PNAS
www.pnas.org

Supplementary Information for

Redox controls metabolic robustness in the gas-fermenting acetogen *Clostridium autoethanogenum*

Vishnuvardhan Mahamkali, Kaspar Valgepea, Renato de Souza Pinto Lemgruber, Manuel Plan, Ryan Tappel, Michael Köpke, Séan Dennis Simpson, Lars Keld Nielsen, and Esteban Marcellin

Corresponding author: Esteban Marcellin
Email: e.marcellin@uq.edu.au

This PDF file includes:

Supplementary text
Figures S1 to S6
Table S1
SI References

Other supplementary materials for this manuscript include the following:

Dataset S1

Supplementary Information Text

Material and Methods

Thermodynamic metabolic flux analysis (tMFA)

Incorporating component contribution method in tMFA to estimate thermodynamically consistent flux profiles

To calculate Gibbs free energy ranges ($\Delta_r G'_{min}$, $\Delta_r G'_{max}$) for a metabolic reaction, it is necessary to calculate the standard Gibbs free energy values ($\Delta_r G°$) and the corresponding metabolite concentrations (1). We used component contribution method (2) for calculating the standard transformed Gibbs free energies of formation of metabolites and the respective standard deviations at given pH and ionic strength. The problem formulation is represented as:

$$\Delta_r \mathrm{G}' = N^T \Delta_f G'_{cc} + RT\, N^T \ln(x) + \Delta_r G_{transport}$$
$$N.V = 0$$
$$v_i - v_{max} * z_i < 0$$
$$\Delta_r G'_i - \mathrm{K} + \mathrm{K} * z_i < 0$$

, where $\Delta_f G'_{cc}$ is the standard transformed Gibbs free energy of a compound varying between two respective standard deviations, x is activity of metabolites.

Constraints on the tMFA model

We calculated the Gibbs free energy change for acetate transport (ΔG_transport) and pH changes (ΔG_pH) as described previously (1, 3), while assuming a cytosolic pH of 6 (4) and ionic strength of 0.1 M. Our tMFA model was further constrained with experimentally measured metabolite concentrations of syngas-grown (CO, H_2 and CO_2) CO-limited steady-state chemostat cultures of *C. autoethanogenum* (5). For non-gaseous metabolites, expect ferredoxins (see below), whose intracellular concentrations were not determined, the minimum and maximum metabolite activities were set to 0.01 mM and 20 mM (1), respectively. The minimum activities for dissolved gases (H_2, CO_2, and CO) were set to 0.00001 mM, and the maximum activities were set to the saturation constants at standard conditions, i.e., 0.79, 34.45, and 0.91 mM, respectively. Thermodynamic variability analysis (TVA) was used to obtain the Gibbs free energy ranges and the metabolite activity ranges.

Bounds on ferredoxin concentrations

In addition to the "standard" redox pairs of NAD(P)H/NAD(P), ferredoxins (reduced and oxidised) are crucial metabolites in acetogens (e.g. *Clostridium autoethanogenum*) as they are involved in several key reactions and complexes such as Aldehyde:Fd oxidoreductase (AOR), Pyruvate:ferredoxin oxidoreductase (PFOR), Nfn, and Rnf complexes. However, we are not aware of any study where ferredoxin ratios have been experimentally determined. Thus, we placed relaxed bounds on both reduced and oxidized ferredoxin concentrations and calculated their ranges using TVA. It was observed that the mean of the calculated range for the total ferredoxin concentration pool and for the ratio of oxidised to reduced ferredoxin differed maximally only by ~6% within the CO-limited steady-state chemostat culture data used for constraining our tMFA model (5). Hence, ferredoxin species seem to have a weak quantitative effect on our thermodynamic analysis.

Based on the latter results, new constraints on the ferredoxin pool and the ratio of oxidized to reduced ferredoxin were introduced as follows:

$$Ferredoxin_{reduced} + Ferredoxin_{Oxidized} = c$$
$$\frac{Ferredoxin_{reduced}}{Ferredoxin_{Oxidized}} = d$$

, where c and d are the pool and ratio constraints of ferredoxin, respectively.

Finally, TVA was performed with the thermodynamic, metabolite concentration, and ferredoxin constraints to calculate reaction directionalities.

Mechanisms of acetate transport through the cell membrane

To investigate the effect of different mechanisms of acetate transport on thermodynamics of the system, the standard Gibbs free energies of acetate transport ($\Delta G^{o}_{acetate}$) for each mechanism were calculated. We considered four mechanisms for acetate transport that each require different concentration gradients (Fig. 1B): transport of the undissociated acid (*i.e.* passive diffusion); transport of anion *via* uniport; symport of the anion with a proton; and ATP-consuming transport (*i.e.* ABC-type).

The $\Delta G^{o}_{acetate}$ depends on the electrochemical potential ($\Delta\psi$), the pH gradient and the concentration gradient of acetate across the cell membrane according to:

$$\Delta G^{o}_{acetate} = \Delta G^{o}_{\Delta\psi} + \Delta G^{o}_{pH} + \Delta G^{o}_{concentration}$$

$$\Delta G^{o}_{\Delta\psi} = cF\Delta\psi$$

, where c is the net charge transported from outside the cell into the cell, and F is the Faraday constant in kcal/mV mol; $\Delta\psi$ is assumed to be -100 mV based on *Clostridium acetobutylicum* membrane potential (6).

and

$$\Delta G^{o}_{pH} = 2.303hRT\Delta pH$$

, where h is the number of protons transported, R is the universal gas constant and T is the temperature in K.

and

$$\Delta G^{o}_{concentration} = RT \ln \frac{A^{out}}{A^{in}}$$

, where A^{out} and A^{in} are acetate extracellular and intracellular concentrations, respectively.

For our analysis, we used the extracellular acetate concentrations of the same cultures (5) to constrain the acetate gradient across the membrane. Intracellular acetate concentrations were estimated using the Henry's law and based on the mechanism of transport. The tMFA model was then constrained with intracellular acetate concentrations and the Gibbs free energy of transport of respective transport mechanism while leaving bounds on extracellular acetate concentrations unrestricted.

Our analysis yielded thermodynamically feasible solutions only for uniport (see main text). Since the energy for acetate transport might be penalized from the transmembrane ATP synthase, the cost of transport of one molecule of acetate would be 0.25 ATP because the ATP synthase requires 4 protons to make one ATP molecule. The intracellular acetate concentration for uniport was calculated using the following equation:

$$\Delta G = -Z \log\left(\frac{A_{in}}{A_{out}}\right) + \Delta\psi$$

LC-MS/MS analysis for intracellular metabolomics

Sample details

Firstly, samples were randomised for analysis to reduce batch effect-based bias. Secondly, samples were flanked with repeat injections of standards and pooled QC samples to monitor instrument stability and ensure data integrity (7). Finally, 5 µL of samples were analysed at different levels of concentration (5x, 10x, and 20x) compared to concentrations in the initial samples from chemostat cultures to assure that measured intensities would fall within the range of the standard curves.

Instrumentation and acquisition details

Targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS) metabolomics analyses were performed using a Shimadzu UHPLC System coupled to a Shimadzu 8060 triple quadrupole (QqQ) mass spectrometer.

The UHPLC (Nexera X2, Shimadzu Corp., Kyoto, Japan) consisted of LC-30AD pump units, DGU-20ASR degassing units, a SIL-30AC autosampler, a CTO-20AC column oven, a CBM-20A communications BUS module, and an FCV-20AH2 diverter valve unit. Liquid chromatography was performed using a guard column (SecurityGuard Gemini-NX C18, 4 x 2 mm, PN: AJO-8367,

Phenomenex) and a Shim-pack Velox SP-C18 UHPLC column (1.8 um, 2.1 x 150 mm, PN: 227-32001-04, Shimadzu Corp.) operated at 35 °C. An ion-pairing buffer system (8) of 7.5 mM tributylamine (PN: 471313, Sigma-Aldrich) adjusted to pH 4.95 using glacial acetic acid (PN: 27225, Fluka) as Solvent A and acetonitrile (Lichrosolv, PN: 1142914000, Merck) as Solvent B were used at a flow rate of 400 µL/min the following gradient chromatography: Solvent B 1% over 0.01-1.00 min; 1-50% over 1.00-15.00 min; 50-70% over 15.00-16.00 min; 70-98% over 16.00-17.00 min; 98% over 17.00-22.00 min; 98%-1% over 22.00-22.10 min; and 1% over 22.10-25.00 min.

The Shimadzu 8060 QqQ system had an electrospray ion source (ESI), and used nitrogen (>99.999 vol %, BOC Australia) as the drying gas and argon (>99.999 vol %, Coregas Pty Ltd) as the collision gas. Further instrument details include: drying gas flow of 10 L/min; nebulising gas flow of 3.0 L/min; heating gas flow of 10 L/min; desolvation line of 250 °C; heat block temperature of 400 °C; CID of gas at 270 kPa; and interface temperature of 300 °C. Interface potential was optimised by performing scheduled multiple reaction monitoring (sMRM) experiments on the standard mix (sample of a mix of analytical standards) at -0.5, -0.75, -1.0, -2.0, -3.0 kV to determine peak response and to lower the limit of detection. Data were collected from 0-18 min and diverted to waste from 18-25 min during column clean up and re-equilibration.

sMRM transients were optimised on negative ionisation mode (m/z –H) for 11 compounds including: acetyl-phosphate, acetolactate, acetyl-CoA (Acetyl-CoA), 3-hydoxybutyric acid (3HB), flavin adenine dinucleotide (FAD), flavin adenine mononucleotide (FMN), nicotinamide adenine dinucleotide (NAD), reduced NAD (NADH), nicotinamide adenine dinucleotide phosphate (NADP), reduced NADP (NADPH), and azidothymidine (AZT). All analytical standards were purchase from Sigma-Aldrich.

Data analysis

Quantification of metabolite concentrations was based on calibration curves from 20 concentration of 1-in-2 serial dilutions (L1 to L20) of the standard mix where L1 = 250 µM, L2 = 125, L3 = 62.5…, and L20 = 0.000477 µM (or 0.477 nM). Depending on the limit of detection for each compound, standard curves consisted of at least five data points. Data were processed using the software LabSolutions Insight Version 3.2 SP1 and LabSolution Postrun/QuantBrowser Version 5.95.

Proteomics

Sample details

High-resolution temporal sampling for proteomics was carried out throughout the oscillations. For this, 2 mL of culture was pelleted by immediate centrifugation (25,000 × *g* for 1 min at 4 °C) and stored at -80 °C until analysis. For comparing protein expression between two extreme states of the culture—just prior to recovery and crashing—, samples at lowest (*i.e.* recovery; bioreplicate #1 at 214 and 383 h, bioreplicate #2 at 252 and 407 h) and highest (*i.e.* crash; bioreplicate #1 at 347 and 477 h, bioreplicate #2 at 347 and 503 h) biomass concentration values (eight in total) were used for further analysis (see Fig. 2 or 3).

Sample preparation was performed as described before (9) with the following modifications: 1) thawed and washed pellets were resuspended in 450 µL of lysis buffer; 2) protein concentration in cell lysates was determined using the Direct Detect® Infrared Spectrometer (DDHW00010-WW; Merck); 3) 50 µg of protein was used for protein digestion; 4) no desalting was performed; 5) 5 µg of peptide material (~11-17 µL) from protein digestion was concentrated using a vacuum centrifuge (Concentrator Plus; Eppendorf) at 30 °C for 60 min and reconstituted in 15 µL of 0.1% formic acid in 5% acetonitrile.

LC method

For generating the spectral library using data dependent acquisition (DDA) and for the data independent acquisition (DIA) sample runs, a ThermoFisher Scientific UltiMate 3000 RSLCnano UHPLC system was used. Each sample was initially injected onto a ThermoFisher Acclaim PepMap C_{18} trap reversed-phase column (300 µm x 5 mm nano viper, 5 µm particle size) at a flow rate of 15 µL/min using 2% acetonitrile (ACN) for 5 min with the solvent going to waste. The trap column was switched in-line with the separation column (ThermoFisher EasySpray Pepmap RSLC C_{18}, 150 µm x 150 mm, 2 µm) and the peptides were eluted using a flowrate of 1.5 µL/min using 0.1% FA in water (buffer A) and 80% ACN in buffer A (buffer B) as mobile phases for gradient elution.

Peptide elution employed a 3-30% ACN gradient for 63 min followed by 30-50% ACN for 10 min and 50-95% ACN for 1 min at 40 °C. The total elution time was 110 min including a 95% ACN wash and a re-equilibration step. For each DDA and DIA sample run, a volume of 3 µL equating to 1 µg of peptide material from protein digestion was injected.

DDA MS spectral library generation

Two samples of one biological replicate culture (bioreplicate #1 "recovery sample" at 214 h and "crash sample" at 347 h) were analysed on the Thermo Fisher Scientific Q-Exactive HF-X mass spectrometer in DDA mode to yield the spectral library for DIA MS data analysis.

The eluted peptides from the C_{18} column were introduced to the MS *via* a nano-ESI and analysed using the Q-Exactive HF-X. The electrospray voltage was 1.8 kV in positive ion mode, and the ion transfer tube temperature was 250 °C. Employing a top-20 ddMS2 acquisition method, full MS-scans were acquired in the Orbitrap mass analyzer over the range m/z 400–1500 with a mass resolution of 120,000 (at m/z 200). The AGC target value was set at 3.00E+06. The 20 most intense peaks with a charge state between 2 and 6 were fragmented in the high energy collision dissociation (HCD) cell with a normalised collision energy (NCE) of 30. MSMS spectra were acquired in the Orbitrap mass analyzer with a mass resolution of 30,000 at m/z 200. The AGC target value for MSMS was set to 1.0E+05 while the ion selection threshold was set to 5E+03 counts. The maximum accumulation times were 50 ms for full MS-scans and MSMS. For all the experiments, the dynamic exclusion time was set to 25 s, and undetermined charge state species were excluded from MSMS.

Identification results of DDA analysis for subsequent creation of a spectral library for DIA MS data confirmation and quantification using the software Skyline (10) were generated using the Proteome Discoverer 2.2 software (Thermo Fisher Scientific) and its SEQUEST HT search as described before (9). The final .pd result file contained peptide-spectrum matches (PSMs) with q-values estimated at 1% FDR for peptides ≥4 AAs.

DIA MS data acquisition

As for the DDA method, eluted peptides were introduced to the MS *via* a nano-ESI and analysed using the Q-Exactive HF-X. The electrospray voltage was 1.8 kV, and the ion transfer tube temperature was 250 °C. Full MS-scans were acquired in the Orbitrap mass analyser over the range m/z 400–1200 with a mass resolution of 30,000 (at m/z 200). The AGC target value was set at 3.00E+06 with a maximum injection time of 60 ms. DIA was achieved using an inclusion list: m/z 395–1100 in steps of 15 amu and scans cycled through the list of 48 isolation windows with a loop count of 48. All DIA scans were recorded with an NCE collision energy of 30 and MSMS detection at a resolution setting of 15,000 (at m/z 200). The AGC target was set to 1.0E+06 with a maximum injection time set at 'auto'. A first fixed mass of m/z 200 was applied, and default charge state of 2 was set for scanning MS2 events.

DIA MS data analysis

DIA MS data analysis was performed with Skyline (10) as described before (9) with the exception of using 12 manually picked endogenous peptides for iRT alignment. Shortly, the .pd result file from Proteome Discoverer was used to build the DIA MS spectral library and the mProphet peak picking algorithm (11) within Skyline was used to separate true from false positive peak groups (per sample) and only peak groups with q-value<0.01 (representing 1% FDR) were used for further quantification. We confidently quantitated 3,595 peptides and 596 proteins on average within each sample with at least two peptides per protein.

Analysis of differential protein expression

Protein expression fold changes with p- and q-values were determined using the software MSstats (12) as described before (9). Proteins were considered to be differentially expressed by a fold-change>1.5 and a q-value<0.05 after FDR correction (13).

Data availability

Proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD016381.

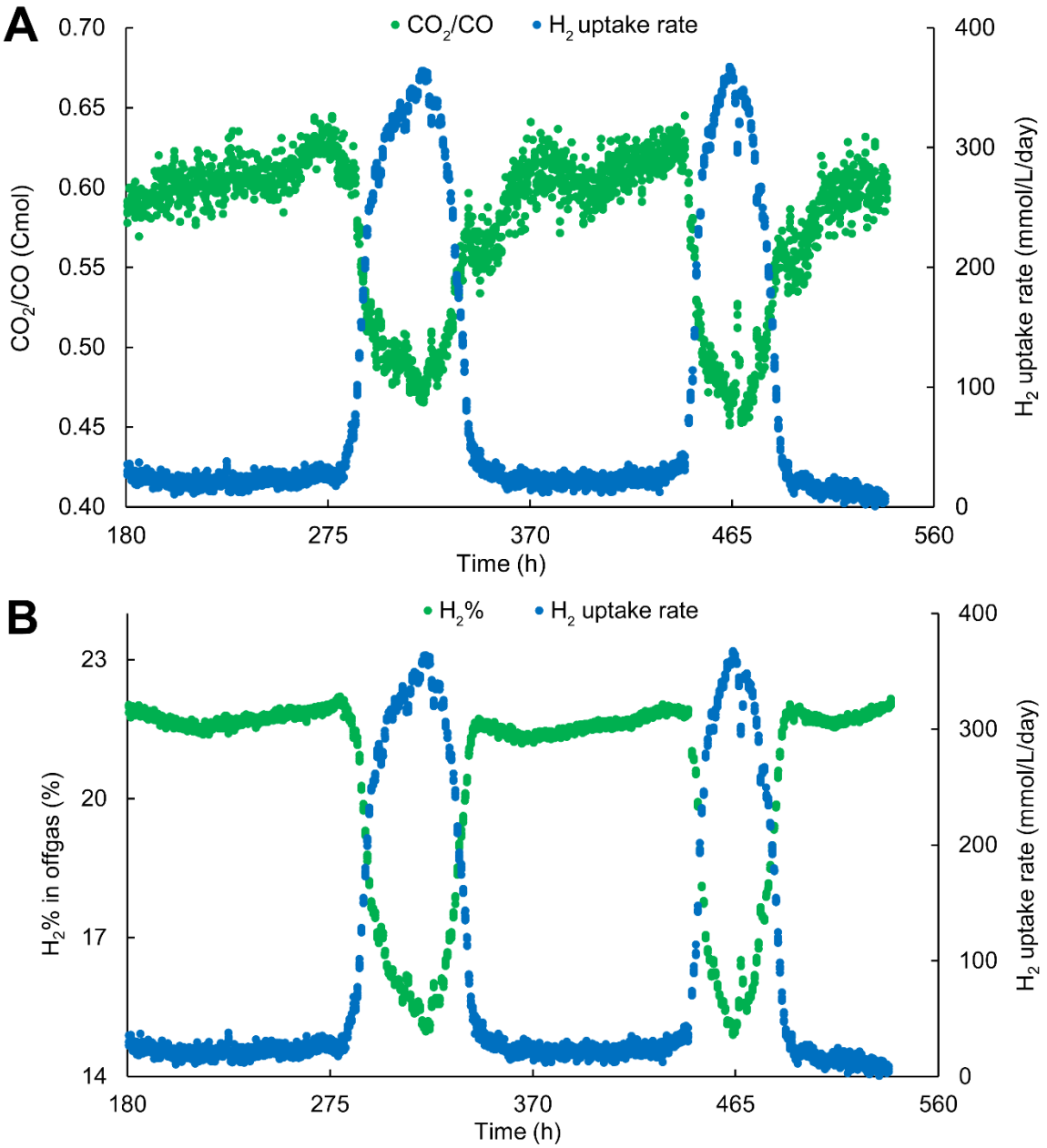

Fig. S1. Balancing supply of reduced ferredoxin (Fd_{red}) between H_2 uptake and CO_2 dissipation in *Clostridium autoethanogenum* continuous cultures. (*A*) Temporal profiles of H_2 uptake and CO_2/CO. (*B*) Temporal profiles of H_2 uptake and H_2% in bioreactor offgas. Data for one bioreplicate culture is shown. Time, fermentation duration from culture inoculation; CO_2/CO, CO_2 production rate (Cmmol/L/day) divided by CO uptake rate (Cmmol/L/day).

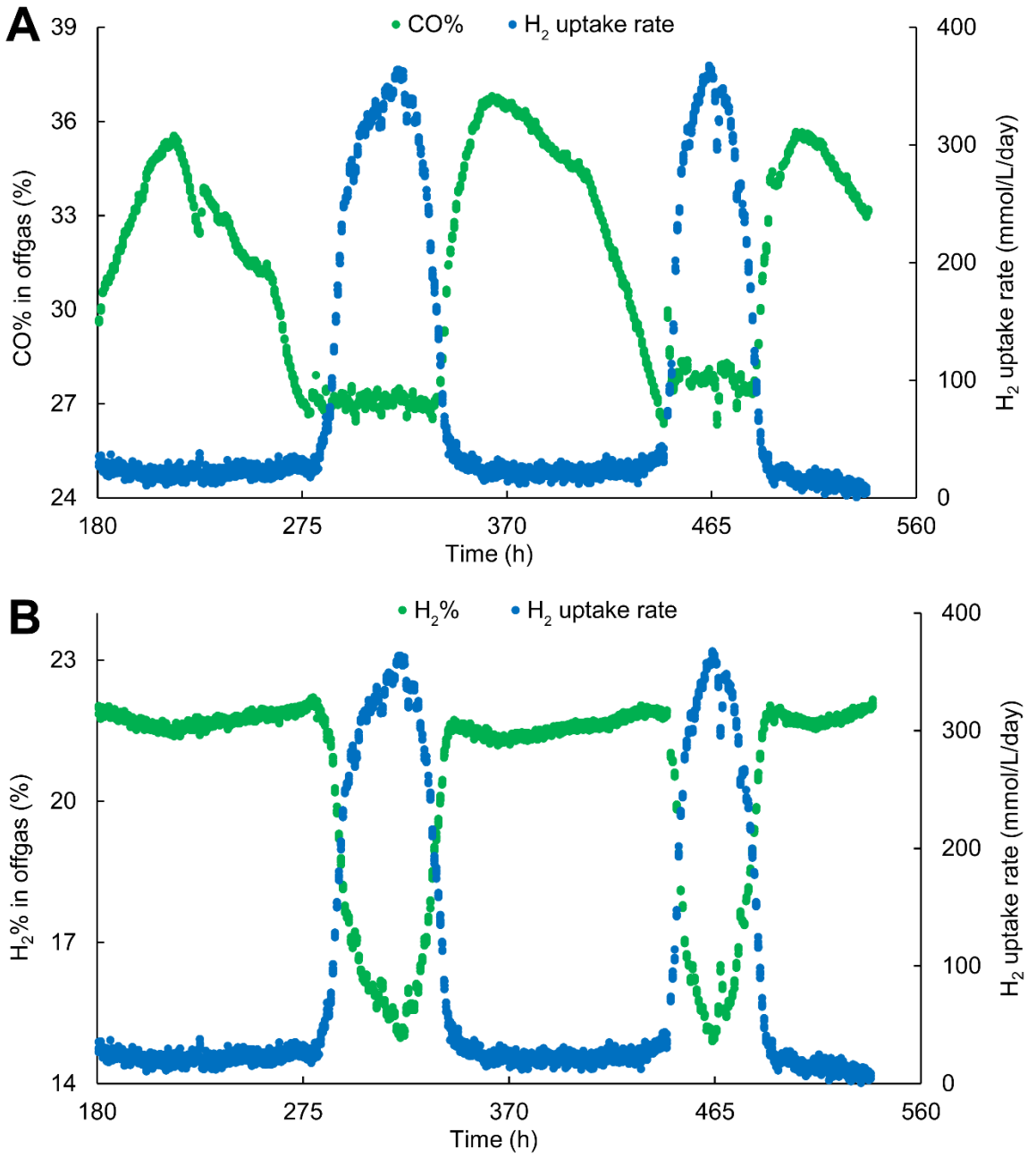

Fig. S2. Recovery of H_2 uptake is potentially triggered by the *Clostridium autoethanogenum* continuous culture becoming CO-limited. (*A*) Temporal profiles of H_2 uptake and CO% in bioreactor offgas. (*B*) Temporal profiles of H_2 uptake and H_2% in bioreactor offgas. Data for one bioreplicate culture is shown. Time, fermentation duration from culture inoculation.

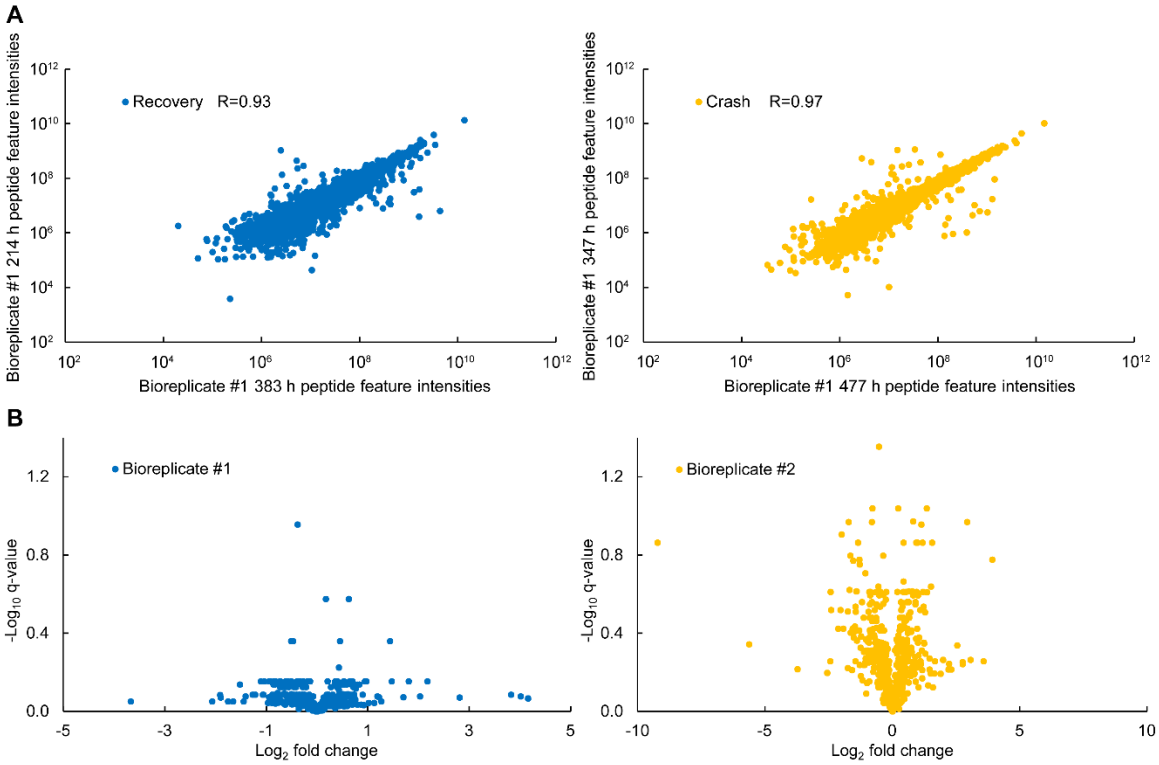

Fig. S3. Quantitative proteomics of oscillating continuous cultures of *Clostridium autoethanogenum*. (*A*) Correlation between confidently quantified (q-value<0.01) peptide feature MS intensities of two samples with either lowest (*i.e.* recovery; left plot) or highest (*i.e.* crash; right plot) biomass concentration in bioreplicate culture #1 (see Fig. 1 or 2). R, Pearson correlation coefficient. (*B*) Volcano plots showing no proteins being significantly differentially expressed between samples with lowest (*i.e.* recovery) and highest (*i.e.* crash) biomass concentration within each bioreplicate culture. Significant differential expression was defined as fold-change>1.5, *i.e.* Log_2 fold-change>~0.6; q-value<0.05, *i.e.* $-Log_{10}$ q-value>~1.3 after false discovery rate [FDR] correction (13).

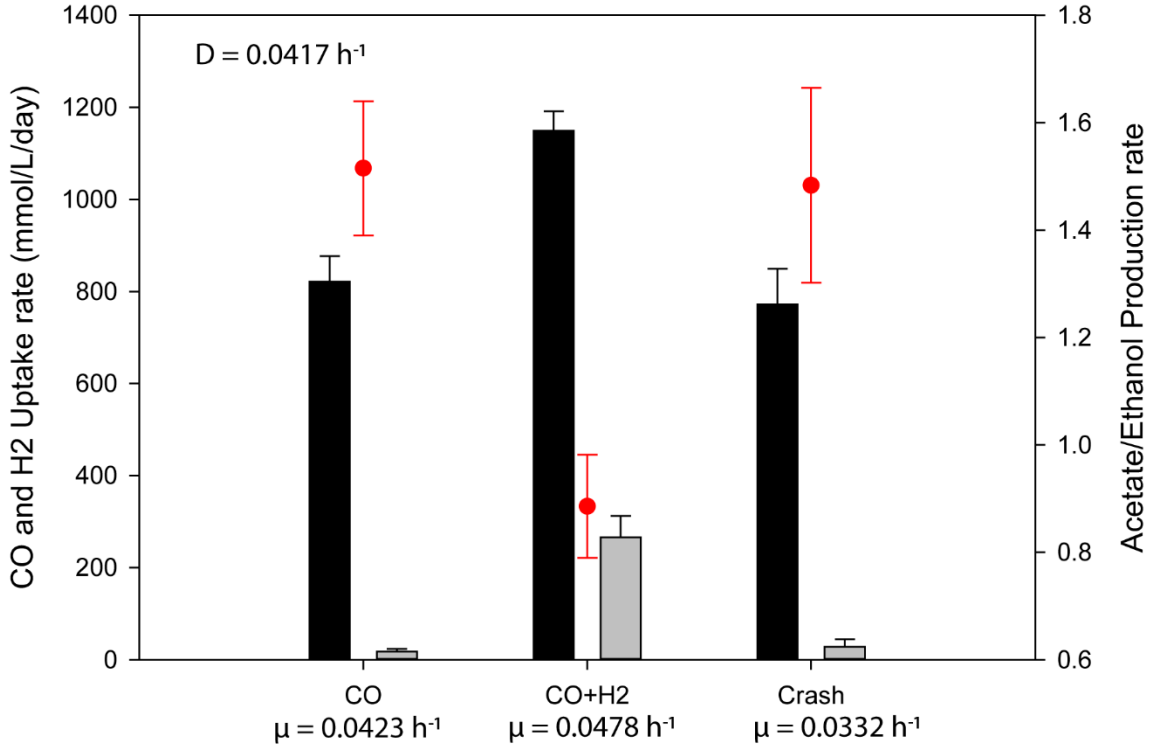

Fig. S4. CO and H_2 uptake rates across three phases of growth; 1) Growth only on CO 2) Growth on CO and H_2 3) Crash; (bars). Acetate/Ethanol production ratios in above three growth phases are shown (red circles). Specific growth rates across three growth phases are shown as well (Reactor dilution rate is maintained at 0.0417 h^{-1}). To calculate the growth rates and production rates over different growth phases on unequally spaced data, we have used Lagrange polynomial.

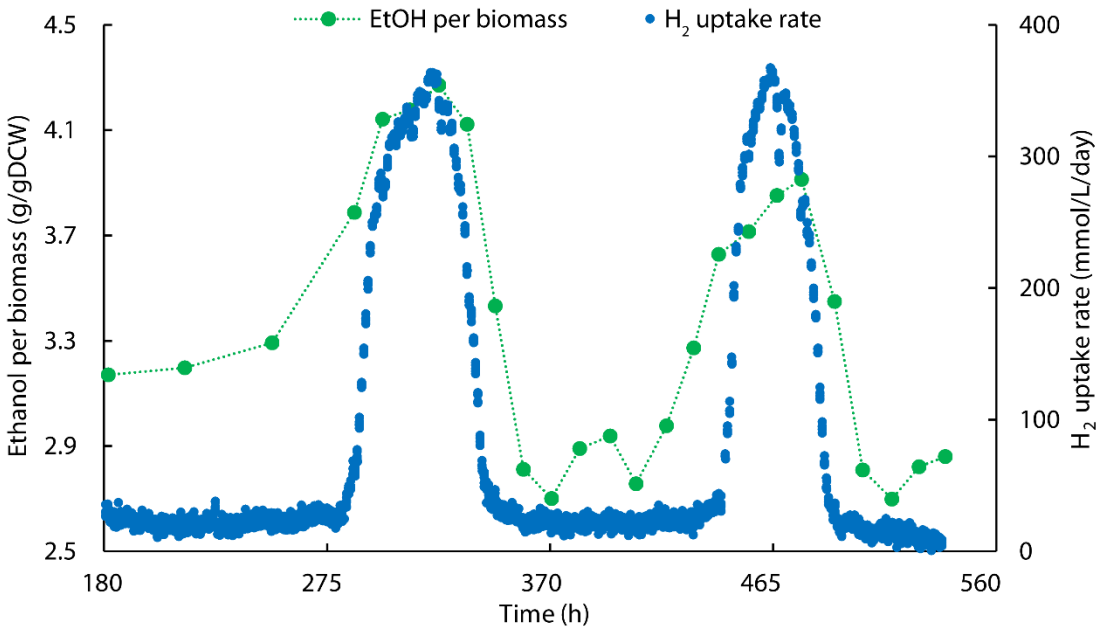

Fig. S5. H_2 uptake leads to a higher ethanol yield per biomass in *Clostridium autoethanogenum* continuous cultures. Data for one bioreplicate culture is shown. Time, fermentation duration from culture inoculation; gDCW, grams of dry cell weight; Ethanol per biomass, ethanol concentration (g/L) divided by biomass concentration (gDCW/L).

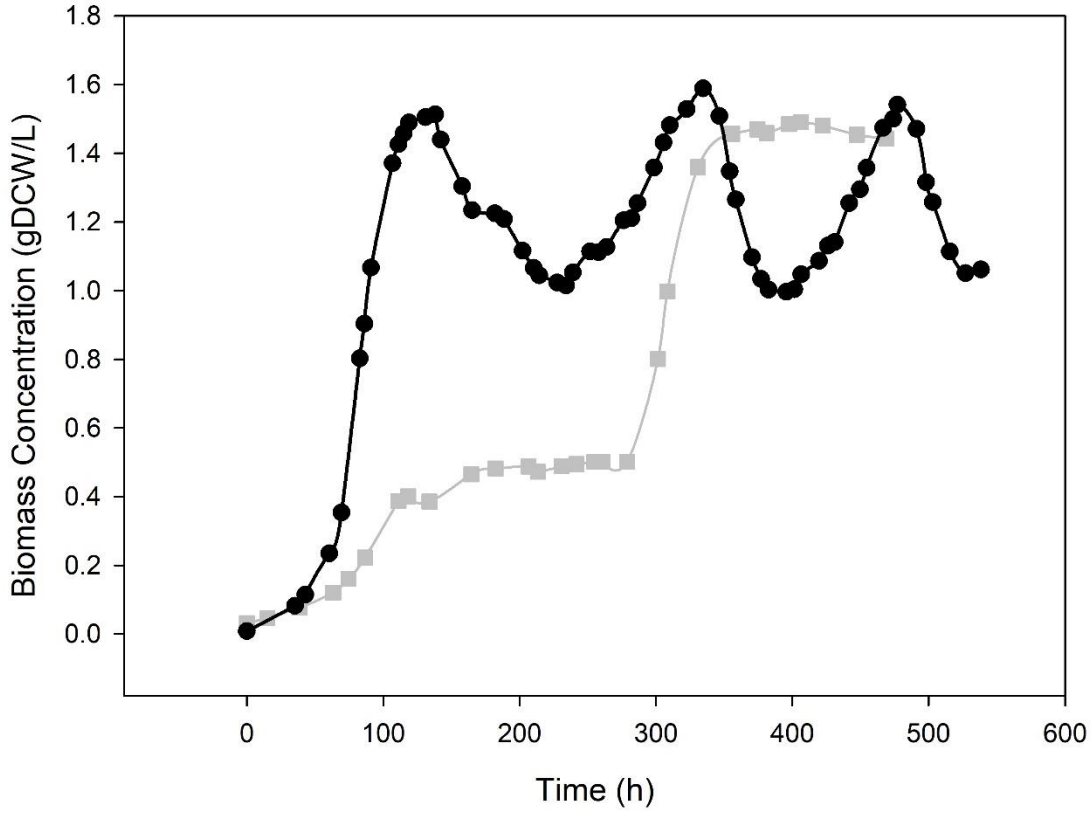

Fig. S6. Biomass concentration displaying oscillations in this work (black circles; feed 50% CO, ~20% H_2, ~20% CO_2, and ~10% Ar) and a steady state culture on CO only in a previous work (9) (grey squares; 60% CO, 40% Ar). As illustrated by the CO culture, a steady state can be reached on pure CO at a comparable biomass concentration. For clarity, only data for one biological replicate is shown. Time, fermentation duration from culture inoculation; gDCW, grams of dry cell weight.

Reaction	CO		H2		Crash	
	Min (kJ/mol)	Max (kJ/mol)	Min (kJ/mol)	Max (kJ/mol)	Min (kJ/mol)	Max (kJ/mol)
W1	-34.533	0.000	-37.967	0.000	-25.611	0.000
W2	-71.124	0.000	-70.631	0.000	-60.509	0.000
W4	-46.500	0.000	-54.488	0.000	-40.487	0.000
W5	-19.355	0.000	-19.355	0.000	-19.355	0.000
W6	-33.744	0.000	-33.250	0.000	-35.626	0.000
W7	-89.292	-20.099	-67.661	-3.807	-66.768	-3.391
W1-Acoa_pre	-85.076	0.000	-89.675	0.000	-80.723	0.000
Hydrogenase	-24.484	0.000	-26.878	0.000	-7.006	0.000
Ac1-Pta	-3.530	0.000	-3.739	0.000	-2.135	0.000
Ac2-Ack	-30.083	0.000	-38.963	0.000	-25.203	0.000
E1-AdhE	-26.499	-1.428	-0.477	19.463	0.124	14.021
E2-AOR	-24.484	0.000	-26.878	0.000	-7.006	0.000
E3-Adh	-39.177	-27.023	-15.549	-9.727	-15.687	-9.311
P1-PFOR	-39.655	25.107	-35.265	32.866	-27.066	27.304
B1-AlsS	-91.500	16.185	-89.044	0.187	-57.871	-3.344
B2-BudA	-74.150	2.418	-58.121	-0.051	-58.117	-0.218
B3-Bdh	-61.097	-3.664	-39.466	12.628	-39.046	13.044
Rnf	-30.909	4.021	-46.161	-14.176	-31.405	-19.296
Nfn	-77.735	-35.858	-52.984	-17.119	-38.950	-16.727
ATP_synthase	-25.036	0.000	-25.038	0.000	-25.017	0.000

Table. S1. Gibbs free energy change ranges predicted by Thermodynamic metabolic flux balance analysis across three growth phases described in the manuscript. The reaction driving force is visibly smaller in crash conditions compared to other two growth phases.

References

1. C. S. Henry, L. J. Broadbelt, V. Hatzimanikatis, Thermodynamics-based metabolic flux analysis. *Biophys. J.* **92**, 1792–1805 (2007).
2. E. Noor, H. S. Haraldsdóttir, R. Milo, R. M. T. Fleming, Consistent estimation of Gibbs energy using component contributions. *PLoS Comput. Biol.* **9** (2013).
3. S. J. Jol, A. Kümmel, V. Hatzimanikatis, D. A. Beard, M. Heinemann, Thermodynamic calculations for biochemical transport and reaction processes in metabolic networks. *Biophys. J.* **99**, 3139–3144 (2010).
4. J. Mock, *et al.*, Energy conservation associated with ethanol formation from H_2 and CO_2 in *Clostridium autoethanogenum* involving electron bifurcation. *J. Bacteriol.* **197**, 2965–2980 (2015).
5. K. Valgepea, *et al.*, Maintenance of ATP homeostasis triggers metabolic shifts in gas-fermenting acetogens. *Cell Syst.* **4**, 505–515.e5 (2017).
6. L. Huang, L. N. Gibbins, C. W. Forsberg, Transmembrane pH gradient and membrane potential in *Clostridium acetobutylicum* during growth under acetogenic and solventogenic conditions. *Appl. Environ. Microbiol.* **50**, 1043–1047 (1985).
7. T. Sangster, H. Major, R. Plumb, A. J. Wilson, I. D. Wilson, A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. *Analyst* **131**, 1075 (2006).
8. T. S. McDonald, K. N. Tan, M. P. Hodson, K. Borges, Alterations of hippocampal glucose metabolism by even versus uneven medium chain triglycerides. *J. Cereb. Blood Flow Metab.* **34**, 153–160 (2014).
9. K. Valgepea, *et al.*, H_2 drives metabolic rearrangements in gas-fermenting *Clostridium autoethanogenum*. *Biotechnol. Biofuels* **11**, 55 (2018).
10. L. C. Gillet, *et al.*, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **11**, O111.016717 (2012).
11. L. Reiter, *et al.*, mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* **8**, 430–435 (2011).
12. M. Choi, *et al.*, MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).
13. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).