**Replies to reviewers’ comments**

**Reviewer 1**

The study combines theoretical metabolic network analysis with thermodynamics and metabolomics to predict flux limiting reactions in an anaerobic bacterial system. The goal of the approach is to identify rational intervention targets to increase flux from cellobiose to ethanol. The study extends previous experimental and theoretical work by the Ctherm group which previously considered the thermodynamics of the pyruvate to ethanol portion of glycolysis based on metabolite levels.

The study is interesting and combines a range of expertise to tackle a complex problem. The computational approach is useful and will be of interest to a wide range of quantitative biologists and engineers. The material is publishable but only after some substantial adjustments are made.

Major:

The current work does not account for synthesis of byproducts acetate, formate and lactate etc. According to supp file 5, the cultures produce more carbon moles of these byproducts than ethanol so leaving them out of the analysis introduces a very large source of error, additionally the model doesn’t consider biomass although this would be a small flux compared to the byproducts. The authors make an argument (line416+) they are not interested in the byproducts because they aren’t important with high ethanol selectivity strains (Olson, 2015, 2017). This isn’t a good argument because the current study uses metabolite concentrations from a strain that makes byproducts and therefore the data should not be projected onto strains that behave differently. This really needs to be addressed in a more satisfactory manner.

The thermodynamic analysis is based on the assumption that the metabolite concentration experienced by any reaction is only dependent on the available metabolite pool which is homogeneous throughout the cell. Thus, the metabolite concentrations will not vary if applied to a smaller core model or imposed on a genome-scale model. Our objective in this study was to evaluate the thermodynamics of ethanol production pathway for the given set of metabolite concentrations.

However, we also re-analyzed the thermodynamics of wild-type pathway by adding acetate production pathway and did not observe any change in the model predictions. This was because the acetate production was not thermodynamically constrained by the measured metabolite pools. Thus, in the absence of additional metabolite measurements any additional reaction would be unconstrained and hence always feasible.

Byproduct synthesis and growth might facilitate the problems with the infeasible 3pg concentrations calculated here.

If any given reaction is infeasible in our workflow, it implies that all possible variations of its constituent metabolite pools is unable to make this reaction feasible. Addition of alternate pathways or reactions can constrain the search space of the participating metabolites and make the pathway infeasible. The only ways to increase thermodynamic favorability of the pathway would be to either relax the metabolite pool constraints or remove the bottleneck reactions. Thus, addition of biomass production pathways will not resolve the 3pg infeasibility problem.

The authors don’t fix the cofactor pools for simulations and instead allow them to fluctuate within a preset range. An approach that is reasonable when lacking data, but it is unclear what data is lacking. Have the authors compared the predicted cofactor levels to the cofactor data reported in Tian 2017? Looking at the NADH and NAD+ values in the simulation data (supp fig 2), the ratio is NADH/NAD+ = 1/100, this is two orders of magnitude smaller than the experimental values in Tian 2017 (~1/1 ratio); the simulation for NADPH/NADP = 1/25 while the experimental ratio reported in Tian 2017 data is 15/1, again two orders of magnitude different. There seems to be some major issues with the analysis aligning with experimental data or there is confusion as to what experimental data is relevant.

<Answer>

The supplementary figure 2 shows the experimentally quantified values of various metabolite pools. In Tian et al. 2017b, the metabolite data comes from two different experiments, with slightly different experimental conditions. The data shown in supplementary figure 2 and supplementary file 1 was determined as the better dataset due to its high ATP charge values and hence used in our analysis.

Line 126: it is stated NADH was not observed, however in Tian 2017b, figure 2 there is NADH/NAD+ and NADPH/NAD+ data. It is not clear what data was or was not used from the previous work leading to confusion.

In Tian et al. 2017b, the metabolite data comes from two different experiments, with slightly different experimental conditions. In our initial analysis, we did not want to combine the data from the different datasets since we are not sure they are thermodynamically consistent (this is a frequent problem when you are trying to re-use data originally collected for a different purpose). In light of the reviewer’s comments, however, we have decided to investigate the effect of introducing these constraints. The two datasets had different biomass concentrations, thus we aligned then based on the amount of externally added ethanol concentration. We observed that inclusion of NAD(P) pools as constraints led to infeasibility because this NADH/NAD ratio was about four-fold higher than the value predicted by our feasible model. This increase in NADH/NAD ratio caused GAPDH infeasibility which can only be resolved by adding an uncertainty of ~ 90% to our model concentration bounds. This demonstrates the incompatibility of the two datasets.

The concentration bounds have been included in supplementary file XX. The main text has been modified to clarify the existence and the variation between the different datasets in Tian et al.

Minor:

Does the model distinguish between CO2 and bicarbonate and their equilibrium? Do the carboxylase enzymes all use CO2 or do some use bicarbonate?

The model does not distinguish between the various forms of carbon dioxide or bicarbonate ([Noor et al., 2014](#_ENREF_7); [Noor et al., 2013](#_ENREF_8)) . The model considers a single combined metabolite pool for CO2 and all its hydrated forms (i.e., aqueous CO2, carbonate, carbonic acid, and bicarbonate). All these forms are assumed to be in equilibrium in biological systems ([Radzicka and Wolfenden, 1995](#_ENREF_11)).

Line 343 does the model account for the spontaneous decarboxylation of OAA, how does this rate compare to the enzyme catalyzed reactions?

Although spontaneous decarboxylation of OAA to pyruvate can be measured in vitro, we do not think it plays a physiological role in *C. thermocellum*. There is no enzymatic conversion of OAA to pyruvate. For a detailed discussion, see ([Olson et al., 2017](#_ENREF_9)).

Fig 2 only shows 4 of the 8 considered metabolites, adding the other profiles would be useful

We have added few metabolites to highlight the trends in upper and lower glycolysis. All metabolite concentration ranges for various timepoints are listed in supplementary table zzz.

Table 1: what is the error of the metabolite measurements, why aren’t all three time points listed to facilitate communication?

The metabolite measurements were associated with 20% error. Concentrations bounds for all the timepoints are now listed in supplementary table YY. We have updated the main text to clarify this.

Table 2: is this the EFM model? The caption calls it the ‘list of possible reactions in any given EFM’. The unbalanced metabolites should be listed in the table caption.

Table 2 lists the pool of reactions from which the EFMs were constructed. A given EFM will include only a subset of the reactions listed in Table 2 which convert Cellobiose to ethanol. This has been clarified in the table caption. A complete list of EFMs is presented in Supplementary file 4.

Table 4: the values need units

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

The H2 levels in the experiment could have a large effect on the redox cofactor equilibria, was the H2 measured or is there an estimate of the range?

The dataset from Tian 2017b does not include H2 measurement, so we were not able to include that data in our model. There is recent evidence that H2 in the liquid phase is not in equilibrium with the gas phase ([Blunt et al., 2015](#_ENREF_1)), and measurement of dissolved H2 requires specialized equipment which we currently do not have access to.

Line 216 vs 324: is this supposed to be positive or negative? MDF (i.e., - 0.13 kJ/mol vs 0.13 kj/mol)

Line 216 talks about the EFMs which are thermodynamically infeasible and thus have negative MDF while Line 324 talks about the EFMs which carry positive MDF and are thermodynamically feasible.

EFMs play a large role in the document, yet they are only introduced in the results section. I would recommend introducing the concept in the introduction and providing a brief summary of earlier work which is relevant. The current work should be framed in terms of some of the earlier EFM thermodynamics and metabolic engineering work: [https://www.sciencedirect.com/science/article/pii/S1096717606000103](https://nam01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.sciencedirect.com%2Fscience%2Farticle%2Fpii%2FS1096717606000103&data=02%7C01%7Ccdm8%40psu.edu%7C43ce4adf338b45f1cc4b08d6c04b0f72%7C7cf48d453ddb4389a9c1c115526eb52e%7C0%7C1%7C636907825949193046&sdata=4ZSzL%2FO5LGlxnVnqLfF5T0AI%2Bb%2B86qxBR1OjXk%2BThMk%3D&reserved=0)

[https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0171440](https://nam01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fjournals.plos.org%2Fplosone%2Farticle%3Fid%3D10.1371%2Fjournal.pone.0171440&data=02%7C01%7Ccdm8%40psu.edu%7C43ce4adf338b45f1cc4b08d6c04b0f72%7C7cf48d453ddb4389a9c1c115526eb52e%7C0%7C1%7C636907825949203054&sdata=Pv9MraUKUAWn05e2GafHze2Lq1x29DfkdnZsvypLBP8%3D&reserved=0)

[https://link.springer.com/article/10.1007/s00253-008-1770-1](https://nam01.safelinks.protection.outlook.com/?url=https%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00253-008-1770-1&data=02%7C01%7Ccdm8%40psu.edu%7C43ce4adf338b45f1cc4b08d6c04b0f72%7C7cf48d453ddb4389a9c1c115526eb52e%7C0%7C1%7C636907825949203054&sdata=3BfE0U7tmHQfo%2B03yAiX7FgQYBp619GhlKp4x0b73Qo%3D&reserved=0)

<Answer>

We have modified the manuscript to introduce EFMs and highlight their usefulness in the introduction section.

Line 211 references citation [17] which isn’t consistent with the name and date format of the other references.

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Line 333, 335 mu is used on one line and u on the other, double check consistent use of ‘micro’ designator

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Line 62: missing the word ‘more’

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Line 191: missing an ‘s’ on reaction

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

-**Reviewer 2**  
  
Overall:

The authors performed a thermodynamic analysis of ethanol production in C. thermocellum with the goal of understanding why production is inhibited at high ethanol concentrations. The analysis is very high-quality and leads to meaningful metabolic explanations. The authors could likely go further in the discussion/conclusions section and address whether their findings change the industrial potential of C. thermocellum for ethanol production.

Questions/Points:

Table 3: Any issues with replacing NAD(H) with NADP(H)? Should cofactor stability be considered?

NADPH is generally used as a cofactor for enzymes involved in biosynthesis. Thus, using NADPH as a cofactor for fermentation reactions has the potential to reduce its availability for biosynthesis. Taken to extremes, this could be problematic, however wild-type *C. thermocellum* produces high levels of cell biomass, and reducing this value slightly might have a beneficial impact on biofuel yield. Previously, we have observed that introducing NADPH-ADH activity increases ethanol production ([Hon et al., 2017](#_ENREF_4)).

Ethanol toxicity has been thought to occur due to cell membrane leaking and possibly enzyme inhibition or unfolding. How do you know these factors are not involved in this case with C. thermocellum?

Wild type *C. thermocellum* can initiate growth in ethanol concentrations only up to about 20 g/l ([Brown et al., 2011](#_ENREF_2)). Inhibition of thermophiles by low levels of ethanol (3-4%) is thought to be due to enzyme inhibition, while inhibition at higher levels (4-8%) may be due to membrane effects ([Lovitt et al., 1988](#_ENREF_5)).

After learning that *T. saccharolyticum* has more favorable metabolism for ethanol production, which organism should be engineered? Shouldn’t the favorable characteristics of *C. thermocellum* be applied to *T. saccharolyticum* and not vice-versa?

A major advantage of *C. thermocellum’s* metabolism is its ability to solubilize cellulose. The reviewer raises an interesting point, however there are several challenges associated with cellulosome synthesizing genes. See ([Currie et al., 2013](#_ENREF_3)) for a discussion of some of the pitfalls of this approach.

It is somewhat unclear what conclusions are being made, and I feel there is plenty of room to establish metabolic and organismal guidelines for ethanol production based on the thermodynamic analysis. Some example questions follow…. Metabolic interventions for C. thermocellum were found, but can these be implemented? If so, what are the expected improvements in production?

We are currently working on implementing some of the suggestions presented in this paper. Once we have created these strains, and collected metabolite data from them, we plan to revisit this analysis.

Is the metabolic engineering strategy too difficult for this organism?

We have generated several different engineered strains of C. thermocellum with a half-dozen or more genetic modifications, so the set of three modifications proposed in this work is certainly reasonable.

Should continuous ethanol stripping during fermentation be used over metabolic engineering?

Ethanol stripping does keep ethanol titer low and increases volumetric productivity. While this approach has merit, and has indeed been investigated by the Lynd group for *C. thermocellum* fermentations ([Lynd et al., 1991](#_ENREF_6" \o "Lynd, 1991 #617)), improving titer via metabolic engineering also has merit and would be less expensive to implement if successful.

Or, does this analysis suggest C. thermocellum is not the best candidate for ethanol production?

Wild type *C. thermocellum* is not very good at producing ethanol, but we’re hoping to change this. For example, there is a fundamental tradeoff between energy conservation and thermodyanamic driving force. We may need to disrupt some of *C. thermocellum’s* energy conservation strategies to improve ethanol titer.

Should candidate organisms for ethanol production have a set of core metabolic characteristics to be considered for further engineering? If so, what are they? Can libraries of organisms (especially thermophiles) be scanned for these characteristics?

Most organisms that produce ethanol at high titer use a pathway for converting pyruvate to ethanol that involves the pyruvate decarboxylase reaction. There are only one or two examples of organisms that can produce ethanol at high titer using the pyruvate ferredoxin oxidoreductase pathway. With so few examples, it’s hard to draw general conclusions ([Olson et al., 2015](#_ENREF_10" \o "Olson, 2015 #605)).

-**Reviewer 3**

**(Please note that this review was sent directly to me and is not in the system)**

*Clostridium thermocellum* does show arguably the best cellulose degradation capabilities for bioethanol production from cellulosic material using consolidated bioprocessing approaches. While mutant strains have been successfully designed to produce ethanol to near theoretical yields per hexose, its capacity to produce high titers of ethanol is very limited. Based on metabolite concentration measurements and thermodynamic modelling, the present manuscript provides a plausible explanation for this observation, and suggests possible gene replacements that would alleviate thermodynamic pathway bottlenecks at high ethanol titers. From an evolutionary point of view, this research permits one to appreciate how this organism has evolve to maximize energy conservation per mole substrate rather than maintain biological energy pools through fluxing high amounts of substrates to end-products like most better-known fermenting organisms.

This is a very interesting set of findings with significant biotechnological applications for enhancing biofuel production from consolidated bioprocessing. There are, however, several improvements with respect to the clarity of the data presentation that can be suggested along with a few questions about the way some of the data are presented.

Major comments

1. Figure 2 legend. The positive slope of PFK, FBA, GAPDH, ALDH, and ADH reactions are not really visible within any part of this figure. A table of individual delta G values in the supplementals that you could refer the reader to would be helpful in order to appreciate these important findings.

We have added a supplementary table XX listing the change in delta G values and the cumulative delta G values for the pathway shown in Figure 2.

Minor comments

Line 62 – “…is **more** thermodynamically favorable…”

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Supplemental file 1 – indicating Units for amounts would be helpful to the readers?

In Supplemental file 1, the column “amount” refers to the concentration of the metabolite in the extraction solution in units of µM and the column “amount\_int” refers to the concentration of the metabolite calculated to have been present in the cytoplasm, also in units of µM.

Line 271 – Looks like an “is” might be missing between “MDF” and “constrained”

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Line 405 – “NADPH liked ALDH” should be “linked”

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Please look at figures 1 and 4 carefully. Some of the arrows are going in the wrong direction. For example, you are indicating in the green arrows that the conversion of acetaldehyde to ethanol used NADP to generate NADPH and you have 1,3pg +ATP yielding 3pg+ATP. I may have missed some other errors

We thank the reviewer for pointing these out. We have updated these corrections in the manuscript.

Blunt, W. A., Gapes, D. J., Sparling, R., Levin, D. B., Cicek, N., Quantitative assessment of H2 and CO2 supersaturation during thermophilic cellobiose fermentation with Clostridium thermocellum. 2015 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, 2015, pp. 1.

Brown, S. D., Guss, A. M., Karpinets, T. V., Parks, J. M., Smolin, N., Yang, S., Land, M. L., Klingeman, D. M., Bhandiwad, A., Rodriguez, M., Jr., Raman, B., Shao, X., Mielenz, J. R., Smith, J. C., Keller, M., Lynd, L. R., 2011. Mutant alcohol dehydrogenase leads to improved ethanol tolerance in Clostridium thermocellum. Proc Natl Acad Sci U S A. 108**,** 13752-7.

Currie, D. H., Herring, C. D., Guss, A. M., Olson, D. G., Hogsett, D. A., Lynd, L. R., 2013. Functional heterologous expression of an engineered full length CipA from Clostridium thermocellum in Thermoanaerobacterium saccharolyticum. Biotechnol Biofuels. 6**,** 32.

Hon, S., Olson, D. G., Holwerda, E. K., Lanahan, A. A., Murphy, S. J. L., Maloney, M. I., Zheng, T., Papanek, B., Guss, A. M., Lynd, L. R., 2017. The ethanol pathway from Thermoanaerobacterium saccharolyticum improves ethanol production in Clostridium thermocellum. Metab Eng. 42**,** 175-184.

Lovitt, R. W., Shen, G. J., Zeikus, J. G., 1988. Ethanol production by thermophilic bacteria: biochemical basis for ethanol and hydrogen tolerance in Clostridium thermohydrosulfuricum. J Bacteriol. 170**,** 2809-15.

Lynd, L. R., Ahn, H.-J., Anderson, G., Hill, P., Kersey, D. S., Klapatch, T., 1991. Thermophilic ethanol production investigation of ethanol yield and tolerance in continuous culture. Applied Biochemistry and biotechnology. 28**,** 549.

Noor, E., Bar-Even, A., Flamholz, A., Reznik, E., Liebermeister, W., Milo, R., 2014. Pathway thermodynamics highlights kinetic obstacles in central metabolism. PLoS Comput Biol. 10**,** e1003483.

Noor, E., Haraldsdottir, H. S., Milo, R., Fleming, R. M., 2013. Consistent estimation of Gibbs energy using component contributions. PLoS Comput Biol. 9**,** e1003098.

Olson, D. G., Horl, M., Fuhrer, T., Cui, J., Zhou, J., Maloney, M. I., Amador-Noguez, D., Tian, L., Sauer, U., Lynd, L. R., 2017. Glycolysis without pyruvate kinase in Clostridium thermocellum. Metab Eng. 39**,** 169-180.

Olson, D. G., Sparling, R., Lynd, L. R., 2015. Ethanol production by engineered thermophiles. Curr Opin Biotechnol. 33**,** 130-41.

Radzicka, A., Wolfenden, R., 1995. A proficient enzyme. Science. 267**,** 90-3.