O, etc.)	Select target compound from SimPheny Database Set the number of network levels n to constrain network size Set molecule size limit to further constrain network Calculate network: for each level n, apply reaction operators to compounds generated in $n-1$ level[a]	Add secondary metabolites to network reactions Calculate physiologically dominant charged species at neutral pH Balance reactions Calculate thermodynamic properties Flag existing or 'known' biochemical reactions and metabolites[c]	Manually select starting metabolites from calculated network Manually select maximum pathway length i Automated tracing of pathways from selected start compound(s) to target[d] Automated calculation of net reactions and pathway thermodynamics	Calculate maximum theoretical yield of all pathways in the context of a metabolic model Select and rank pathways by selection criteria[e] Manually evaluate results: high-priority pathways are typically high-yielding and contain one or more novel reaction steps similar to naturally occuring enzymes

Figure 1 | Overview of the Biopathway Predictor network calculation and analysis procedure. [a]Procedure for calculating new compounds at level n based on reaction operators and compounds present at level $n-1$ is given in ref. 16. [b]The SimPheny database contains manually curated metabolic models, enzymes, reactions and metabolites. Calculated Biopathway Predictor networks, composed of simple substrate-product reactions, must be further processed into balanced chemical reactions composed of physiologically relevant species. [c]Reaction and metabolite flags, used for pathway ranking and display, allow the scientist to quickly distinguish between known and novel metabolites and enzymatic reactions. [d]Traced pathways vary in length from 1 to i reaction steps. [e]Selection criteria include number of pathway steps, number of known metabolites or enzymes, ΔG_{rxn} of pathway, contains or excludes reaction operator or metabolite of interest, and maximum theoretical yield of product (according to constraint-based flux analysis).

SimPheny Biopathway Predictor software to elucidate all potential pathways from *E. coli* central metabolites to BDO. The Biopathway Predictor algorithm is based on transformation of functional groups by known chemistry, termed 'reaction operators', rather than by known enzyme reactions. Therefore, we are not restricted by what enzymes are known, leaving the flexibility to identify novel enzyme activities or to engineer enzymes for a particular substrate. The concept is similar to others previously reported[15–17] (see Methods). The algorithm identified over 10,000 pathways of four to six steps for the synthesis of BDO from common central metabolites such as acetyl-CoA, α-ketoglutarate, succinyl-CoA and glutamate (**Supplementary Results**, **Supplementary Fig. 1**). We next used in-house pathway-visualization and selection software to sort and rank the pathways. The evaluation process involved the iterative ranking of pathways on the basis of various attributes including maximum theoretical BDO yield, pathway length, number of non-native steps, number of novel steps and thermodynamic feasibility (**Fig. 1**). We first eliminated pathways with unfavorable thermodynamics (based on group-contribution theory[18]) or reduced theoretical yield when evaluated by constraint-based modeling. We prioritized the remaining pathways (approximately 10% of the initial number) using the following criteria (in order of weighting): the number of steps without currently

characterized enzymes (based on the Kyoto Encyclopedia of Genes and Genomes, EcoCyc and internal databases), the number of non-native steps required, and the total number of steps from central metabolism. Overall, this process selected the BDO production pathways proceeding through the 4-hydroxybutyrate intermediate (**Scheme 1** and **Supplementary Fig. 1**) as the highest priority for construction and *in vivo* testing. Several additional pathways, as well as candidate enzymes for catalyzing each pathway step, are disclosed[19] (**Supplementary Figs. 2–4**).

Two artificial routes for BDO biosynthesis (**Scheme 1**) converge at the common intermediate 4-hydroxybutyrate (4HB). For the purpose of development and validation, we divided the pathway into upstream enzymes for the production of 4HB and downstream enzymes for the conversion of 4HB to BDO.

Upstream pathway: biosynthesis of 4HB from glucose

The first route to 4HB starts from the tricarboxylic acid (TCA)–cycle intermediate succinate, which is activated as succinyl-CoA by the native *E. coli* enzyme succinyl-CoA synthetase (SucCD). After two sequential reduction steps catalyzed by CoA-dependent succinate semialdehyde dehydrogenase (SucD) and 4HB dehydrogenase (4HBd), respectively, the CoA derivative converts to 4HB
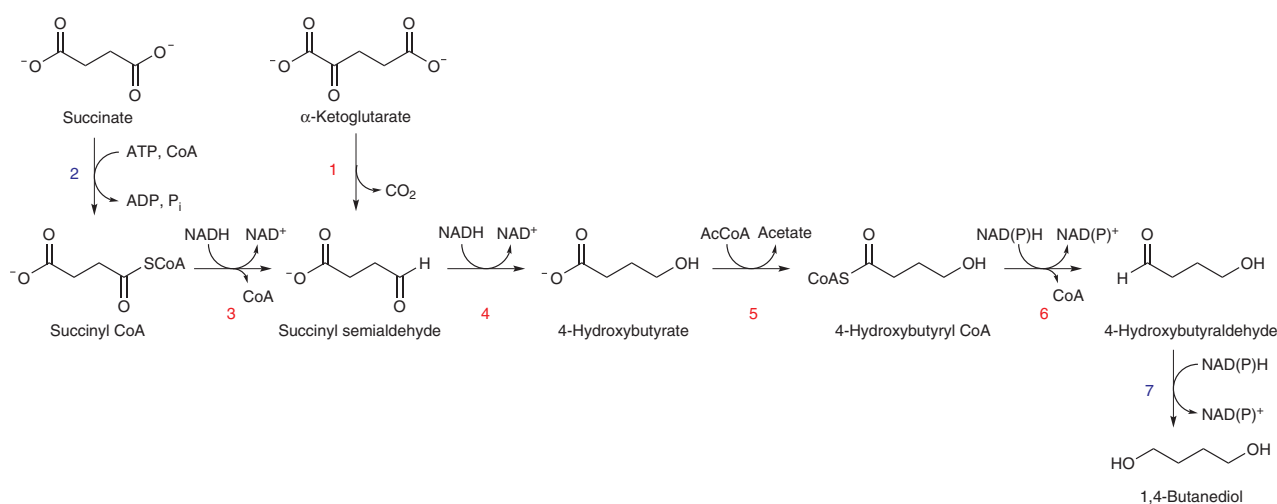
Scheme 1 | BDO biosynthetic pathways introduced into *E. coli*. Enzymes for each numbered step are as follows: (1) 2-oxoglutarate decarboxylase; (2) succinyl-CoA synthetase; (3) CoA-dependent succinate semialdehyde dehydrogenase; (4) 4-hydroxybutyrate dehydrogenase; (5) 4-hydroxybutyryl-CoA transferase; (6) 4-hydroxybutyryl-CoA reductase; (7) alcohol dehydrogenase. Steps 2 and 7 occur naturally in *E. coli*, whereas the others are encoded by heterologous genes introduced in this work.

via succinate semialdehyde. SucD activity is not found in wild-type *E. coli* K-12 strains, but it has been reported in other bacterial species. *E. coli* already possesses a 4HBd activity[20], so we first tested whether introduction of heterologous SucD alone would result in 4HB production. *E. coli* W3110 was transformed with p99sucD, which harbors the *Clostridium kluyveri sucD* gene expressed from the *trc* promoter. After 24 h cultivation in glucose minimal medium, 0.13 mM of 4HB appeared in the culture broth. A control culture without p99sucD did not produce any detectable 4HB. To improve upon this rate of production, we screened SucD and 4HBd enzymes from *C. kluyveri* and *Porphyromonas gingivalis*[21–23] and an additional 4HBd enzyme from *Ralstonia eutropha* for biochemical properties and the ability to be expressed in *E. coli* (**Supplementary Table 1**). Both SucD and 4HBd from *P. gingivalis* were expressed well and resulted in the highest specific activities in host strain MG1655*lacI*[Q]. We then assembled a synthetic operon by sequentially cloning *sucCD* (*E. coli*), *sucD* and *4hbd* (*P. gingivalis*) together with individual ribosomal binding sites between the PA1 promoter, which is regulated by the lactose repressor protein, and the rrnB T1 transcriptional terminator. We tested this operon in three expression vectors with copy numbers ranging from low to high (**Table 1**), transformed into MG1655*lacI*[Q]. All three *E. coli* constructs enabled synthesis of 4HB from glucose, whereas the host strain transformed with empty plasmid backbones only accumulated succinate (**Supplementary Fig. 5**). The highest 4HB concentration (11 mM), highest ratio of 4HB to succinate and highest enzyme activities occurred with the strain containing the medium–copy number plasmids (**Supplementary Fig. 5** and **Supplementary Table 2**).

The second route for 4HB synthesis branches from *E. coli* central metabolism at the key oxidative TCA-cycle intermediate, α-ketoglutarate. The pathway consists of two consecutive reactions catalyzed by an α-keto acid decarboxylase, encoded by the *Mycobacterium bovis sucA* gene[24], and the 4HB dehydrogenase described above. This pathway is thermodynamically more favorable than the succinate route, owing to the irreversible decarboxylation step, and it also consumes one less reducing equivalent. To demonstrate the synthesis of 4HB using α-ketoglutarate as the intermediate, we examined three *E. coli* constructs (**Supplementary Fig. 6**). Similar concentrations of succinate accumulated in all constructs including the control strain, whereas only the three constructs with SucA overexpression produced 4HB. There was also a positive correlation between 4HB accumulation and gene copy number.

Downstream pathway: conversion of 4HB to BDO in *E. coli*

The conversion of 4HB to BDO requires two reduction steps, catalyzed by dehydrogenases (**Scheme 1**). Alcohol and aldehyde dehydrogenases (ADH and ALD, respectively) are NADH- and/or NADPH-dependent enzymes that together can reduce a carboxylic acid group (derivatized with Coenzyme A) to an alcohol group. This biotransformation can occur by addition of exogenous 4HB to wild-type *Clostridium acetobutylicum*[25], but neither the enzymes nor the genes responsible have been identified. We developed a list of candidate enzymes from *C. acetobutylicum* and related organisms on the basis of known activity with the nonhydroxylated analogs of 4HB (for example, butyrate) and other pathway intermediates, or by sequence similarity to characterized genes with these activities (**Supplementary Table 3**). We tested candidates for activity by expressing the corresponding genes on the high-copy pZA33S plasmid in the *E. coli* host MG1655*lacI*[Q]. As 4HB-CoA is not available commercially, the nonhydroxylated analog butyryl-CoA served as the substrate for aldehyde dehydrogenase assays. Activity was determined by the oxidation of NADH as measured by the change in absorbance at 340 nm (A_{340}; **Supplementary Methods**). The ratio of activity with a four-carbon (butyryl-CoA) substrate over activity with a two-carbon (acetyl-CoA) substrate was 1.65 or 0.73 for the enzymes encoded

Table 1 | Strains and plasmids

Designation	Genotype or description	References
***E. coli* strains**		
W3110	Wild-type *E. coli* K-12	CGSC 4474
MG1655 *lacI*[Q]	Wild-type *E. coli* K-12	ATCC 47076
MG1655 Δ*ldhA lacI*[Q]	MG1655 Δ*ldhA lacI*[Q]	This work
AB3	MG1655 *lacI*[Q] Δ*adhE* Δ*ldhA* Δ*pflB*	This work
ECKh-138	AB3 Δ*lpdA*::*K.p.lpdD354K*[a]	This work
ECKh-401	ECKh-138 Δ*mdh* Δ*arcA*	This work
ECKh-422	ECKh-401 *gltAR163L*, Str[R]	This work
ECKh-463	ECKh-422 *rrnC*::*cscAKB*	This work
Plasmids[b]		
pZS*13S	pSC101 origin, Amp[R], P[A1lacO-1]	ref. 40
pZA33S	P15A origin, Chlor[R], P[A1lacO-1]	ref. 40
pZE13S	ColE1 origin, Amp[R], P[A1lacO-1]	ref. 40
pZE23S	ColE1 origin, Kan[R], P[A1lacO-1]	ref. 40
pTrc99A	ColE1 origin, Amp[R], P[trc]	Vector Database

[a]Native *lpdA* gene replaced by *K. pneumoniae lpdA* modified with a D354K mutation. [b]Plasmid backbones are listed here. Plasmids expressing genes in an operon are described using the convention 'backbone-gene1-gene2-gene3' in the main text.

by *C. acetobutylicum adhE2* (gene 002) and *E. coli adhE* (gene 011), respectively (**Supplementary Table 3**). These results are similar to those previously reported[5,26]. However, only dehydrogenase 002 was able to synthesize BDO *in vivo*. We also synthesized a codon-optimized version of 002 for use in *E. coli* to improve expression[27] (**Supplementary Fig. 7**), which was used in many of the following experiments. This gene is referred to as 002C.

For the BDO production experiments, we expressed the 4-hydroxybutyryl-CoA transferase (*cat2*) gene from *P. gingivalis* W83 (gene 0034)[28] on the medium-copy plasmid pZA33S for the conversion of 4HB to 4HB-CoA, and the candidate dehydrogenase genes on the high-copy plasmid pZE13S. We then cultured the cells aerobically to an A_{600} of about 0.5 in medium supplemented with 10 mM 4HB, induced expression with 0.25 mM IPTG and took culture-broth samples after 24 h. This resulted in 138 μM BDO in the culture broth, as determined by gas chromatography–mass spectrometry. In cultures of a control strain containing the empty vector pZA33S, no detectable BDO resulted from 10 mM 4HB.

Complete synthesis of BDO from glucose

The final step of pathway validation was to express both the 4HB and BDO segments of the pathway together in *E. coli* and demonstrate production of BDO in glucose minimal medium. We reassembled the complete BDO pathway on two plasmids, pZA33S-sucCD-sucD-002 and pZE13S-cat2-4hbd, and transformed them first into the MG1655*lacI*[Q] background. Cells grown anaerobically in M9 minimal medium and supplemented with 0.25 mM IPTG approximately 15 h after inoculation produced 2.1 mM 4HB and 0.6 mM BDO from 20 g l[−1] glucose in 48 h, along with 6.2 mM succinate. Improved production of both 4HB and BDO resulted from rearrangement of the pathway to express all 4HB synthesis genes from the medium-copy plasmid pZA33S, and the downstream genes *cat2* and 002C from the high-copy plasmid pZE23S. We also included the *M. bovis sucA* gene to enable use of the α-ketoglutarate route discussed above. This resulted in the production of 1.3 mM BDO and 3.4 mM 4HB in 40 h (**Fig. 2a**, left set of bars). We therefore used these plasmids, pZA33S-sucCD-sucD-4hbd/sucA and pZE23S-cat2-002C, in all the host-development work discussed below.
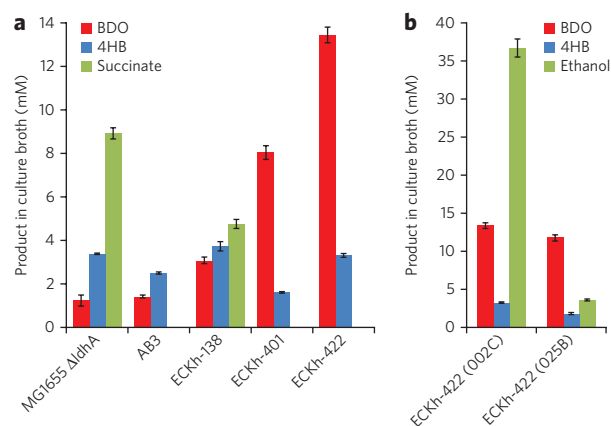
Figure 2 | Production of BDO from glucose in engineered *E. coli* strains. (**a**,**b**) Host strains are as described in **Table 1**. All strains contain the complete BDO pathway on plasmids as follows: *E. coli* sucCD, *P. gingivalis* sucD, *P. gingivalis* 4hbd and *M. bovis* sucA were expressed on the medium-copy plasmid pZA33S (pZA33S-sucCD-sucD-4hbd/sucA); *P. gingivalis* Cat2 and *C. acetobutylicum* AdhE2 (**a**) or *C. beijerinckii* Ald (**b**) were expressed on the high-copy plasmid pZE23S (pZE23S-cat2-002). Cells were grown microaerobically in M9 minimal medium containing 10 g l^{-1} glucose, induced with 0.25 mM IPTG at an A_{600} of approximately 0.5 and harvested 40 h after induction. Bars represent extracellular concentrations of compounds indicated in key. Data are averages from three replicate cultures; error bars show s.d.

Metabolic engineering of the host strain

Achieving a functional BDO pathway in *E. coli* was the first step in engineering an organism for high-level BDO production. The second major step involved optimizing the host strain to channel carbon and energy resources into the pathway. The OptKnock algorithm[29] identified strain designs that couple BDO production to growth while maintaining the overall maximum theoretical BDO yield. Evaluation of the OptKnock results revealed one highly promising four-knockout strategy, removing alcohol dehydrogenase (*adhE*), pyruvate formate lyase (*pfl*), lactate dehydrogenase (*ldh*), and malate dehydrogenase (*mdh*) genes. The model predicted that this design could achieve a BDO yield of 0.37 g per g at maximum growth assuming anaerobic conditions (**Supplementary Fig. 8**). In this design, formation of the natural fermentation products ethanol, formate, lactate and succinate is blocked, thereby forcing production of BDO to balance redox. A deletion analysis of the *in silico* mutant strain revealed that pyruvate dehydrogenase, citrate synthase and aconitase are essential to achieving high yields of BDO in this four-deletion background. Absence of these individual activities reduced the predicted maximum BDO yields to 46%, 33% and 33%, respectively.

We first introduced the pathway into an intermediate strain, Δ*adhE* Δ*pflB* Δ*ldhA* (referred to as AB3). AB3 did not grow under strict anaerobic conditions, probably owing to a combination of redox imbalance and insufficient pyruvate dehydrogenase activity. Pyruvate dehydrogenase in *E. coli* naturally has very low activity under oxygen-limited or anaerobic conditions owing to NADH sensitivity of the E3 subunit encoded by *lpdA*[30,31], leading to poor anaerobic growth of strains with pyruvate formate lyase (*pflB*) deletions. We therefore replaced the native *lpdA* gene with the *Klebsiella pneumonia lpdA* gene, which is greater than 90% identical at the nucleotide level to the *E. coli* gene and is known to function anaerobically[32]. The modified gene also contained the D354K mutation, reported to reduce NADH sensitivity[31]. This strain, called ECKh-138, still did not grow under strict anaerobic conditions, but did

grow much better microaerobically (see Methods for a description of the microaerobic culturing procedure). After transformation with plasmids encoding the entire BDO pathway, this strain produced more BDO in culture than both MG1655 and AB3 (**Fig. 2a**). A fed-batch fermentation of this strain resulted in 82 mM (7.5 g l^{-1}) BDO in 48 h (**Supplementary Fig. 9**). However, high concentrations of acetate, pyruvate and ethanol, and moderate concentrations of pathway intermediates 4HB and γ-butyrolactone (GBL) accumulated as well. GBL is the lactonized form of 4HB, hypothesized to be produced spontaneously from 4HB-CoA. Reduction of these byproducts is discussed below.

In accordance with the OptKnock design, we deleted *mdh* to block reductive TCA cycle flux and channel all carbon through the oxidative TCA cycle. To enable greater expression of the endogenous genes encoding citrate synthase, aconitase and isocitrate dehydrogenase activities, we also deleted the transcriptional repressor of several aerobically expressed genes, *arcA* (ref. 33) (strain ECKh-401). Additionally, we introduced a R163L mutation into the gene encoding citrate synthase, *gltA*, to reduce inhibition by NADH and therefore improve TCA-cycle flux[34,35], generating the strain ECKh-422. This mutation involves residues believed to take part in a complex hydrogen-bonding network with the pyrophosphate moiety of NADH. *In vitro* citrate synthase assays performed on extracts from this strain showed little inhibition by 2 mM NADH, compared to almost complete inhibition in the parent strain (**Supplementary Fig. 10**). When transformed with plasmids encoding the full BDO synthesis pathway, including α-ketoglutarate decarboxylase (*sucA*), ECKh-422 produced substantially more BDO and 4HB than ECKh-401, and both were improved relative to ECKh-138 (**Fig. 2a,b**). Removal of the reductive TCA cycle also eliminated succinate production. Furthermore, ^{13}C labeling analysis confirmed that over 95% of the carbon flux fed into the BDO pathway was via the oxidative TCA cycle and α-ketoglutarate decarboxylase (**Fig. 3**).

Further optimization of BDO pathway

With the host strain constructed according to the computational design, we then turned our attention to improving overall throughput to BDO. On the basis of several lines of evidence, we hypothesized that the functions of the downstream aldehyde and alcohol dehydrogenases were limiting BDO production. First, the alcohol and aldehyde dehydrogenase step is the only one catalyzed by enzymes not using their native substrates. Second, increased expression of the 002 gene owing to codon optimization resulted in higher BDO production. Third, pulse-labeling of a culture with ^{13}C-glucose resulted in a substantial lag between the appearance of label in 4HB and in BDO (**Supplementary Fig. 11**). We therefore set out to improve this downstream activity by rescreening our bank of aldehyde and alcohol dehydrogenases (**Supplementary Table 3**) for those with better BDO production *in vivo* and by improving expression through codon optimization[27]. We found that, like 002C, a codon-optimized version of the *C. beijerinckii* aldehyde dehydrogenase gene[36], 025B, was very active in the conversion of 4HB-CoA to BDO *in vivo*, but unlike 002C, it did not produce substantial amounts of ethanol (presumably owing to lower activity of the enzyme encoded by 025B on acetyl-CoA; **Supplementary Fig. 12**). When expressed in the strain ECKh-422, 025B generated much less ethanol than 002C while slightly increasing the conversion of 4HB to BDO (**Fig. 2b**). 025B encodes a monofunctional enzyme, containing only aldehyde dehydrogenase activity, as opposed to the bifunctional aldehyde and alcohol dehydrogenase activity of enzyme 002 (ref. 36). The alcohol dehydrogenase activity(s) responsible for converting 4-hydroxybutyraldehyde to BDO (**Scheme 1**) in this strain was thus native to *E. coli*.
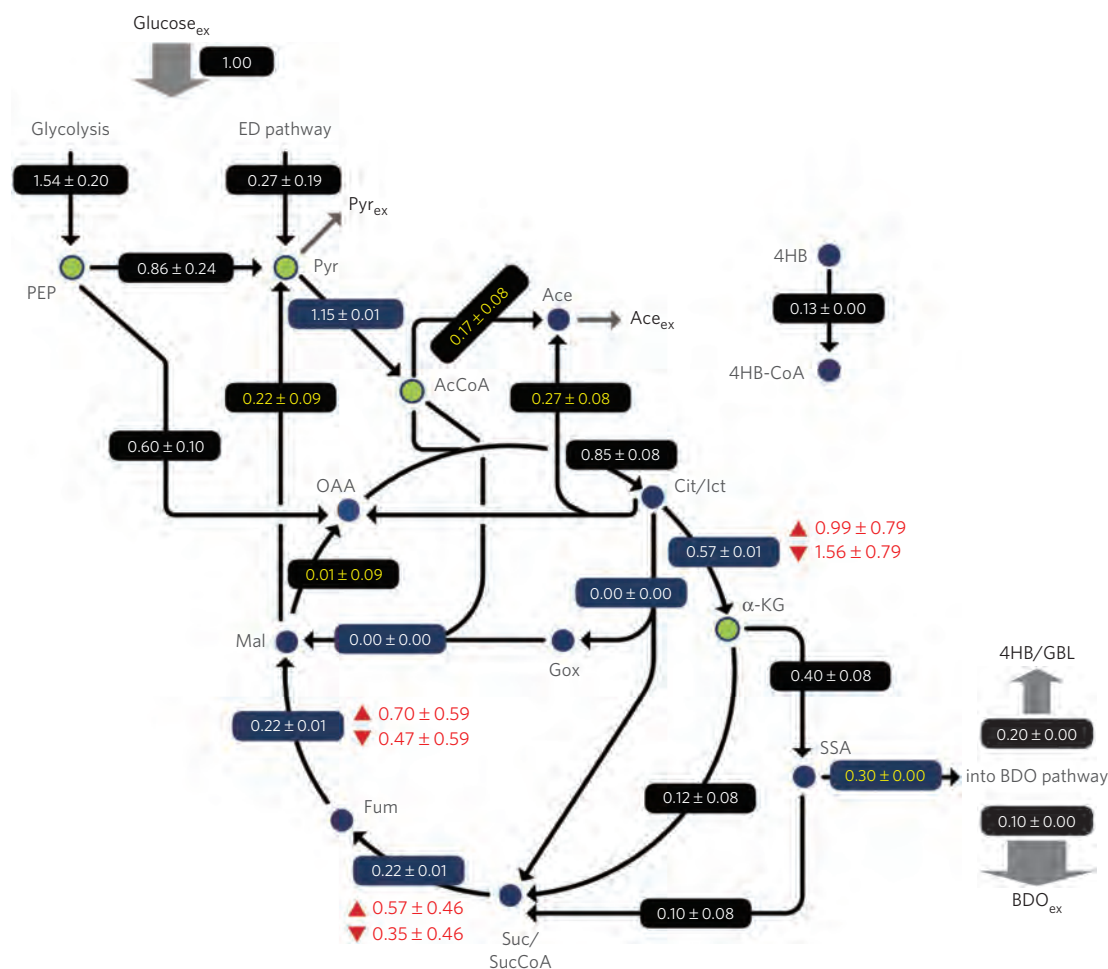
Figure 3 | Flux distribution of ECKh-422 pZS*13S-sucCD-sucD-4hbd/sucA pZE23-OO2C-Cat2 determined by ¹³C metabolic flux analysis. Cells were cultured in microaerobic bottles using 2 g l⁻¹ glucose of various label distributions and analyzed as described in the **Supplementary Methods**. Net fluxes in the direction of arrows are labeled; confidence intervals represent s.d. Metabolites are shown together (separated by a slash) when isotopomer data cannot distinguish between them (for example, succinate and succinyl-CoA). Numbers in red next to a reaction indicate relative forward (top) and reverse (bottom) rates for reversible reactions. PEP, phosphoenolpyruvate; Pyr, pyruvate; Ace, acetic acid; AcCoA, acetyl CoA; OAA, oxaloacetic acid; Cit, citric acid; Ict, isocitric acid; Mal, malic acid; Gox, glyoxalate; α-KG, α-ketoglutarate; Fum, fumaric acid; Suc, succinic acid; SucCoA, succinyl-CoA; SSA, succinic semialdehyde.

We were not surprised by this, considering the large number of *E. coli* genes annotated as alcohol dehydrogenases[33,37].

Production of BDO from different carbohydrate feedstocks

Sucrose represents a potential feedstock for production of chemicals and is a globally traded commodity that in some parts of the world is cheaper per pound than glucose of the same purity. *E. coli* K-12 strains do not naturally take up and utilize sucrose, though sucrose operons of related organisms such as *E. coli* W[38] and *Klebsiella pneumoniae*[39] have been characterized. On the basis of model predictions for maximum theoretical yield, we chose the sucrose operon of *E. coli* W, which does not use the phosphotransferase system to import the sucrose molecule. The operon consists of a sucrose permease (*cscB*), D-fructokinase (*cscK*), sucrose hydrolase (*cscA*) and a LacI-related sucrose-specific repressor (*cscR*). We cloned the sucrose-utilization operon into the *rrnC* locus of ECKh-422 to generate ECKh-436. The synthetic operon also lacked the section of *cscR* encoding the first 53 amino acid residues, thus rendering expression constitutive. ECKh-436 grew as well on sucrose as on glucose and produced concentrations of BDO similar to those of ECKh-422 on glucose (**Fig. 4a**).

Our exploration of feedstock flexibility extended beyond purified substrates to the use of crude, mixed sugar streams derived from hydrolyzed biomass. To this end, we tested the final BDO production strain ECKh-422 pZS*13-sucCD-sucD-4hbd/sucA pZE23-025B-Cat2 on xylose, a typical 5-carbon sugar, and on biomass hydrolysate containing glucose and xylose as well as various oligosaccharides and other impurities (**Supplementary Table 4**). The results indicated approximately equal efficiency of BDO production from all these substrates (**Fig. 4b**). Although the absolute BDO concentrations from xylose and mixed sugars were slightly lower than from glucose, the cell densities of these cultures were also lower, resulting in reduced concentrations of all products and approximately equal specific production rates.

Fermentation of BDO production strains

We further studied one of the best BDO-producing strains, ECKh-422 pZS*13-sucCD-sucD-4hbd/sucA pZE23-025B-Cat2, in a 2-l microaerobic fed-batch fermentation (**Fig. 5a**). The culture produced over 18 g l⁻¹ BDO from glucose in 5 d. To measure the improvement resulting from strain engineering, we compared these results to those from the ancestral strain ECKh-138, which produced 7.5 g l⁻¹ BDO strictly via succinate (**Supplementary Fig. 9**). Acetate, pyruvate and 4HB dominated the ECKh-138 fermentation, whereas BDO was the dominant product with ECKh-422 (**Fig. 5b**).
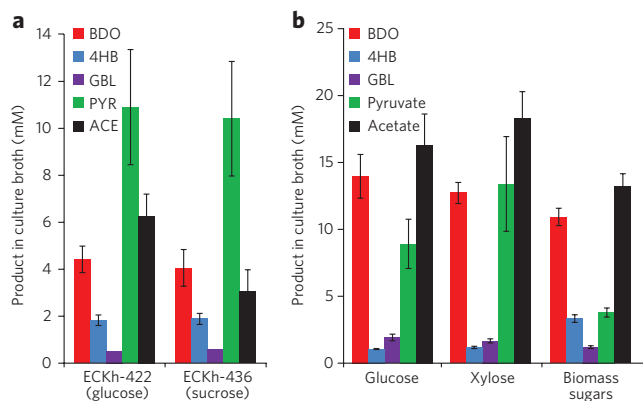
Figure 4 | Production of BDO from various carbohydrate sources.
(**a,b**) Cultures were grown microaerobically in M9 minimal medium containing 20 g l⁻¹ glucose (**a**, ECKh-422), 20 g l⁻¹ sucrose (**a**, ECKh-436), 10 g l⁻¹ glucose (**b**), 10 g l⁻¹ xylose (**b**) or mixed sugars from biomass hydrolysate (**b**), induced with 0.25 mM IPTG and harvested after 24 h. Composition of the biomass hydrolysate is listed in **Supplementary Table 4**. The host strain was ECKh-422 for all substrates except sucrose, in which case it was ECKh-436. Both strains contain the plasmids pZS*13S-sucCD-sucD-4hbd/sucA and pZE23S-025B-34. Bars represent, from left to right, BDO, 4HB, GBL, pyruvate and acetate. Data represent averages from three replicate cultures; error bars show s.d.

DISCUSSION

In this work we present a microbial system for the production of a large-volume commodity chemical, 1,4-butanediol, from sugar feedstocks, representing the biological synthesis of this non-naturally occurring product. This endeavor met with two major challenges: the construction and expression of a heterologous pathway to BDO from central metabolic intermediates, and the engineering of the host *E. coli* metabolism to channel carbon flux and energy into this NADH-intensive pathway. The challenges met here are likely to be characteristic of the production of any commodity chemical biologically, as these compounds tend to be highly reduced, not produced biologically in nature and toxic to the cell. Our integrated computational-experimental approach to this problem can therefore enable biocatalytic production of more challenging yet commercially relevant targets.

The BDO pathway we selected uses a combination of native *E. coli* enzymes, heterologous enzymes working on their native substrates, and heterologous enzymes made to act on non-natural substrates. We searched the literature for candidate enzymes for each of those steps and then used bioinformatics to identify, in databases of sequenced organisms, homologs that in some cases functioned better. We cloned these genes in synthetic operons under control of the *lac*-based P_{A01} promoter[40]. This promoter has the advantage that it is IPTG-inducible in strains containing *lacI*^Q, such as those used in this study, and constitutive in strains lacking this element. Therefore, one could control expression during the research stage of a project, then use the same constructs in the production strain (lacking *lacI*), where adding an inducer is impractical. Expression of six genes under a strong promoter in multiple copy results in a substantial metabolic burden on the cell. Rearrangement of gene order on two plasmids of different copy number optimized the expression balance so that the cell allocated limited cellular resources primarily to those steps that presented a bottleneck for BDO production. Codon optimization further improved the expression of these crucial genes.

E. coli synthesis of 1,4-butanediol entails the reduction of highly oxidized starting material, carbohydrates, into a more reduced target molecule. We used constraints-based modeling and the OptKnock framework to design a strain in which production of BDO was the only way to balance redox and enable anaerobic growth. The first step was deleting pathways to the common fermentation products lactate, formate and ethanol. A less intuitive manipulation was to engineer the oxidative branch of the TCA cycle to be fully functional under the anaerobic or microaerobic cultivation conditions. This entailed a combination of genetic manipulations, including the deletion of the *arcA* gene in the ArcAB two-component system to alleviate the transcriptional repression of genes that encode enzymes functioning in the oxidative branch of the TCA cycle; the overexpression of an NADH inhibition–insensitive version of citrate synthase, which directs the carbon flux into the TCA cycle in the presence of high intracellular NADH concentration; and the deletion of malate dehydrogenase, encoded by *mdh*, to reduce the flux through the reductive TCA cycle under oxygen-limited cultivation conditions. Anaerobic operation of the TCA cycle in the oxidative direction has been reported[33], though no prior examples of high TCA flux to produce a reduced product exist in the peer-reviewed literature. To the best of our knowledge, this is the first example of using the oxidative TCA cycle to supply reducing equivalents under anaerobic or microaerobic cultivation conditions for the production of small molecules in an *E. coli* host strain. This strategy provides a new avenue for metabolic engineering and could apply to the design of other biocatalysts.

The new host was not able to grow anaerobically, even with the BDO pathway expressed. This was not unexpected, as it is well documented that strains with similar knockout combinations do not grow under anaerobic or near-anaerobic conditions[41,42]. However, with microaeration these cells grew well and produced much more
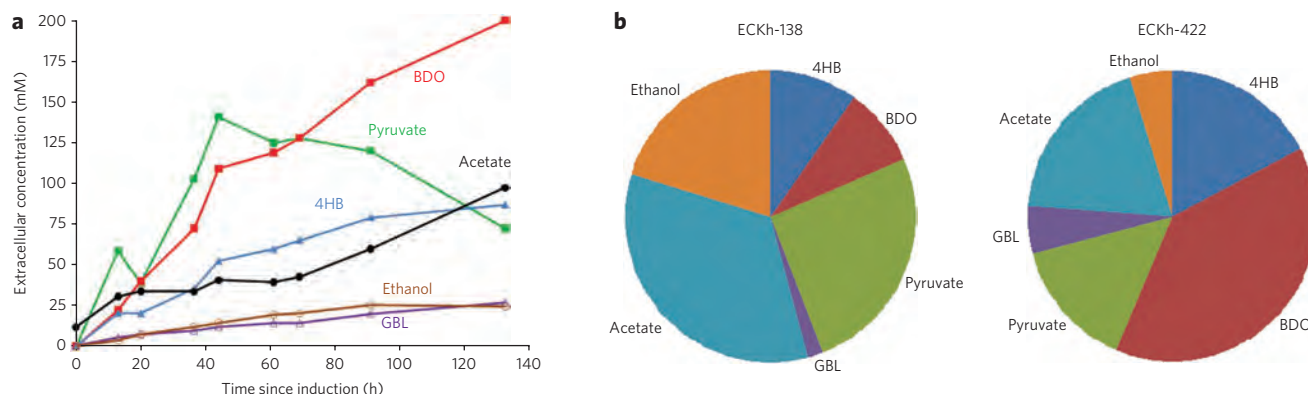

Figure 5 | Production of BDO from glucose in 2-l fed-batch fermentation using the OptKnock strain ECKh-422. The host was transformed with plasmids pZS*13S-sucCD-sucD-4hbd/sucA and pZE23S-025B-34. (**a**) Fermentation profile. Extracellular concentrations are in fermentation broth. (**b**) Comparison of product profile at the end of this fermentation run (right) with that from a fermentation with ECKh-138 (left).

BDO than the wild-type host. We chose microaerobic conditions such that sufficient NAD(P)H is available to drive the BDO pathway, but any excess beyond the capacity of the pathway is respired to generate ATP that may be necessary for cell growth, maintenance and product transport. Dissolved oxygen concentration in the fermentation, measured by a polarographic probe, was always 0, indicating that the cell experienced a severely oxygen-limited environment. Further adjustments to the BDO pathway will improve the ability of this host to consume NAD(P)H and thus reduce the host's aeration requirement.

The work outlined here shows production of BDO by a heterologous pathway, followed by an increase in BDO titer by over three orders of magnitude, to nearly 20 g l^{-1}. We recognize that commercialization will require another three- to five-fold increase, and to this end we are taking steps to raise the rates of key steps in the pathway, remove metabolic inefficiencies and substantially reduce byproducts. Systems-biology approaches that combine modeling technologies with 'omics' data measurements can identify and address bottlenecks that are obstacles to commercialization[43]. Fermentation process engineering can also optimize the feeding and aeration strategy to achieve higher cell densities and improve specific productivity. Finally, the ability to use lower-cost crude feedstocks will make biological BDO production more economically attractive, and feedstock flexibility will allow manufacturers to take advantage of geographic differences in sugar availability. Here we have described the production of BDO from pure and mixed C5 and C6 sugars and show that these sugars do not pose any additional technical barriers.

METHODS

Cultivation conditions. Twenty-milliliter bottle cultivations for metabolite production or bioconversion were performed in M9 minimal salts medium (6.78 g l^{-1} Na$_2$HPO$_4$, 3.0 g l^{-1} KH$_2$PO$_4$, 0.5 g l^{-1} NaCl, 1.0 g l^{-1} NH$_4$Cl, 1 mM MgSO$_4$, 0.1 mM CaCl$_2$) supplemented with 10 mM NaHCO$_3$, 20 g l^{-1} D-glucose and 100 mM MOPS to improve the buffering capacity, 10 µg ml^{-1} thiamine and the appropriate antibiotics for plasmid maintenance (**Table 1**). *E. coli* strain MG1655 *lac*^{IQ}+ was grown anaerobically, and we obtained anaerobic conditions by flushing capped anaerobic bottles with nitrogen for at least 5 min. Microaerobic conditions were used for all other strains, which we established by initially flushing capped anaerobic bottles with nitrogen for 5 min, then piercing the septum with a 23G needle (Becton-Dickenson) after inoculation. The needle was kept in the bottle during growth to allow a small amount of air to enter. Protein expression was induced with 0.2 mM IPTG when the culture reached mid-log growth phase, unless otherwise indicated in the text. For 4HB-to-BDO conversion experiments, 50 mM 4HB was added at the time of induction. All cultures were grown at 37 °C.

Fermentations were performed with 1 l initial culture volume in 2-l Biostat B+ bioreactors (Sartorius-Stedim Biotech) using modified M9 minimal medium (6.78 g l^{-1} Na$_2$HPO$_4$, 3.0 g l^{-1} KH$_2$PO$_4$, 0.5 g l^{-1} NaCl, 2.0 g l^{-1} NH$_4$Cl, 1.0 g l^{-1} (NH$_4$)$_2$SO$_4$, 1 mM MgSO$_4$, 0.1 mM CaCl$_2$) supplemented with 20 g l^{-1} D-glucose. The temperature was held at 37 °C, and the pH was held at 7.0 using 2 M NH$_4$OH (aerobic phase) or Na$_2$CO$_3$ (microaerobic phase). Cells were grown aerobically to an A_{600} of approximately 10, at which point the cultures were induced with 0.2 mM IPTG. One hour after induction, the airflow rate was reduced to 0.02 standard liters per minute for microaerobic conditions, and NaHCO$_3$ was added to a concentration of 10 mM. The agitation rate was set at 700 r.p.m. Concentrated glucose was fed in to maintain the glucose concentration in the vessel between 0.5 and 5.0 g l^{-1}.

Bacterial strains and plasmids. Genomic DNA from bacterial strains was isolated with the PureLink Genomic DNA Mini Kit (Invitrogen) according to the manufacturer's instructions. Recombinant DNA manipulations were conducted as described[44]. Genes and open reading frames in the study were amplified from appropriate genomic DNA templates using the high-fidelity KOD DNA polymerase enzyme (EMD Chemicals). Analytical PCR experiments for genotyping and sequencing were conducted according to standard molecular biology protocols with Taq polymerase (Promega). DNA sequencing was provided by Genewiz.

Development of expression vectors for BDO pathway. Vector backbones were obtained from R. Lutz (Expressys) and are based on the pZ Expression System[40]. The vectors obtained were pZE13luc, pZA33luc, pZS*13luc and pZE22luc, and these contained the luciferase gene as a stuffer fragment. The luciferase stuffer fragment was replaced as described in the **Supplementary Methods**. All vectors have the pZ designation followed by letters and numbers indicating the origin of replication, antibiotic resistance marker and promoter or regulatory unit.

The origin of replication is the second letter and is denoted by E for ColE1-based origins, A for p15A-based origins or S for pSC101-based origins. A lower–copy number version of pSC101 is designated S*. The first number represents the antibiotic resistance marker (1 for ampicillin, 2 for kanamycin, 3 for chloramphenicol). The final number defines the promoter that regulated the gene of interest (1 for P$_{LtetO-1}$, 2 for P$_{LlacO-1}$ and 3 for P$_{A1lacO-1}$). For the work discussed here, we used three base vectors, pZS*13S, pZA33S and pZE13S, modified for the biobricks insertions as discussed above.

Codon optimization of genes 002C and 025B. Codon optimization was performed using the neighbor-correlation method[27], which optimizes each codon in the specific context of its neighboring codons on the basis of *E. coli* usage. The optimized nucleotide sequences of each of these genes are shown in **Supplementary Figures 13** and **14** and were synthesized commercially (Geneart).

Constraint-based modeling and OptKnock. The OptKnock strain-design methodology[29] was applied within an iterative framework to identify multiple sets of knockouts leading to the coupling of BDO production to cell growth in *E. coli*. A genome-scale metabolic model of *E. coli* based on *iJR904*[14] was implemented for this analysis using Genomatica's in-house SimPheny computational platform (http://www.genomatica.com). A non–growth-associated energetic maintenance requirement of 7.6 mmol gDW^{-1} h^{-1} was assumed, along with a maximum specific glucose uptake rate of 20 mmol gDW^{-1} h^{-1}. BDO was assumed to be exported via facilitated diffusion. The final result was a list, ranked in terms of the maximum BDO yield at the optimal biomass yield, of distinct deletion strategies that differed from one another by at least one knockout. Knockout strategies derived by OptKnock, along with the metabolic model used in this study, are provided in **Supplementary Data 1–3**.

Design of new BDO pathways using the BioPathway Predictor algorithm. BioPathway Predictor is a computational tool implemented in Genomatica's SimPheny platform for enumerating and evaluating networks of enzyme-catalyzed reactions, with the goal of identifying novel pathways for producing a chemical of interest (**Fig. 1**). Reactions in a BioPathway network are generated by a defined set of reaction rules that operate repeatedly on a specified starting material and its derivatives[8]. Reaction rules are derived from generalized enzyme chemistry functions on the basis of recommendations by the International Union of Pure and Applied Chemistry (IUPAC)[45]. The set of reactions used is provided in **Supplementary Data 4**. Using the calculated reaction network, BioPathway Predictor generates potential biosynthetic routes between any two molecules contained in the network. Predicted biopathways are then ranked using criteria such as percentage of reactions catalyzed by characterized enzymes, maximum theoretical product yields, and thermodynamic feasibility. A flowchart of this process is provided in **Figure 1**.

Additional methods. Details of chromosomal manipulations, enzyme assays and analytical procedures are provided in the **Supplementary Methods**.

References

1. Carole, T.M., Pellegrino, J. & Paster, M.D. Opportunities in the industrial biobased products industry. *Appl. Biochem. Biotechnol.* **113–116**, 871–885 (2004).
2. Kroschwitz, J.I. & Howe-Grant, M. *Encyclopedia of Chemical Technology* (John Wiley and Sons, New York, 1993).
3. Nakamura, C.E. & Whited, G.M. Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotechnol.* **14**, 454–459 (2003).
4. Altaras, N.E. & Cameron, D.C. Metabolic engineering of a 1,2-propanediol pathway in *Escherichia coli*. *Appl. Environ. Microbiol.* **65**, 1180–1185 (1999).
5. Atsumi, S. *et al.* Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab. Eng.* **10**, 305–311 (2008).
6. Lindberg, P., Park, S. & Melis, A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab. Eng.* **12**, 70–79 (2010).
7. Qian, Z.G., Xia, X.X. & Lee, S.Y. Metabolic engineering of *Escherichia coli* for the production of putrescine: a four carbon diamine. *Biotechnol. Bioeng.* **104**, 651–662 (2009).
8. Lee, S.J., Song, H. & Lee, S.Y. Genome-based metabolic engineering of *Mannheimia succiniciproducens* for succinic acid production. *Appl. Environ. Microbiol.* **72**, 1939–1948 (2006).
9. Moon, T.S., Yoon, S.H., Lanza, A.M., Roy-Mayhew, J.D. & Prather, K.L. Production of glucaric acid from a synthetic pathway in recombinant *Escherichia coli*. *Appl. Environ. Microbiol.* **75**, 589–595 (2009).
10. Rathnasingh, C., Raj, S.M., Jo, J.E. & Park, S. Development and evaluation of efficient recombinant *Escherichia coli* strains for the production of 3-hydroxypropionic acid from glycerol. *Biotechnol. Bioeng.* **104**, 729–739 (2009).

11. Lee, K.H., Park, J.H., Kim, T.Y., Kim, H.U. & Lee, S.Y. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* **3**, 149 (2007).

12. Park, J.H., Lee, K.H., Kim, T.Y. & Lee, S.Y. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc. Natl. Acad. Sci. USA* **104**, 7797–7802 (2007).

13. Hong, S.H. & Lee, S.Y. Metabolic flux analysis for succinic acid production by recombinant *Escherichia coli* with amplified malic enzyme activity. *Biotechnol. Bioeng.* **74**, 89–95 (2001).

14. Reed, J.L., Vo, T.D., Schilling, C.H. & Palsson, B.O. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).

15. Cho, A., Yun, H., Park, J.H., Lee, S.Y. & Park, S. Prediction of novel synthetic pathways for the production of desired chemicals. *BMC Syst. Biol.* **4**, 35 (2010).

16. Hatzimanikatis, V. *et al.* Exploring the diversity of complex metabolic networks. *Bioinformatics* **21**, 1603–1609 (2005).

17. Henry, C.S., Broadbelt, L.J. & Hatzimanikatis, V. Discovery and analysis of novel metabolic pathways for the biosynthesis of industrial chemicals: 3-hydroxypropanoate. *Biotechnol. Bioeng.* **106**, 462–473 (2010).

18. Mavrovouniotis, M.L. Group contributions for estimating standard gibbs energies of formation of biochemical compounds in aqueous solution. *Biotechnol. Bioeng.* **36**, 1070–1082 (1990).

19. Burk, M.B., Burgard, A.P., Osterhout, R.E. & Sun, J. Microorganisms for the production of 1,4-butanediol. International Bureau patent WO/2010/030711 (2010).

20. Saito, N. *et al.* Metabolite profiling reveals YihU as a novel hydroxybutyrate dehydrogenase for alternative succinic semialdehyde metabolism in *Escherichia coli. J. Biol. Chem.* **284**, 16442–16451 (2009).

21. Söhling, B. & Gottschalk, G. Purification and characterization of a coenzyme-A-dependent succinate-semialdehyde dehydrogenase from *Clostridium kluyveri. Eur. J. Biochem.* **212**, 121–127 (1993).

22. Söhling, B. & Gottschalk, G. Molecular analysis of the anaerobic succinate degradation pathway in *Clostridium kluyveri. J. Bacteriol.* **178**, 871–880 (1996).

23. Valentin, H.E., Zwingmann, G., Schonebaum, A. & Steinbuchel, A. Metabolic pathway for biosynthesis of poly(3-hydroxybutyrate-co-4-hydroxybutyrate) from 4-hydroxybutyrate by *Alcaligenes eutrophus. Eur. J. Biochem.* **227**, 43–60 (1995).

24. Tian, J., Bryk, R., Itoh, M., Suematsu, M. & Nathan, C. Variant tricarboxylic acid cycle in Mycobacterium tuberculosis: identification of alpha-ketoglutarate decarboxylase. *Proc. Natl. Acad. Sci. USA* **102**, 10670–10675 (2005).

25. Jewell, J.B., Coutinho, J.B. & Kropinski, A.M. Bioconversion of propionic, valeric, and 4-hydroxybutyric acids into the corresponding alcohols by *Clostridium acetobutylicum* NRRL 527. *Curr. Microbiol.* **13**, 215–219 (1986).

26. Nair, R.V., Bennett, G.N. & Papoutsakis, E.T. Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol.* **176**, 871–885 (1994).

27. Sivaraman, K., Seshasayee, A., Tarwater, P.M. & Cole, A.M. Codon choice in genes depends on flanking sequence information–implications for theoretical reverse translation. *Nucleic Acids Res.* **36**, e16 (2008).

28. Scherf, U. & Buckel, W. Purification and properties of 4-hydroxybutyrate coenzyme A transferase from *Clostridium aminobutyricum. Appl. Environ. Microbiol.* **57**, 2699–2702 (1991).

29. Burgard, A.P., Pharkya, P. & Maranas, C.D. OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657 (2003).

30. Snoep, J.L. *et al.* Differences in sensitivity to NADH of purified pyruvate dehydrogenase complexes of *Enterococcus faecalis, Lactococcus lactis, Azotobacter vinelandii* and *Escherichia coli*: implications for their activity in vivo. *FEMS Microbiol. Lett.* **114**, 279–283 (1993).

31. Kim, Y., Ingram, L.O. & Shanmugam, K.T. Dihydrolipoamide dehydrogenase mutation alters the NADH sensitivity of pyruvate dehydrogenase complex of *Escherichia coli* K-12. *J. Bacteriol.* **190**, 3851–3858 (2008).

32. Menzel, K., Zeng, A.P. & Deckwer, W.D. Enzymatic evidence for an involvement of pyruvate dehydrogenase in the anaerobic glycerol metabolism of *Klebsiella pneumoniae. J. Biotechnol.* **56**, 135–142 (1997).

33. Iuchi, S. & Lin, E.C. arcA (dye), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. *Proc. Natl. Acad. Sci. USA* **85**, 1888–1892 (1988).

34. Stokell, D.J. *et al.* Probing the roles of key residues in the unique regulatory NADH binding site of type II citrate synthase of *Escherichia coli. J. Biol. Chem.* **278**, 35435–35443 (2003).

35. Pereira, D.S., Donald, L.J., Hosfield, D.J. & Duckworth, H.W. Active site mutants of *Escherichia coli* citrate synthase. Effects of mutations on catalytic and allosteric properties. *J. Biol. Chem.* **269**, 412–417 (1994).

36. Toth, J., Ismaiel, A.A. & Chen, J.S. The ald gene, encoding a coenzyme A-acylating aldehyde dehydrogenase, distinguishes *Clostridium beijerinckii* and two other solvent-producing clostridia from *Clostridium acetobutylicum. Appl. Environ. Microbiol.* **65**, 4973–4980 (1999).

37. Riley, M. *et al. Escherichia coli* K-12: a cooperatively developed annotation snapshot–2005. *Nucleic Acids Res.* **34**, 1–9 (2006).

38. Orencio-Trejo, M. *et al.* Metabolic regulation analysis of an ethanologenic *Escherichia coli* strain based on RT-PCR and enzymatic activities. *Biotechnol. Biofuels.* **1**, 8 (2008).

39. Sprenger, G.A. & Lengeler, J.W. Analysis of sucrose catabolism in *Klebsiella pneumoniae* and in Scr⁺ derivatives of *Escherichia coli* K12. *J. Gen. Microbiol.* **134**, 1635–1644 (1988).

40. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).

41. Donnelly, M.I., Millard, C.S., Clark, D.P., Chen, M.J. & Rathke, J.W. A novel fermentation pathway in an *Escherichia coli* mutant producing succinic acid, acetic acid, and ethanol. *Appl. Biochem. Biotechnol.* **70–72**, 187–198 (1998).

42. Kim, Y., Ingram, L.O. & Shanmugam, K.T. Construction of an *Escherichia coli* K-12 mutant for homoethanologenic fermentation of glucose or xylose without foreign genes. *Appl. Environ. Microbiol.* **73**, 1766–1771 (2007).

43. Mahadevan, R., Burgard, A., Famili, I., Van Dien, S. & Schilling, C. Applications of metabolic modeling to drive bioprocess development for the production of value-added chemicals. *Biotechnol. Bioprocess Eng.* **10**, 408–417 (2005).

44. Sambrook, J., Fritsh, E.F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989).

45. Webb, E.C. *Enzyme Nomenclature 1992: Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes* (Academic Press, San Diego, 1992).

Author contributions

H.Y., R.H. and W.N. cloned and expressed BDO-pathway genes, performed bottle experiments and wrote the manuscript; C.P.-B. constructed the host strain and wrote the manuscript; A.B. conceived the project, performed simulations and wrote the manuscript; J.B. cloned and expressed BDO-pathway genes; J.K. and R.S. performed analytical work; J.D.T. conceived the project and performed simulations; R.E.O. performed simulations and wrote the manuscript; J.E. constructed the host strain and performed bottle experiments; S.T. and H.B.S. performed fermentations; S.A. developed and carried out enzyme assays; T.H.Y. analyzed ¹³C data for flux analysis; S.Y.L. conceived the project and wrote the manuscript; M.J.B. and S.V.D. conceived and directed the project and wrote the manuscript.

Competing financial interests

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturechemicalbiology/.

Additional information

Supplementary information is available online at http://www.nature.com/naturechemicalbiology/. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to S.V.D.

Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol

Harry Yim, Robert Haselbeck, Wei Niu, Catherine Pujol-Baxley, Anthony Burgard, Jeff Boldt, Julia Khandurina, John D Trawick, Robin E Osterhout, Rosary Stephen, Jazell Estadilla, Sy Teisan, H Brett Schreyer, Stefan Andrae, Tae Hoon Yang, Sang Yup Lee, Mark J Burk & Stephen Van Dien

EDITORIAL SUMMARY

AOP: The design and implementation of a high-yielding enzymatic route to 1,4-butanediol—a compound not known to be produced naturally—provides a compelling example of how metabolic engineering can be harnessed for the microbial conversion of carbohydrate feedstocks to desired small molecules.

Biopathway prediction

Metabolic design and engineering

Microbial biocatalysts for chemical production