nature biotechnology

Conversion of proteins into biofuels by engineering nitrogen flux

Yi-Xin Huo[1], Kwang Myung Cho[1,2], Jimmy G Lafontaine Rivera[1], Emma Monte[1], Claire R Shen[1], Yajun Yan[1] & James C Liao[1,2]

Biofuels are currently produced from carbohydrates and lipids in feedstock. Proteins, in contrast, have not been used to synthesize fuels because of the difficulties of deaminating protein hydrolysates. Here we apply metabolic engineering to generate *Escherichia coli* that can deaminate protein hydrolysates, enabling the cells to convert proteins to C4 and C5 alcohols at 56% of the theoretical yield. We accomplish this by introducing three exogenous transamination and deamination cycles, which provide an irreversible metabolic force that drives deamination reactions to completion. We show that *Saccharomyces cerevisiae*, *E. coli*, *Bacillus subtilis* and microalgae can be used as protein sources, producing up to 4,035 mg/l of alcohols from biomass containing ~22 g/l of amino acids. These results show the feasibility of using proteins for biorefineries, for which high-protein microalgae could be used as a feedstock with a possibility of maximizing algal growth[1] and total CO_2 fixation.

Current biorefinery schemes are suboptimal for several reasons. First, schemes based on algae have limited efficiency because cultures used for biofuel production must be starved so that they produce lipid feedstocks, resulting in less cell growth and less total CO_2 fixation. Second, all existing schemes, including sugar-based or cellulosic biorefining, lead to the accumulation of protein by-products, but there are no strategies to convert these by-products into liquid fuels. These protein by-products are typically used as animal feed. But despite the current profitability of animal feed, the feed market has a limited ability to absorb the increasing protein by-products from the fast-expanding biorefinery industry[2]. Third, in all existing schemes, reduced nitrogen is not recycled, resulting in a net loss of reduced nitrogen[3,4] and an increase in nitrous oxide production, which is a greenhouse gas almost 300 times more potent than CO_2 (ref. 5). Moreover, the lost reduced nitrogen must be replaced by supplementing future crops with reduced fertilizer nitrogen, which is produced by the energy-intensive and environmentally unfriendly Haber-Bosch process[6].

A heretofore unexplored solution to these limitations would be to use proteins as a feedstock. If proteins were deaminated and converted to fuel or chemicals, the reduced nitrogen could be recycled to close the nitrogen loop. More importantly, proteins are the major component of the photosynthesis apparatus, CO_2 fixation pathways and other biosynthetic and cell growth machinery. Proteins are the dominant fraction in fast-growing photosynthetic microorganisms[1,7] (**Supplementary Fig. 1**) and industrial fermentation residues (**Supplementary Table 1**). Thus, using proteins as feedstock might maximize growth and CO_2 fixation rates. However, current schemes aim to increase production of carbohydrates or lipids, rather than proteins.

Before protein hydrolysates, which are mixtures of peptides and amino acids, can be converted to biofuels or chemicals, their carbon skeletons must be released from amino groups by enzymatic reactions that involve deamination, transamination or dehydrogenation. However, deamination of protein hydrolysates is limited by both thermodynamic reversibility and biological regulation that favors anabolism rather than catabolism[8]. But if a metabolic driving force could be devised and the biological regulation could be properly reprogrammed, some amino acids could be directly deaminated to 2-keto acids, which can be converted to many chemicals, such as aldehydes, by 2-keto acid decarboxylases that have a broad range of substrate specificity, and then to alcohols by alcohol dehydrogenases[9] (**Supplementary Fig. 2**). Other amino acids could be deaminated to tricarboxylic acid cycle intermediates, which can be directed to pyruvate by gluconeogenic enzymes, such as malic enzymes or phosphoenolpyruvate carboxykinase. Pyruvate is a central metabolite in sugar-based biorefining, and it can be further extended to longer keto acids by acetohydroxy acid synthase or isopropylmalate synthase chain elongation pathways[9,10] (**Supplementary Fig. 2**). These synthetic routes enable proteins to be utilized for biorefining and fuel production. However, because of the complexities of nitrogen regulation, the native pathways for converting amino acids to higher alcohols support only a small amount of product formation[8]. Thus, a metabolic engineering strategy involving disrupted regulation, rewired metabolism and a driving force is needed to enable protein hydrolysates to be used for biorefining and fuel production.

We chose *E. coli* as the host organism for engineering because of its versatility[9,11–13]. To test its efficiency at metabolizing protein hydrolysates, we grew the cells in yeast (*Saccharomyces cerevisiae*) extract or mixtures of 20 amino acids, which simulated algae hydrolysates. Although *E. coli* grew well in these rich media, it has been observed that under similar conditions the amino acids were used sequentially and incompletely[14]. When we introduced the isobutanol synthesis pathway into *E. coli* by overexpressing the genes *alsS*, *ilvC*, *ilvD*, *KivD* and *AdhA*, the cells could produce only 0.17 g/l of isobutanol from medium containing 4% yeast

[1]Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, Los Angeles, California, USA. [2]Institute for Genomics and Proteomics, University of California, Los Angeles, Los Angeles, California, USA. Correspondence should be addressed to J.C.L. (liaoj@seas.ucla.edu).

Received 30 November 2010; accepted 26 January 2011; published online 6 March 2011; doi:10.1038/nbt.1789

Table 1 Higher alcohol (C ≥ 4) production in *E. coli* from a yeast extract medium containing 21.6 g/l amino acids, of which 14 amino acids could be converted to higher alcohols

No. of strain	Host strain	Genes overexpressed	Consumed AA (g/l)	Biomass (g/l)	Alcohol C ≥ 4 (mg/l)[a]	Percent of theoretical yield	Remarks
(1)	JCL16	None	19.8 ± 0.4	3.6 ± 0.3	0	0	Wild type
(2)	YH19	None	18.3 ± 0.5	3.8 ± 0.4	0	0	High AA consumer
(3)	JCL16	*alsS, ilvC, ilvD, kivD, AdhA*	17.7 ± 0.9	1.1 ± 0.2	170 ± 15	2.3	(1) + iBOH pathway
(4)	YH19	*alsS, ilvC, ilvD, kivD, AdhA*	16.6 ± 0.8	1.2 ± 0.2	406 ± 31	5.6	(2) + iBOH pathway
(5)	YH19Δ*luxS*	*alsS, ilvC, ilvD, kivD, AdhA*	16.7 ± 0.7	2.0 ± 0.3	895 ± 105	12.4	(4) + quorum-sensing deletion
(6)	YH19Δ*lsrA*	*alsS, ilvC, ilvD, kivD, AdhA*	16.8 ± 0.5	1.9 ± 0.2	952 ± 56	13.1	(4) + quorum-sensing deletion
(7)	YH19Δ*glnA*Δ*gdhA*	*alsS, ilvC, ilvD, kivD, AdhA*	10.0 ± 0.5	0.6 ± 0.1	420 ± 46	5.8	(4) + NH_3-assimilation deletion
(8)	YH19Δ*glnA*Δ*gdhA*Δ*luxS*	*alsS, ilvC, ilvD, kivD, AdhA*	9.5 ± 0.5	0.7 ± 0.1	1,334 ± 178	18.4	(5) + NH_3-assimilation deletion
(9)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*	*alsS, ilvC, ilvD, kivD, AdhA*	9.7 ± 0.4	0.7 ± 0.1	1,490 ± 135	20.6	(6) + NH_3-assimilation deletion
(10)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*Δ*ilvE*	*alsS, ilvC, ilvD, kivD, AdhA*	8.9 ± 0.5	0.6 ± 0.1	1,207 ± 107	16.7	(9) + BCAA-transamination deletion
(11)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*	*alsS, ilvC, ilvD, kivD, yqhD*	9.5 ± 0.5	0.7 ± 0.1	1,277 ± 158	17.6	(9) but *yqhD* instead of *AdhA*
(12)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*	*alsS, ilvC, ilvD, kivD, yqhD, LeuDH*	17.0 ± 0.3	0.7 ± 0.1	3,823 ± 235	52.8	(11) + 1st transamination cycle
(13)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*Δ*ilvE*	*alsS, ilvC, ilvD, kivD, yqhD, LeuDH*	11.1 ± 0.4	0.6 ± 0.1	3,557 ± 278	49.1	(12) + BCAA-transamination deletion
(14)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*	*alsS, ilvC, ilvD, kivD, yqhD, LeuDH, AvtA*	17.2 ± 0.2	0.7 ± 0.1	3,898 ± 326	53.8	(12) + 2nd transamination cycle
(15)	YH19Δ*glnA*Δ*gdhA*Δ*lsrA*	*alsS, ilvC, ilvD, kivD, yqhD, LeuDH, AvtA ilvE, ilvA, sdaB*	18.0 ± 0.3	0.7 ± 0.1	4,035 ± 294	55.7	(14) + 3rd transamination cycle

[a]The theoretical maximum titer is 7,240 mg/l (see also **Supplementary Table 4b**). The products were identified by gas chromatography–mass spectrometry and quantified by gas chromatography–flame ionization detector. Quantities given as mean ± s.d. ($n = 3$). BCAA, branched-chain amino acids.

extract supplemented with M9 salts (**Table 1**), representing 2.3% of the theoretical yield.

To improve amino acid utilization, we first performed several rounds of chemical mutagenesis followed by growth on either single or multiple amino acids[15,16], obtaining strain YH19, which could utilize 13 individual amino acids (compared to four for the wild type) as the sole carbon source for growth (**Supplementary Fig. 3**). We then tested 20 single-gene deletion strains, each missing a gene thought to regulate carbon and nitrogen metabolism or participate in pathways that compete with fuel production for common intermediates. We introduced the isobutanol pathway into these strains and screened them for improved production (**Fig. 1a**). Strains missing the quorum-sensing genes *luxS* or *lsrA* showed increased production of isobutanol. LuxS is the enzyme that synthesizes the signaling molecule autoinducer-2 (AI-2). LsrA is one of the subunits of the AI-2 transporter (LsrABCD)[17], which is involved in the re-uptake of extracellular AI-2 during transition into the stationary phase[17]. The production of isobutanol causes stress to the cell, and AI-2 re-uptake acts to reduce growth in the presence of stress[18]. Deletion of either AI-2 synthesis or its reuptake system in YH19 yielded a strain that grew despite the stress generated by increased fuel production (**Supplementary Fig. 4** and **Fig. 1b,c**). Because AI-2 re-uptake is inhibited by glucose[19], these deletions had no effect on the strains containing the isobutanol pathway that were grown in glucose medium. Cells harboring the deletion but missing the isobutanol pathway did not have a growth pattern noticeably different from that of wild type (data not shown).

To further drive amino acid degradation, we designed a strategy focusing on redirecting nitrogen flux (**Fig. 2** and **Supplementary Fig. 5**), rather than carbon flux (**Supplementary Fig. 6**). We noted that amino acid utilization is complicated by the nitrogen flux between various amino acids[20,21]. The cell preferentially degrades amino acids in sequence, transferring the amino group from some amino acids (Ser, Asn, Asp, Thr, Arg, Cys and Ala) to synthesize others, such as glutamic acid and branched chain amino acids, by transamination reactions in order to store nitrogen in the cell. All of the amino acids except aspartic acid that are preferentially degraded are typically directly deaminated[14] (**Supplementary Fig. 5**). After ammonia is produced by the deamination reactions, it is taken up by ammonia assimilation enzymes GdhA and GlnA[20,22].

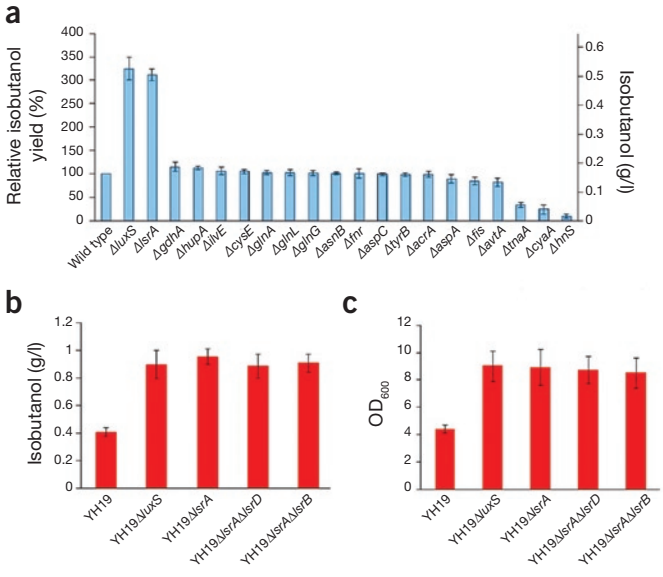

Figure 1 Screening of regulatory mutants for improved isobutanol production. All strains contain overexpressed isobutanol pathway genes (*alsS, ilvC, ilvD, kivD, AdhA*). Error bars indicate s.d. ($n = 3$). (**a**) Isobutanol titer from 20 strains containing different single-gene knockouts relative to the titer produced by the wild-type host (JCL16). (**b**,**c**) Deletion of quorum-sensing genes in YH19 (the high amino acid–consuming mutant) improved isobutanol production (**b**) and growth (**c**).

To block ammonia re-uptake and drive the nitrogen flux toward deamination, we deleted the two ammonium-assimilation genes, *gdhA* and *glnA*. This created a metabolic driving force for deamination and improved the degradation of several amino acids (Ser, Asn, Thr, Arg, Cys, Gly and Gln) that directly generate ammonia when they are degraded[20] (**Supplementary Fig. 5**). Indeed, deletion of *gdhA* and *glnA* increased the production of alcohols in the presence of the keto-acid pathway (**Table 1**). Other amino acids (Asp, Ala, Leu, Ile and Val) are degraded by transamination, and the amino groups are transferred to 2-ketoglurate to form glutamate[20] (**Fig. 2a**). The amino group on glutamate is further redistributed to branched-chain amino acids. Thus, glutamate and branched-chain amino acids accumulated in the medium as a result of transamination during the course of the bacteria cultivation. It appears that these groups of amino acids serve as a nitrogen reservoir to keep reduced nitrogen inside the cell in the presence of keto acid pathway. Because this reduced nitrogen reservoir is attached to carbon skeletons, the latter cannot be utilized for fuel production without deamination.

To release the carbon skeletons of amino acids, thereby draining the intracellular nitrogen reservoir, we introduced three exogenous transamination and deamination cycles by expressing *LeuDH* (from *Thermoactinomyces intermedius*) *ilvE*, *avtA*, *ilvA* and *sdaB* (all from *E. coli*) on plasmids. These three cycles all start from glutamate, which is also the transamination product of aspartic acid and alanine (**Fig. 2a**). The first cycle (**Fig. 2b**) utilizes the enzyme IlvE to transfer the amino group from glutamate to 2-keto methylvalerate and 2-keto isocaproate, forming isoleucine and leucine, respectively. Then, LeuDH catalyzes the deamination reaction of isoleucine and leucine to release the carbon skeletons, 2-keto methylvalerate and 2-keto isocaproate, respectively. The ammonia released in this reaction is excreted and cannot be reutilized because of the deletion of the ammonium-assimilation genes *glnA* and *gdhA*. As a result of this cycle, ammonia release becomes an irreversible trap, driving the transamination and deamination reactions. The

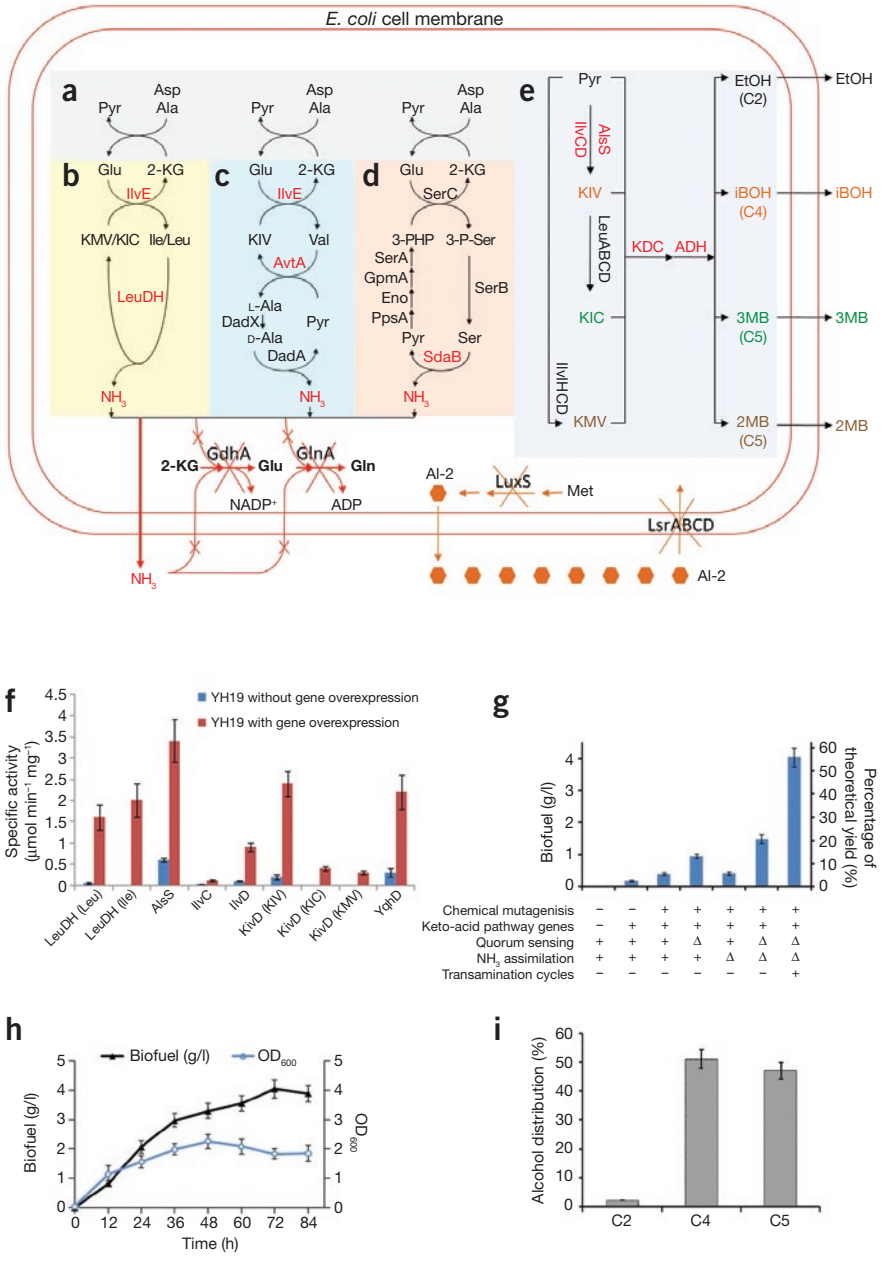

Figure 2 Nitrogen-centric metabolic engineering strategy in *E. coli*. Error bars indicate s.d. ($n = 3$). Amino acids that are directly deaminated are shown in **Supplemental Figure 5**. Others are deaminated through the reprogrammed transamination and deamination cycles (**a**–**d**). Overexpressed enzymes shown in red. Deletion of *gdhA* and *glnA* blocks ammonia re-uptake and provides a metabolic driving force to direct the nitrogen flux out of the cell. (**a**) The amino groups in Asp and Ala are transferred to 2-ketoglutarate (2-KG) to yield pyruvate (Pyr) and Glu through a series of reactions (not shown). (**b**) IlvE transfers the amino group from Glu to 2-ketomethylvalerate (KMV) or 2-ketoisocaproate (KIC) to yield Ile and Leu, respectively. The heterologous LeuDH from *Thermoactinomyces intermedius* then deaminates Ile and Leu to regenerate KMV and KIC, respectively. NH_3 is excreted but cannot be reassimilated because of Δ*gdhA* and Δ*glnA*. (**c**) IlvE also can transfer the amino group from Glu to 2-ketoisovalerate (KIV) and generate Val. AvtA then transfers the amino group from Val to Pyr to generate L-Ala, which is then converted to D-Ala and deaminated to yield Pyr. Excretion of NH_3 is used as a driving force. (**d**) SerC transfers the amino group from Glu to 3-phosphohydroxypyruvate (3-PHP) to yield 3-phosphoserine (3-P-Ser), which is then converted to Ser by SerB. Ser is deaminated by SdaB to generate Pyr. Pyr can be recycled to 3-PHP or used for fuel synthesis. Again, NH_3 excretion is used to drive the nitrogen flux. (**e**) Engineered keto acid pathways that produce ethanol (EtOH), isobutanol (iBOH), 2-methyl-1-butanol (2MB) and 3-methyl-1-butanol (3MB)[9]. (**f**) Enzyme activities of the overexpressed genes. For LeuDH and KivD, the substrates of specific reactions are given in parentheses. The substrate for YqhD is isobutyraldehyde. (**g**) Biofuel produced (EtOH, iBOH, 2MB, 3MB) from various engineering steps. Biofuel titers and the corresponding percentage of theoretical yield are marked. (**h**,**i**) Biofuel production and optical density (OD_{600}) (**h**) and chain-length distribution (**i**) from *E. coli* engineered using the nitrogen-centric metabolic engineering strategy shown in (**a**–**e**). The strain is YH83 (YH19Δ*glnA*Δ*gdhA*Δ*lsrA*/ overexpression of *alsS, ilvC, ilvD, avtA, LeuDH, kivD, yqhD, ilvE, ilvA* and *sdaB*).

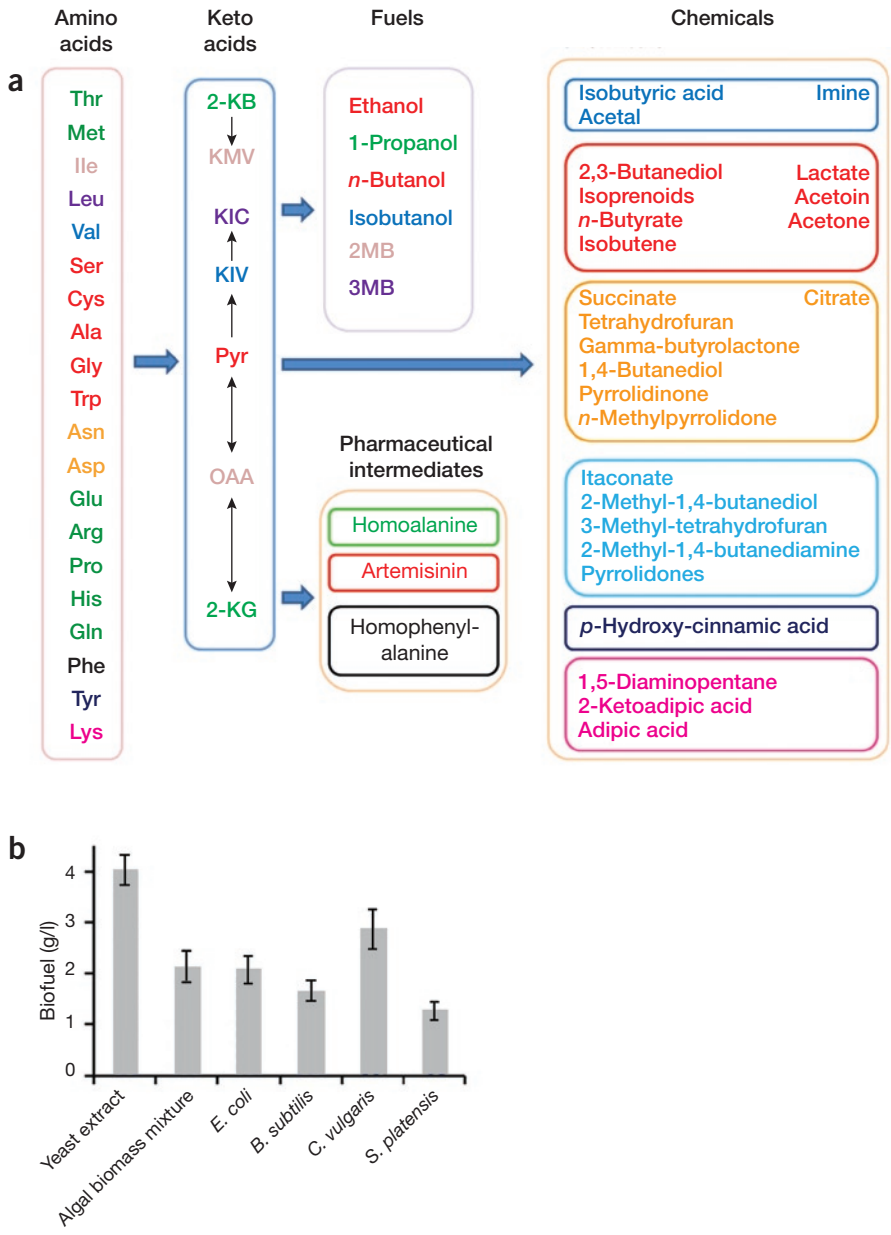

Figure 3 Biofuel production and biorefining scheme from algal or bacterial protein sources. (**a**) The proposed protein-based biorefinery scheme. Amino acids are deaminated to various keto acids, which are then used to produce fuels, chemicals and pharmaceutical intermediates. The colors link products and intermediates to the amino acids from which they are derived. (**b**) Biofuel (EtOH, iBOH, 2MB, 3MB) produced from the engineered *E. coli* strain YH83 grown in flasks using algal or bacterial cell hydrolysates. Small laboratory-scale reactors (1 liter or 30 liters) were used to grow bacterial and algal cells individually. The algal biomass mixture includes *C. vulgaris*, *P. purpureum*, *S. platensis* and *S. elongatus*. All protein sources were adjusted to contain 21.6 g/l peptides and amino acids. Error bars indicate s.d. ($n = 3$). OAA, oxaloacetate; 2-KB, 2-ketobutyrate.

carbon skeletons released in this cycle are 2-keto methylvalerate, 2-keto isocaproate, 2-ketoglutarate and pyruvate, which can be converted to higher alcohols (C > 2) or other carbon-based compounds (**Fig. 2b**). The second cycle (**Fig. 2c**) also uses the enzyme IlvE, which transfers the amino group from glutamate to 2-keto isovalerate and produces valine. Then, AvtA transfers the amino group from valine to pyruvate to form L-alanine, which is converted to D-alanine by DadX protein. Finally, D-alanine is deaminated by DadA protein to form pyruvate and drain nitrogen as ammonia, which is excreted (**Fig. 2c**). Similarly, the third cycle (**Fig. 2d**) uses SerBC, to transfer the amino group from glutamate to 3-phosphohydroxypyruvate to form serine. Also in this cycle, SdaB deaminates serine and drains nitrogen as ammonia, which is excreted.

The three engineered cycles all produce ammonia, pyruvate and keto acids. Ammonia can be collected and utilized as fertilizers. Pyruvate can be synthesized into higher alcohols or recycled to 3-phosphohydroxypyruvate by a series of enzymes (PpsA, Eno, GpmA and SerA) (**Fig. 2d**). Keto acids can be converted to various alcohols, such as ethanol, isobutanol, 2-methyl-1-butanol and 3-methyl-1-butanol[9] (**Fig. 2e**). To produce 2-methyl-1-butanol and 3-methyl-1-butanol, we replaced AdhA in the isobutanol synthesize pathway with *E. coli* alcohol dehydrogenase YqhD, which has a broader substrate range[23]. All of the overexpressed enzymes are functional, as shown by their measured activities (**Fig. 2f**). Using the best strain, YH83, the production of biofuels (isobutanol, 2-methyl-1-butanol and 3-methyl-1-butanol) was 4.0 g/l (**Fig. 2g,h**), which represents 56% of the theoretical yield (**Table 1** and **Fig. 2g**). The product is comprised of ~50% isobutanol, 47% C5 alcohols and 3% ethanol (**Fig. 2i**).

These results demonstrate that reengineering nitrogen flux for 14 amino acids, which we refer to as fuel-convertible amino acids, enables protein hydrolysates to be used for fuel production. The remaining amino acids include Lys, Met, His and the three aromatic amino acids. The transamination and deamination cycles for degrading these remaining amino acids are best conducted under fully aerobic conditions, which is very different from the micro-aerobic conditions used for degrading the fuel-convertible amino acids. Therefore, we converted them into a mixture of all 20 amino acids by using them as carbon and nitrogen sources for growing organisms such as *Pseudomonas* (**Supplementary Fig. 7**). The theoretical yield for using the six remaining amino acids to accumulate *Pseudomonas* protein is 41.4%, which contains 77.9% fuel-convertible amino acids (**Supplementary Table 2**). *Pseudomonas* protein lysate could then be used as a feedstock for biofuel production. Alternatively, many of the remaining amino acids could be used as raw materials for further biorefining (**Fig. 3a**) or simply as animal feed supplements. With this nitrogen-centric metabolic engineering strategy, all amino acids can be converted to their corresponding keto acids, and can in turn be further converted to fuels or chemicals[24] (**Fig. 3a**), directly or through key metabolites such as acetyl-CoA or succinate. The compounds that could be produced from proteins include bulk chemicals, monomers and pharmaceutical intermediates, in addition to fuels. We calculate that from a typical microbial amino acid composition (**Supplementary Fig. 8**), the maximum

theoretical yield of combined long-chain alcohols from protein biomass is 60% (**Supplementary Fig. 9** and **Supplementary Tables 3** and **4**) or 73% (net yield, excluding nitrogen in calculation), which is higher than that of ethanol from sugar (51%). Thus, the ability to deaminate protein hydrolysates provides a general strategy for protein-based biorefining, which as an added benefit also recycles ammonia for use as fertilizer for algae and plants, supporting further protein production.

Previous metabolic engineering strategies have focused on carbon flux. In contrast, the strategy used here focuses on nitrogen flux. The reprogrammed transamination and deamination cycles introduced in this work enable the effective deamination of protein hydrolysates. Unlike lignocelluloses, which are recalcitrant to degradation, proteins can be readily hydrolyzed by proteases to short peptides and amino acids (**Supplementary Fig. 10a**), which have been used industrially for various applications. Peptide bonds are readily accessible to proteases, and protein hydrolysis does not need to be complete, as microorganisms can take up short peptides and degrade them into amino acids intracellularly. Proteases (K_{cat} on the order of 100–1,000 s^{-1})[25] are efficient enzymes and can be secreted by many microorganisms naturally, facilitating consolidated bioprocessing. Microbial protein biomass can be readily hydrolyzed by a simple process of pretreatment and enzymatic hydrolysis (**Supplementary Fig. 10b**), especially biomass from single-cell organisms like microalgae that do not have thick cell walls. In addition, hydrolysis of proteins does not generate by-products such as furfural and 5-hydroxymethylfurfural that inhibit microbial growth (**Supplementary Fig. 11**), as lignocellulose hydrolysis does.

For large-scale applications, the protein raw material could come from several sources. In the short term, waste proteins generated from the current fermentation, food processing and biofuel production industries could be used (**Supplementary Table 1**). For example, regional bioethanol production facilities may provide about 10 metric tons/day of proteins (from 1,000 metric tons/day of biomass processed) in fermentation residues. These residues could be used as feedstock for a protein-based biorefinery, and the ammonia released in the transamination and deamination reactions could be recycled as fertilizers. In particular, genetically modified organisms used in fermentation cannot be disposed of as fertilizers or animal feed without additional treatment, and thus provide an excellent source material for protein-based biorefining. For long-term, large-scale applications, we envision using algal biomass, because proteins are the major component of fast-growing microalgae[1] (**Supplementary Fig. 1**) in open-pond cultures that are not artificially induced to accumulate lipids. Even with contamination by heterotrophic bacteria, the combined protein sources are still usable as raw material.

To test the feasibility of using algal and bacterial proteins as a feedstock, we grew green algae *Chlorella vulgaris*, red algae *Porphyridium purpureum*, green-blue algae *Spirulina platensis* and cyanobacterium *Synechococcus elongatus,* as well as *E. coli* and *B. subtilis* in 1 liter flasks or 30 liter fish tanks. Biomass was collected, digested with protease and used as feedstock for biofuel production with the engineered *E. coli* strain YH83. In all experiments, the protein concentration was adjusted to 21.6 g/l, which is equivalent to the amount of protein in 4% yeast extract. Except for *S. platensis* and *B. subtilis*, processing this biomass, comprised of either individual species or a mixture of species cultured separately, produced about 50–70% of what was produced from the same amount of yeast extract, even though the protein hydrolysis process was not optimized (**Fig. 3b**). These results demonstrate the feasibility of using protein derived from unicellular organisms, such as mixed algal species or fermentation residues from *E. coli* or *B. subtilis*.

In summary, we have demonstrated the feasibility of using proteins as raw material for biorefining by applying a nitrogen-centric metabolic engineering strategy. Although the protein biomass could also be used by anaerobic digestion for biogas production or thermochemical treatment for heat production, existing processes are not able to produce liquid fuels, bulk chemicals or pharmaceutical intermediates, specifically without high pressure and high-temperature conditions. For large-scale protein production using microalgae, natural selection under frequent or continuous harvesting conditions will favor fast-growing and robust microorganisms, which generally contain high-protein contents and are fully adapted to the local environment. Thus, it is possible that the proposed biorefining scheme can bypass the need for expensive photobioreactors or the lignocellulose recalcitrance problem by using protein biomass from algal cultures[1], waste biotreatment and the fermentation industry as a long-term, sustainable protein source. The recycled nitrogen could be used for future protein production or conventional agriculture. However, several challenges remain to be addressed, including large-scale algal production, harvesting, product purification and nitrogen recycling. In addition, energy and economic factors associated with the above challenges will need to be considered as the technology moves into practical application. It is conceivable that the potential advantages of using proteins as raw material for biorefining may stimulate development in this direction.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This work was partially supported by UCLA–Department of Energy Institute for Genomics and Proteomics.

AUTHOR CONTRIBUTIONS

Y.-X.H. designed and performed experiments, designed the theoretical calculation, analyzed data and wrote the manuscript; K.M.C. designed and performed experiments and analyzed data; J.G.L.R. designed and performed theoretical calculation; E.M. performed part of the experiments in **Table 1**; C.R.S. performed the chemical mutagenesis; Y.Y. designed and performed carbon-flux driven biofuel production; J.C.L. designed experiments, theoretical calculation, analyzed data 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/naturebiotechnology/

Published online at http://www.nature.com/naturebiotechnology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Sheehan, J., Dunahay, T., Benemann, J. & Roessler, P. *A Look Back at the US Department of Energy's Aquatic Species Program: Biodiesel from Algae; Close-Out Report* (National Renewable Energy Laboratory, Golden, CO, USA; 1998). <http://www.nrel.gov/biomass/pdfs/24190.pdf>
2. Wijffels, R.H. & Barbosa, M.J. An outlook on microalgal biofuels. *Science* **329**, 796 (2010).
3. Miller, S.A. Minimizing land use and nitrogen intensity of bioenergy. *Environ. Sci. Technol.*, **44**, 3932–3939 (2010).
4. Melillo, J.M. *et al.* Indirect emissions from biofuels: how important? *Science* **326**, 1397–1399 (2009).
5. Crutzen, P.J., Mosier, A.R., Smith, K.A. & Winiwarter, W. N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. *Atmos. Chem. Phys.* **8**, 389–395 (2008).
6. Erisman, J.W., Sutton, M.A., Galloway, J., Klimont, Z. & Winiwarter, W. How a century of ammonia synthesis changed the world. *Nature Geoscience* **1**, 636–639 (2008).
7. Becker, E.W. Micro-algae as a source of protein. *Biotechnol. Adv.* **25**, 207–210 (2007).
8. Hazelwood, L.A., Daran, J.M., van Maris, A.J.A., Pronk, J.T. & Dickinson, J.R. The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. *Appl. Environ. Microbiol.* **74**, 2259–2266 (2008).
9. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86–89 (2008).

10. Zhang, K., Sawaya, M.R., Eisenberg, D.S. & Liao, J.C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc. Natl. Acad. Sci. USA* **105**, 20653–20658 (2008).
11. Steen, E.J. *et al.* Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559–562 (2010).
12. Biebl, H., Menzel, K., Zeng, A.P. & Deckwer, W.D. Microbial production of 1, 3-propanediol. *Appl. Microbiol. Biotechnol.* **52**, 289–297 (1999).
13. Ingram, L.O., Conway, T., Clark, D.P., Sewell, G.W. & Preston, J.F. Genetic engineering of ethanol production in *Escherichia coli*. *Appl. Environ. Microbiol.* **53**, 2420–2425 (1987).
14. Pruss, B.M., Nelms, J.M., Park, C. & Wolfe, A.J. Mutations in NADH: ubiquinone oxidoreductase of *Escherichia coli* affect growth on mixed amino acids. *J. Bacteriol.* **176**, 2143–2150 (1994).
15. Connor, M.R., Cann, A.F. & Liao, J.C. 3-Methyl-1-butanol production in *Escherichia coli*: random mutagenesis and two-phase fermentation. *Appl. Microbiol. Biotechnol.* **86**, 1155–1164 (2010).
16. Cann, A.F. & Liao, J.C. Pentanol isomer synthesis in engineered microorganisms. *Appl. Microbiol. Biotechnol.* **85**, 893–899 (2010).
17. Xavier, K.B. & Bassler, B.L. Regulation of uptake and processing of the quorum-sensing autoinducer AI-2 in *Escherichia coli*. *J. Bacteriol.* **187**, 238–248 (2005).
18. DeLisa, M.P., Valdes, J.J. & Bentley, W.E. Mapping stress-induced changes in autoinducer AI-2 production in chemostat-cultivated *Escherichia coli* K-12. *J. Bacteriol.* **183**, 2918–2928 (2001).
19. Wang, L., Hashimoto, Y., Tsao, C.Y., Valdes, J.J. & Bentley, W.E. Cyclic AMP (cAMP) and cAMP receptor protein influence both synthesis and uptake of extracellular autoinducer 2 in *Escherichia coli*. *J. Bacteriol.* **187**, 2066–2076 (2005).
20. Neidhart, F.C. *Escherichia coli and Salmonella* (American Society for Microbiology, 1996).
21. Atkinson, M.R., Blauwkamp, T.A., Bondarenko, V., Studitsky, V. & Ninfa, A.J. Activation of the glnA, glnK, and nac promoters as *Escherichia coli* undergoes the transition from nitrogen excess growth to nitrogen starvation. *J. Bacteriol.* **184**, 5358–5363 (2002).
22. Reitzer, L. Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu. Rev. Microbiol.* **57**, 155–176 (2003).
23. Atsumi, S. Wu. T.-Y., Eckl, E.-M., Hawkins, S.D., Buelter, T. & Liao, J.C. Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Appl. Genet. Mol. Biotechnol.* **85**, 651–657 (2010).
24. Werpy, T. & Petersen, G., eds. *Top Value Added Chemicals From Biomass*, vol. I (Pacific Northwest National Laboratory, National Renewable Energy Laboratory and Office of Biomass Program for the Office of the Biomass Program of US Department of Energy, 2004). <www1.eere.energy.gov/biomass/pdfs/35523.pdf>
25. Wan, M.Y., Wang, H.Y., Zhang, Y.Z. & Feng, H. Substrate specificity and thermostability of the dehairing alkaline protease from *Bacillus pumilus*. *Appl. Biochem. Biotechnol.* **159**, 394–403 (2009).

ONLINE METHODS

Reagents. Restriction enzymes and Antarctic phosphatase were from New England Biolabs. KOD DNA polymerase was from EMD Chemicals. Rapid DNA ligation kit was from Roche. Yeast extract was from Becton, Dickinson. Amino acids, 2-ketobutyrate, 2-ketoisovalerate, 2-ketovalerate, 2-keto-3-methyl-valerate and 2-keto-4-methyl-pentanoate were from Sigma. Oligonucleotides were from IDT. Amino acid standard (0.25 nmol/μl) and *o*-phthaldialdehyde (OPA) were from Agilent Technologies.

Strains and plasmids. The JCL16 strain is a BW25113 ($rrnB_{T14}$ $\Delta lacZ_{WJ16}$ *hsdR514* $\Delta araBAD_{AH33}$ $\Delta rhaBAD_{LD78}$) derivative with F′ transduced from XL-1 blue to supply $lacI^q$. Twenty Keio collection strains[26] were used as shown in **Figure 1a**. Genome-wide random mutagenesis is done with *N*-methy-*N*'-nitro-*N*-nitrosoguanidine (NTG)[15]. After each round of the NTG treatment, mutations were screened for growth on an agar plate containing one individual nBCAA as sole carbon source. In addition, L-valine analog DL-norvaline, which is toxic to the cell partly due to its incorporation into polypeptides, was added (2 g/l). Some mutant strains could survive such a challenge through overproducing 2-keto isovalerate (precursor of valine and isobutanol) to outcompete the analog for polypeptide biosynthesis. Thirty single colonies were transformed with keto-acid pathway (AlsS, IlvCD, KivD and AdhA) and tested for isobutanol production. The best production strain was selected for the next round of NTG mutagenesis. YH19 was the final strain obtained after the NTG mutagenesis. YH19 derivatives with certain gene deletions (such as Δ*luxS*, Δ*lsrA*, Δ*lsrD*, Δ*lsrB*, Δ*glnA*, Δ*gdhA*, Δ*ilvE*) were created through P1 transduction or λ phage recombination. The strains containing both Δ*glnA* and Δ*gdhA* can use amino acids, but not ammonium salts as sole nitrogen source for growth. In yeast extract medium, the strains containing both Δ*glnA* and Δ*gdhA* grow at approximately one-third of the rate of the wild-type strain and stop growing after the medium's glutamine supply is depleted. Phages were prepared from the Keio collection[26]. Plasmid pCP20 was transformed into single colonies containing the correct deletions to remove the kanamycin-resistance marker. A list of the plasmids used is given in **Supplementary Table 5**. Construction of plasmids is described in **Supplementary Table 5**, and the primers used are listed in **Supplementary Table 6**. All constructions were verified by sequencing.

Medium and cultivation. Unless stated otherwise, 1× modified M9 salt (31.5 g/l $NaHPO_4$, 15 g/l KH_2PO_4, 2.5 g/l NaCl, 120 mg/l $MgSO_4$, 11 mg/l $CaCl_2$ and 10 mg/l vitamin B1 per liter water) containing 40 g/l BD bacto yeast extract (containing 21.64 g/l amino acids, 4.48 g/l ash, 3.05 g/l various salts, 1.24 g/l H_2O as well as 6.53 g/l carbohydrate, which is nondegradable by *E. coli*[27]) was used for all biofuel production assays. Ampicillin (100 μg/ml), kanamycin (50 μg/ml), chloramphenicol (30 μg/ml) and spectinomycin (50 μg/ml) were added as appropriate. Pre-culture in test tubes containing 3 ml of medium was performed at 37 °C overnight on a rotary shaker (250 r.p.m.). Overnight culture was diluted 1:100 into 20 ml of fresh medium in a 250-ml screw-cap conical flask. Cells were grown 2 h at 37 °C before adding 0.1 mM isopropyl-β-D-thiogalactoside (IPTG). Cultivation was performed at 37 °C on a rotary shaker (250 r.p.m.). In certain cases, two-phase fermentation was performed as described[15] to reduce the products' toxicity effect to the *E. coli* cells but no improvement of biofuel titer was observed. To obtain the biomass used in **Figure 3b**, the *E. coli* and *B. subtilis* were grown in LB medium. The *Chlorella vulgaris* (ATCC 13482^T), *Porphyridium purpureum* (ATCC 50161^T), *Spirulina platensis* (UTEX LB2340) and *Synechococcus elongatus* PCC7942 were grown in ATCC medium 5, ATCC medium 1495, UTEX spirulina medium and BG-11 medium, respectively. After harvest, some biomass was treated by a mini beadbeater for 1 min or 0.5 N hot NaOH for 30 min to release the proteins for a concentration measurement through Bradford assay. The rest of protein biomass was hydrolyzed by heating in 60 or 80–100 °C water for 10–20 min followed by an overnight protease hydrolysis at 50 °C. The amount of protease was 1–3% of the biomass's dry weight (0.3–0.9 mg/ml). The concentrations of free amine groups before and after the protease treatment were measured by Ninhydrin Assay Kit (Sigma). All protease-treated protein biomass were then filtered for medium preparation. Gas chromatography–mass spectrometry (GC–MS), gas chromatography–flame ionization detector (GC–FID) and high-performance liquid chromatography (HPLC) were used to analyze the substrates and products.

Enzyme assays. For crude extract enzyme assays, YH19 strain with or without the relevant plasmids were grown at 30 °C in 10 ml 4% yeast extract media 12 h after IPTG induction. Crude extracts were prepared by concentrating the cultures by tenfold in 0.1 M phosphate buffer (pH 7.1) and lysing them with 0.1 mm glass beads. Total protein concentrations were measured by Bradford assays. The YH19 without gene overexpression was used as the background control. For LeuDH activity, the YH19 strain was transformed with pYX51 and the assay was performed as described in ref. 28. For AlsS, IlvC and IlvD activities, the YH19 strain was transformed with pYX90 and the assays were performed as described in ref. 29. For KivD and *E. coli* alcohol dehydrogenase YqhD activities, the YH19 strain was transformed with pYX97 and the assays were performed as described in refs. 23 and 30, respectively.

GC analysis. Alcohol compounds produced by our strains were identified by GC–MS and quantified by GC-FID, as described previously[9,15].

HPLC analysis. Amino acids (except L-proline) were quantified using ZORBAX Eclipse AAA column (Agilent Technologies) with OPA (*o*-phthaldialdehyde) derivatization method. L-Proline was derivatized with FMOC (9-fluorenylmethyl chloroformate). Derivatized amino acids were analyzed using PDA detector (338 nm, 262 nm (L-Pro)) of HPLC.

Theoretical yield calculation. Detailed computational methods for **Supplementary Figure 9** and **Supplementary Tables 2–4** are described in the corresponding legends. The amino acid degradation pathways and the higher alcohol biosynthesis pathways are summarized in **Supplementary Table 7** and **Supplementary Figure 6**.

26. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
27. BD Biosciences. *BD Bionutrients Technical Manual: Advanced Bioprocessing* (Becton, Dickinson & Co., 2006). <http://www.bd.com/ds/technicalCenter/misc/br_3_2547.pdf>
28. Ohshima, T. *et al.* The purification, characterization, cloning and sequencing of the gene for a halostable and thermostable leucine dehydrogenase from *Thermoactinomyces intermedius*. *Eur. J. Biochem.* **222**, 305–312 (1994).
29. Atsumi, S., Higashide, W. & Liao, J.C. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.* **27**, 1177–1180 (2009).
30. Smith, K.M., Cho, K.M. & Liao, J.C. Engineering *Corynebacterium glutamicum* for isobutanol production. *Appl. Genet. Mol. Biotechnol.* **87**, 1045–1055 (2010).

doi:10.1038/nbt.1789