Diagnosing and predicting mixed culture fermentations with unicellular and guild-based metabolic models

Running Title: MCFA Metabolic Models

Matthew J. Scarborough,[1,2,3*] *Joshua J. Hamilton*[4] *Elizabeth A. Erb,*[5] *Timothy J. Donohue*[1,6] *and Daniel R. Noguera*[1,2]

1. The Great Lakes Bioenergy Research Center, UW-Madison, Madison, WI
2. Department of Civil and Environmental Engineering, UW-Madison, Madison, WI
3. Department of Civil and Environmental Engineering, University of Vermont, Burlington, VT
4. Department of Biochemistry, UW-Madison, Madison, WI
5. Department of Chemical and Biological Engineering, UW-Madison, Madison, WI
6. Department of Bacteriology, UW-Madison, Madison, WI

*corresponding author e-mail: mscarbor@uvm.edu

ABSTRACT

Multi-species microbial communities determine the fate of materials in the environment and can be harnessed to produce beneficial products from renewable resources. In a recent example, fermentations by microbial communities have produced medium-chain fatty acids (MCFAs). Tools to predict, assess, and improve the performance of these communities, however, are limited. To provide such tools, we constructed two metabolic models of MCFA-producing microbial communities based on available genomic, transcriptomic and metabolomic data. The first model is a unicellular model (iFermCell215), while the second model (iFermGuilds789) separates fermentation activities into functional guilds. Ethanol and lactate are fermentation products known to serve as substrates for MCFA production, while acetate is another common co-metabolite during MCFA production. Simulations with iFermCell215 predict that low molar ratios of acetate-to-ethanol favor MCFA production, whereas the products of lactate and acetate co-utilization are less dependent on the acetate-to-lactate ratio. In simulations of an MCFA-producing community fed a complex organic mixture derived from lignocellulose, iFermGuilds789 predicted that lactate was an extracellular co-metabolite that served as a substrate for butyrate (C4) production. Extracellular hexanoic (C6) and octanoic acids (C8) were predicted by iFermGuilds789 to be from community members that directly metabolize sugars. Modeling results provide several hypotheses that can improve our understanding of microbial roles in a MCFA-producing microbiome and inform strategies to increase MCFA production. Further, these models represent novel tools for exploring the role of mixed microbial communities in carbon recycling in the environment, as well as on beneficial reuse of organic residues.

IMPORTANCE

Microbiomes are vital to human health, agriculture, and protecting the environment. Predicting behavior of self-assembled or synthetic microbiomes, however, remains a challenge. In this work, we used unicellular and guild-based metabolic models to investigate production of medium-chain fatty acids by a mixed microbial community that is fed multiple organic substrates. Modeling results provided insights into metabolic pathways of three medium-chain fatty acid-producing guilds and identified potential strategies to increase production of medium-chain fatty acids. This work demonstrates the role of metabolic models in augmenting multi-omic studies to gain greater insights into microbiome behavior.

INTRODUCTION

Mixed microbial fermentations have benefited humanity throughout history.[1] Since the mid-18th century, mixed culture fermentations have produced valuable chemicals and recovered energy from wastes.[2] Many mixed microbial fermentations involve the synergistic activities of several functional groups to convert organic substrates into valuable products. In anaerobic digestion, fermentation products like acetic acid (C2) and H_2 are used to support microbial production of methane, thus enabling the recovery of a large fraction of the chemical energy originally present in the complex organic substrates. While anaerobic digestion technologies are well established, other strategies to expand the range of potential products of mixed culture fermentations are only beginning to emerge. One example of such a mixed culture fermentation is the carboxylate platform, which produces medium-chain fatty acids (MCFAs) as a potentially valuable set of bio-based products.[3]

MCFAs are attractive carboxylate platform products because of their many industrial uses.[4] Octanoic acid (C8), an 8-carbon linear monocarboxylic acid, is a particularly valuable MCFA because of its high market value and relative ease of recovery. Further, when C8 is reduced to its

66 corresponding alkane, it can substitute for octane in liquid transportation fuels. Recent research
67 has demonstrated the feasibility of using mixed culture fermentations to transform renewable
68 resources into MCFAs.[5-7] While C8 is produced by some self-assembled bioreactor microbial
69 communities, hexanoic acid (C6) is often the primary MCFA, along with two short-chain fatty
70 acids (SCFAs), acetic (C2) and butyric (C4) acids.[5-10]

71 The microbial metabolic networks involved in converting complex organic substrates to
72 MCFAs by mixed culture fermentations are not fully elucidated. Consequently, there is a lack of
73 knowledge on how to obtain a carboxylate platform community that produces C8 as the main
74 MCFA, or how to minimize the accumulation of SCFAs. Initial models of mixed culture
75 fermentations to produce products other than methane have focused on simulating metabolic
76 networks for production of C2, C4, propionic acid, ethanol, lactic acid and H_2 from glucose.[11]
77 Subsequent models took into account that multiple electron carriers (e.g., ferredoxin, NADH) play
78 different roles in these metabolic networks, enabling the simulation of energy-conserving electron-
79 bifurcating reactions during production of C4 by reverse β-oxidation.[12] While a recent model
80 considered the role of homoacetogens as a H_2-consuming member of mixed culture fermentation,
81 it was limited to producing C4 as the major product of reverse β-oxidation and did not include
82 MCFA production.[13] Additionally, recent metabolic models assessed the potential for producing
83 C2, C3, C4 and valerate (C5) from amino acids.[14] While all of these models have enabled
84 predictions of mixed-culture fermentations from a single organic substrate, they lack the ability to
85 simulate production of MCFAs from complex feedstocks.

86 In this study, we constructed a unicellular model (iFermCell215) and a guild-based
87 microbial community model (iFermGuilds789) that offer expanded substrates and metabolic
88 pathways compared to existing models. We used iFermCell215 and iFermGuilds789 to analyze

89 pathways that lead to MCFA production. Simulations were performed with single organic
90 substrates, including sugars, glycerol, lactate, and acetate, with ethanol and acetate or lactate and
91 acetate as co-substrates, and with complex organic mixtures. The iFermGuilds789 was also used
92 to evaluate predictions of observed bioreactor behavior and of the flow of carbon and electrons
93 through different guilds within the mixed microbial community. In total, results of individual
94 model simulations generated new hypotheses on how to increase MCFA production by mixed
95 microbial communities.

96 **MATERIALS AND METHODS**

97 Metabolic networks of mixed culture fermentation that contain reactions that occur during
98 anaerobic metabolism of a previously described lignocellulosic biorefinery residue[15] were
99 assembled. The first network, iFermCell215, contains 215 reactions and 154 metabolites, models
100 mixed culture fermentation as a single unit, and expands the metabolic pathways and substrate
101 range of previously published mixed culture fermentation networks.[11-13] iFermCell215 describes
102 metabolism of exogenous substrates including glucans, glucose, xylans, xylose, and glycerol,
103 along with lactate, ethanol, acetate, CO_2 and H_2, which are potential intermediates in this mixed
104 culture fermentation (**Fig. 1**).

105 The second network, iFermGuilds789, represents activities of different guilds (i.e., groups
106 of organisms with similar functions) within the microbial community. The iFermGuilds789
107 network describes the same reactions and metabolites as in the iFermCell215 model but these are
108 separated into six functional guilds with each guild representing a subset of compartmentalized
109 reactions (**Fig. 2**). iFermGuilds789 also includes additional transport reactions that simulate
110 metabolite exchange among guilds. Combining the six guilds, the iFermGuilds789 network
111 includes 789 reactions and 579 metabolites. The guilds in the iFermGuilds789 model represent

112 previously described organisms that perform fatty acyl chain elongation from sugars (SEOs), sugar
113 fermenting organisms that produce lactate, acetate, and ethanol as fermentation products (SFO),
114 hydrogen-producing sugar fermenters (HSFs), organisms that perform fatty acyl chain elongation
115 from lactate (LEOs) or ethanol (EEOs), and homoacetogenic organisms (HAOs). Examples of
116 organisms falling within each functional guild are described in **Supplementary Text 1**.

117 When constructing each metabolic network, reactions were added to capture the exchange
118 of substrates and products with the extracellular space.[16] In addition to intracellular (denoted as
119 "_c" in the metabolic models) and extracellular ("_e") compartments, a compartment for ion-
120 motive force generation was included (denoted as "_i" in the metabolic models). Where possible,
121 standard metabolite and reaction abbreviations from the Biochemical, Genetic, and Genomic
122 (BiGG) knowledgebase were used.[17] All reactions were balanced for mass and charge assuming
123 an intracellular pH of 7.0 with the dominant ionic form of all metabolites obtained from the BiGG
124 database.[17]

125 Since many members of this mixed microbial community are Firmicutes, simplified
126 biomass equations for both models were created based on the biomass composition of the well-
127 studied gram-positive bacterium *Bacillus subtilis.*[18-20] We simplified all of the biosynthetic
128 pathways for production of biomass components to capture 16 precursor compounds produced by
129 the models. The precursor compounds included α-ketoglutarate, oxaloacetate, glucose-6-
130 phosphate, glucose-3-phosphate, 3-phosphoglycerate, pyruvate, phosphoenolpyruvate, acetyl-
131 CoA, erythrose-4-phosphate, ribulose-5-phosphate, ATP, NADPH, NAD+, ammonium, sulfate,
132 and H_2O. Additional information and calculations for the biomass equation are provided in
133 **Supplementary Text 2, Supplementary Data File 1, Figures S1-S6, and Table S1**. Both

134 metabolic models (iFermCell215 and iFermGuilds789) and instructions for their use are available
135 at https://github.com/mscarbor/Mixed-Culture-Fermentation-Models.

136 The production capabilities of both metabolic networks were simulated using constraint-
137 based methods.[21] These methods predict the steady-state flux distribution through a metabolic
138 network that is subject to physiochemical constraints. In this work, we used flux sampling[22] and
139 parsimonious flux balance analysis (pFBA).[23] Flux sampling was performed using artificial
140 centering hit and run (ACHR)[22] using 101 simulations and randomly generated seeds. A solution
141 with pFBA provides a flux distribution by minimizing the sum of all fluxes predicted by the
142 model.[23] When modeling bioreactors with the presence of solute gradients across the cell
143 membrane, we considered the ATP required for transport as described in **Supplementary Text 3.**
144 The objective function for all modeling scenarios was to maximize biomass production. When
145 simulating the iFermGuilds789 network for a steady-state condition, we required all individual
146 guilds to achieve the same growth rate (hr^{-1}). All simulations were performed in Python 3.7 using
147 the cobrapy package (v 0.18.1).[24]

148 Throughout this work, we present results in terms of electron equivalents (eeq), which
149 report the number of electrons potentially transferred to an electron acceptor if the compound is
150 completely oxidized (using factors provided in **Table S2** to convert from molar concentrations to
151 eeq concentrations). For each scenario, we calculated the overall reaction thermodynamics using
152 the predicted stoichiometry of the model and the standard free energy of formation (ΔG_f^0) of
153 metabolites (**Table S2**) obtained from KBase.[25] Unless otherwise noted, we ensured that at least
154 50 kJ was available per ATP produced under standard conditions at neutral pH.[26, 27]

155 **RESULTS**

156 **Simulation of mixed culture fermentations with iFermCell215**

157 By constraining the iFermCell215 model to produce defined products, we initially tested
158 its predictive ability with well-established fermentation conditions (See **Supplementary Text 4**
159 and **Table S3**). We then investigated predictions derived from an iFermCell215 model under
160 conditions where single substrates were consumed within this unit. The iFermCell215 model
161 depicts a simplified representation of fermentation in a mixed microbial community where
162 interspecies transfer of metabolites is assumed to have no energetic penalty, and is similar to
163 previously published models of mixed culture fermentations.[11-13] For this analysis, we required
164 that all possible solutions released ≥ 50 kJ per mol ATP produced.[26]

165 Using flux sampling, iFermCell215 provided 101 solutions that maximized biomass
166 production depending on the substrates provided to the simulation (**Fig. 3**). For instance, when we
167 modeled the utilization of glucose or xylose as the sole substrate, iFermCell215 predicted multiple
168 solutions that allowed for maximum biomass production (40% of eeq are incorporated into
169 biomass) with H_2, C2, C4, C6, and C8 as predicted products (**Fig. 3A**). None of the solutions in
170 this flux sampling analysis predicted ethanol or lactate as significant products. The predicted
171 biomass yield (eeq incorporated into biomass per eeq consumed) decreased with glycerol (37%),
172 lactate (20%), and ethanol (5.9%) as sole substrates. Consumption of ethanol as a sole substrate
173 was predicted to produce C8 and H_2 as the most energetically favorable solution (**Fig. 3D**), whereas
174 consumption of lactate as a sole substrate produced optimal solutions through heterofermentation
175 (**Fig. 3C**) and consumption of H_2 and CO_2 was predicted to produce C2 for optimal biomass
176 growth, which is consistent with the activity of homoacetogens (**Fig 3E**). These results suggest
177 that sugars, glycerol, and lactate can produce significant amounts of multiple fermentation
178 products, including C2, C4, C6, and C8, while supporting optimal biomass production, and that
179 production of ethanol or lactate, the two common fermentation products described as intermediates

180 in MCFA production,[28] are not as energetically favorable as producing other fermentation
181 products. It further suggests that consumption of ethanol as a sole substrate should favor
182 production of C8, and therefore, that ethanol may be a better intermediate product for optimizing
183 C8 production than lactate. However, since the fermentation of complex organic substrates by
184 mixed communities often results in accumulation of multiple fermentation products, we also
185 simulated conditions in which ethanol and acetate or lactate and acetate were co-substrates for
186 MCFA production.

187 **Simulation of ethanol and acetate utilization with iFermCell215**

188 Ethanol is a substrate that has often been used for chain elongation,[29] and early work on
189 MCFA production was performed using the ethanol- and acetate-consuming *Clositridium*
190 *kluyveri*.[30, 31] During initial work with *C. kluyveri*, it was proposed that C4 could be produced from
191 two moles of ethanol, without using acetate as an electron acceptor.[31] Recent work has confirmed
192 that *C. kluyveri* produces MCFAs from ethanol as a sole substrate in batch experiments,[32] in
193 agreement with the iFermCell215 prediction (**Fig. 3E**). Production of MCFAs from C2 as a sole
194 substrate has not been observed and is not predicted by iFermCell215. Past studies of ethanol
195 elongation have investigated optimal ratios of ethanol-to-acetate for MCFA production.[29, 33, 34]
196 Here, we use iFermCell215 to assess impacts of different molar ratios of acetate-to-ethanol
197 (C2:EtOH) on predicted product formation.

198 When supplied with both ethanol and C2, the iFermCell215 predicted maximum growth
199 rate of 0.11 hr^{-1} is more than double the predicted growth rate with ethanol as the sole substrate
200 (**Fig. 4A**), suggesting an ecological advantage for microorganisms that can co-utilize ethanol and
201 acetate. This growth rate can be achieved at C2:EtOH ratios from 0.56 to 1.28 mol:mol. At ratios
202 of 0.56 mol:mol and lower, C8 and H_2 are the only major predicted products (**Fig. 4**, left panels).

203 At ratios between 0.56 and 1.20, a mixture of C8, C6, and C4 are predicted by iFermCell215, and
204 at ratios of 1.28 and above, C4 is predicted as the sole product. These model predictions are
205 consistent with previous experimental work that shows that low ratios of C2:EtOH improve
206 specificity of C8 production.[33] The overall production rate of MCFAs, however, has been shown
207 to improve at higher ratios of C2:EtOH.[33] The highest overall MCFA production rate was achieved
208 by Spiritio et. al at a ratio of 0.53 mol:mol,[33] which is very close to the model predicted ratio of
209 0.56 that optimizes biomass growth and C8 production. The production of H_2 at low ratios of
210 C2:EtOH raises the question as to the impact of H_2 partial pressure on the thermodynamic
211 feasibility of the model results. We used iFermCell215 to calculate the maximum H_2 partial
212 pressure for all 101 scenarios and found that the lowest allowable H_2 partial pressure for growth
213 is greater than 10^6 atm (**Fig S7**), indicating that the proposed transformations are
214 thermodynamically favorable under all bioreactor H_2 partial pressure conditions.

215 In total, iFermCell215 predicts that the C2:EtOH ratio could either be used to improve
216 production of C8 as a desired product, or could be an indication that the environmental conditions
217 in a bioreactor would lead to a mixture of fermentation products. At low C2:EtOH ratios (<0.56),
218 iFermCell215 predicts that C8 should be the abundant product. At medium C2:EtOH ratios (0.56-
219 1.20) iFermCell215 predicts a mixture of C4, C6, and C8 would be produced, while at high ratios
220 (>1.20) of these substrates, iFermCell215 predicts C4 would be the predominant product.
221 Therefore, in a microbial community fermenting a mixture of substrates, iFermCell215 predicts
222 that different ratios of ethanol and C2 will impact the composition and abundance of products.
223 Additionally, iFermCell215 predicts that, if feeding a reactor with ethanol and C2, attention should
224 be paid to creating a feeding regime that optimizes production of the target product.

225 **Simulation of lactate and acetate utilization with iFermCell215**

226 In addition to ethanol, lactate is often used as a substrate for microbial chain elongation.
227 The well-studied *Megasphaera elsdenii* has been shown to produce MCFAs from lactate,[35] and
228 recent chain elongation studies with microbial communities have explored lactate as both a
229 substrate[9, 36] and a proposed key intermediate that supports chain elongation from carbohydrate
230 fermentation.[15, 37] While ethanol and lactate are both potential fermentation products and substrates
231 for chain elongation, iFermCell215 predicts that the pathways for utilization of these substrates
232 have key differences that impact the terminal fermentation products, as shown in the flux sampling
233 analyses (**Fig. 3D-E**), and may impact the products when C2 is a co-substrate. Therefore, similar
234 to the analysis shown above for C2:EtOH ratios, we used iFermCell215 to assess impacts of
235 different C2-to-lactate ratios.

236 In the case of lactate, the predicted maximum growth rate does not change when C2 is a
237 co-substrate (**Fig. 4E**). This suggests that, unlike ethanol and C2 co-utilization, there is no apparent
238 energetic benefit (reflected in growth rate) for a microorganism to co-utilize lactate and C2. Low
239 ratios of C2:lactate are predicted by iFermCell215 to produce a variety of end products, including
240 C8, C6, C4, and H_2 (**Fig. 4**, right panels). As the C2:lactate ratio increases, the product distribution
241 is predicted by iFermCell215 to change, with a decrease in MCFA and H_2 production accompanied
242 by an increase in C4 production, and at ratios of 0.96 and above, C4 is predicted as the sole product.
243 Like ethanol and C2 co-utilization, lactate and C2 co-utilization is not predicted by iFermCell215
244 to be impacted by bioreactor H_2 partial pressure (**Fig S7**). These results suggest that when lactate
245 is a key intermediate in an MCFA-producing community, controlling the relative amounts of C2
246 and lactate may not be an effective strategy for increasing C6 and C8 production. However, like
247 with ethanol, an excess amount of acetate is predicted by iFermCell215 to increase C4 production.

248 The predicted differences in product formation when acetate is co-metabolized with
249 ethanol or lactate suggests important differences in the metabolism of these substrates. Therefore,
250 to understand these differences, we also used iFermCell215 to assess the predicted metabolic
251 pathways of ethanol and lactate utilization.

252 **Predicted metabolic pathways for ethanol and lactate utilization**

253 The flux sampling analysis obtained with iFermCell215 consuming ethanol as a sole
254 substrate converged to a solution in which C8 was the only significant fermentation product (**Fig.**
255 **3E**). The flux of electron equivalents in this simulation predicted that reactions catalyzed by
256 ferredoxin hydrogenase (Hyd1), proton-translocating ferredoxin:NAD+ oxidoreductase (aka,
257 *Rhodobacter* nitrogen fixation complex; RNF), a coenzyme A transferase (CoAT), and acetate
258 kinase (AckR) were key enzymes used in the network (**Fig 5A**). The genome of the well-studied
259 ethanol chain-elongating organism, *C. kluyveri,* is predicted to contain genes that encode the
260 proteins for catalyzing all of these reactions.[38] Further, Hyd1 is the only hydrogen-producing gene
261 annotated in the genome of *C. kluyveri*.

262 The flux through the above enzymes predicts extensive use of energy conserving
263 mechanisms that rely on the generation of an ion-motive force (IMF). When ethanol is the sole
264 substrate (**Fig. 5A**), iFermCell215 predicted that the IMF is generated by the ATP synthase while
265 hydrolyzing ATP. The model predicts that the RNF enzyme uses the IMF to generate reduced
266 ferredoxin and that ATP is provided by acetate kinase. The reverse β–oxidation cycle uses an
267 electron bifurcating acyl-CoA dehydrogenase that also generates reduced ferredoxin, and redox
268 balance is maintained by Hyd1 that uses the reduced ferredoxin to generate H_2 **(Fig. 5A)**. When
269 C2 is available in addition to ethanol, iFermCell215 predicts that no net H_2 production is required
270 for redox balance **(Fig. 5B)**. Instead, iFermCell215 predicts that the activities of several enzymes

271 are reversed, with the RNF complex generating IMF, ATP synthase producing ATP, and acetate
272 kinase consuming ATP to generate acetyl-CoA. Thus, the same set of reactions is predicted to be
273 important by iFermCell215 when using ethanol as a sole substrate or ethanol and C2 as co-
274 substrates, but the predicted direction of key reactions change (**Fig 5A-B**). These predictions
275 suggest that organisms that elongate ethanol may readily transition between utilizing ethanol as a
276 sole substrate and co-utilizing ethanol and C2 since the same enzymes are predicted to be used,
277 and the co-utilization of C2 and ethanol improves growth rates (**Fig 4A**).

278 Unlike ethanol utilization, lactate conversion to acetyl-CoA is predicted by iFermCell215
279 to produce CO_2, formate, and reduced ferredoxin (**Fig. 5C-D**). The Wood-Ljungdahl pathway is
280 also predicted by iFermCell215 to use this CO_2, formate and reduced ferredoxin to produce C2
281 (**Fig. 5C-D**). The predicted ability to generate reduced ferredoxin during lactate consumption
282 (compared to only NADH from ethanol consumption) enables iFermCell215 to utilize a much
283 wider range of pathways for generating energy and maintaining redox balance. The reverse
284 β–oxidation cycle also uses an electron bifurcating acyl-CoA dehydrogenase under these
285 conditions, generating reduced ferredoxin, with RNF also predicted to be active in redox
286 balancing. These predictions of iFermCell215 explain the wide range of products that are possible
287 from utilization of lactate as sole substrate and co-utilization of C2 and lactate. Further, this
288 suggests that limiting the available metabolic pathways can impact the predicted products from
289 lactate. Indeed, removing specific reactions in iFermCell215 impacts the predicted product
290 spectrum (**Supplementary Text 5**, **Fig S7**). This demonstrates that strategies to improve MCFA
291 production may depend on the available set of microbial transformations and may differ between
292 bioreactor microbial communities.

293 **Simulating community behavior with iFermCell215 and iFermGuilds789**

294 The iFermCell215 model predicted that multiple substrates could lead to C8 production
295 (**Fig. 3**), but that biomass could be maximized by co-production of other fermentation products.
296 We wanted to assess if this single-unit model could simulate the observed product yields of a
297 mixed microbial community fed lignocellulosic biorefinery residues, a complex organic
298 substrate.[15] The previously described bioreactor was fed lignocellulosic biorefinery residues for
299 96 days, at which time its microbial community was evaluated using metagenomic and
300 metatranscriptomic analyses along with end-product sampling. To represent organic substrate
301 metabolism, the substrate uptake fluxes were set according to the observed consumption of
302 compounds in the feedstock (**Table S6**). We constrained iFermCell215 by setting the growth rate
303 of the model to 0.0069 hr^{-1}, corresponding to the bioreactor retention time of 6 d and we again
304 constrained the model only by requiring a minimum of 50 kJ of free energy release per mol of
305 ATP produced. Flux sampling results from this scenario predicted a range of potential products
306 (**Fig. 6A)**, including ethanol, C2, H_2, even chain carboxylic acids, and odd-chain carboxylic acids.
307 Overall, the amount of MCFAs predicted by iFermCell215 exceeded observed values, exposing a
308 limitation of the single unit model.

309 In an attempt to better represent the compartmentalized nature of activities within mixed
310 microbial communities we developed the functional guild model iFermGuilds789. In
311 iFermGuilds789, the metabolic activities of a mixed microbial community are partitioned among
312 a number of functional guilds. This biological division of labor negates the free exchange of
313 reducing equivalents (e.g., NADH, ferredoxin) and of many pathway intermediates (e.g., acetyl-
314 CoA) between community members, ultimately constraining how individual organisms maintain
315 redox balance, which can potentially impact ATP and product yields. For instance, the single unit
316 model (iFermCell215) cannot predict the production of lactate via a NADH-dependent

317 dehydrogenase and re-consumption of lactate via an electron-confurcating dehydrogenase,
318 activities that may happen in a mixed microbial community that includes lactate producers and
319 lactate consumers. The iFermGuilds789 model contains fermentation pathways in six guilds (**Fig.**
320 **2**), with each guild including predictions on transport reactions to simulate uptake and release of
321 intermediate metabolites by the most abundant members of the community as predicted from
322 previous metagenomics analyses.[15]

323 To simulate the activities in the MCFA-producing bioreactor we used a version of
324 iFermGuilds789 with only four guilds, which represented the most abundant community members
325 in the bioreactor as determined by metagenome sequencing.[15] These guilds included the sugar
326 elongating Cand. *Weimeria bifida* (SEOs), lactate-elongating Cand. *Pseudoramibacter fermentans*
327 (LEOs),[39] five sugar fermenting *Lactobacillus* species (SFOs), and three hydrogenic sugar
328 fermenting *Coriobacteriaceae* species (HSFs).[15] We constrained the abundance of each guild in
329 iFermGuilds789 according to the relative abundance of DNA for each guild in the bioreactor.
330 Guilds were assumed to have all pathways associated with their guild (**Fig. 2**), and we required
331 each guild to have the same growth rate of 0.0069 hr^{-1}. The results of this iFermGuilds789
332 simulation (**Fig. 6B**) predicted a narrower range of products than iFermCell215 (**Fig. 6A**), with
333 even- and odd-chain fatty acid products predicted. With these parameters, the iFermGuilds789
334 model overpredicted the production of C2 and C8 while underpredicting C4 and C6 production.

335 In an additional iFermGuilds789 simulation, we restricted pathways in guild members
336 based on previous metatranscriptomic analyses.[15] SEOs were constrained to use glucose, xylose
337 and glycerol; SFOs were restricted to consuming glucose, xylose, xylans, glucans, and glycerol;
338 HSFs were constrained to only consume glucans, glucose and glycerol; and LEOs were restricted
339 to consuming only lactate and glycerol (**Table S7**). Further, SEOs and LEOs were constrained to

340 not produce odd-chain volatile fatty acids (propionate, pentanoate, and heptanoate). LEOs were
341 further constrained to not use acetate kinase, which is absent from the genome of Cand.
342 *Pseudoramibacter fermentans*.[39] This iFermGuilds789 simulation predicted ranges of ethanol,
343 lactate, C2, and C6 that contained the experimentally observed values (**Fig 6C**). With these
344 parameters, C8 and H_2 production were overpredicted by iFermGuilds789, while C4 production
345 was underpredicted.

346 One potential explanation for the differences in the various iFermGuilds789 predictions
347 may be the specificity of chain elongating enzymes for producing products of defined chain
348 lengths. This could include any of the enzymes catalyzing the four steps of reverse β-oxidation
349 itself or the terminal enzymes that cleave the final products of reverse β-oxidation from CoA. We
350 tested this by constraining the flux through CoAT and acyl-CoA hydrolase reactions for both SEOs
351 and LEOs according to carbon-chain length. When constraining both CoAT and acyl-CoA
352 hydrolase in both SEOs and LEOs to 0.02 mmol $gDCW^{-1}$ hr^{-1} for C8 and 1 mmol $gDCW^{-1}$ hr^{-1} for
353 C6, iFermGuilds789 modeling predictions aligned much better with community performance (**Fig.**
354 **6D**). This suggests that reverse β-oxidation and/or terminal enzymes in this pathway have higher
355 affinity for C4 and C6 than for C8 in this specific microbial community.

356 Taken together, initial predictions from iFermCell215 (**Fig. 6A**) and iFermGuilds789 (**Fig.**
357 **6B**) did not align with observed bioreactor performance. Predictions of iFermGuilds789 improved
358 by incorporating knowledge of specific reactions within each guild (**Fig. 6C**). Further, the best
359 alignment between observations and model predictions resulted from constraining the fluxes
360 related to individual products with iFermGuilds789 (**Fig. 6D**). This suggests that accurate
361 predictions of MCFA production may require incorporation of enzyme affinity and additional
362 characterization of enzymes from chain-elongating species to improve modeling predictions.

363 **Assessing interspecies metabolite flux using iFermGuilds789**

364 Our previous analyses predicted that microbes within the MCFA-producing microbial
365 community perform different functions, compete for resources, and participate in interspecies
366 transfer of metabolites that allows for conversion of lignocellulosic biorefinery residues to SCFAs
367 and MCFAs.[15] We used iFermGuilds789 to predict the function of each guild in the transformation
368 of substrates as well as the production and transfer of intermediate metabolites between microbes.
369 In this case, unlike previous simulations with iFermGuilds789, we did not constrain the growth
370 rate (**Fig. 6B-D**) or fluxes for MCFA production (**Fig. 6D**). Instead, we constrained
371 iFermGuilds789 to accumulate extracellular products at the rates observed in the MCFA-
372 producing bioreactor (**Table S6**), constrained the reactions within each guild as described above,
373 and predicted fluxes using pFBA, which calculates a single flux distribution when the model is
374 constrained to the observed production rates. Under these conditions, iFermGuilds789 predicted a
375 growth rate of 0.0051 hr^{-1}, which is 74% of the expected growth rate (0.0069 hr^{-1}). This growth
376 rate corresponds to a retention time of 8 days (compared to the actual retention time of 6 days) and
377 these differences may reflect difficulty in estimating the amount of metabolically active cells in
378 the reactor or absence of unknown metabolic pathways that may increase the predicted growth
379 rate.

380 By constraining the substrates and products according to observed community behavior,
381 iFermGuilds789 predicted that SFOs and HSFs consumed xylans and glucans, while SEOs
382 consumed xylose, glucose, and 99% of the glycerol (**Fig 7**). SFOs were predicted to produce
383 ethanol, lactate and C2, while HSFs were predicted to produce a small amount of the lactate and
384 C2. SEOs are also predicted to produce lactate and C2. In total, 22% of the eeq in the feedstock
385 are predicted by this instance of iFermGuilds789 to be converted to lactate and 27% are predicted

386 to be converted to C2. All of the lactate and 52% of the C2 is predicted to be consumed by LEOs.
387 LEOs are predicted to produce all of the C4. SEOs are predicted to produce all of the C6 and C8
388 along with all of the H_2. SEOs are further predicted to have a flux for C6 production of 0.46 mmol
389 $gDCW^{-1}$ hr^{-1} and C8 production of 0.033 mmol $gDCW^{-1}$ hr^{-1} (**Supplementary Data File 5**), both
390 of which are lower than the maximum flux constraints applied previously (**Fig. 6D).**

391 **Predicted routes of sugar and lactate elongation**

392 The above prediction that essentially all MCFAs were produced by SEOs consuming
393 xylose, glycerol, and glucose, while LEOs produced C4 from lactate and C2 were different than
394 the predictions made earlier based on transcriptomic analyses alone.[15] To gain further insights into
395 SEO and LEO metabolism, we investigated the specific reactions predicted for both guilds by
396 iFermGuilds789. SEOs are predicted to produce pyruvate via glycolysis. Some pyruvate is
397 predicted to be converted to lactate, while most is converted to acetyl-CoA via pyruvate flavodoxin
398 oxidoreductase (PFOR, **Fig 8A**). iFermGuilds789 also predicted that acetyl-CoA is derived from
399 the phosphoketolase pathway (PKETX, **Fig 8A**). Acetyl-CoA is predicted to be directed to reverse
400 β-oxidation, but a portion is also predicted to be hydrolyzed by an acyl-CoA hydrolase. Acyl-CoA
401 hydrolase enzymes are also predicted as terminal enzymes of reverse β-oxidation rather than
402 energy-conserving CoA transferase enzymes (CoAT). SEOs are further predicted to use Hyd1 for
403 hydrogen production, to generate IMF via RNF1, and produce ATP via ATP synthase (ATPS4r)
404 (**Fig. 8A**).

405 LEOs are predicted to convert lactate to pyruvate via an electron-confurcating lactate
406 dehydrogenase (**Fig. 8B**, ECLDH), and pyruvate is predicted to form acetyl-CoA via PFOR. All
407 of the acetyl-CoA is predicted to undergo reverse β-oxidation and C2 is incorporated by a CoAT
408 that transfers the CoA from butyryl-CoA to acetyl-CoA (**Fig. 8B**). LEOs are predicted to consume

409 equimolar concentrations of C2 and lactate, which was previously predicted to favor C4 production
410 (**Fig. 4**). Further, LEOs are not predicted to produce H_2. iFermGuilds789 predicted that H_2
411 produced by Hyd1 is re-consumed via a proton-translocating energy conserving hydrogenase that
412 produces reduced ferredoxin (ECH, **Fig. 8B**). This ferredoxin augments that produced by PFOR
413 and is predicted to be used for lactate consumption by ECLDH and the RNF complex. The use of
414 an energy conserving hydrogenase for ferredoxin production from H_2 has been demonstrated
415 previously.[40]

416 While both SEOs and LEOs are predicted by iFermGuilds789 to use reverse β-oxidation
417 to elongate acetyl-CoA, the energy-conserving and redox-balancing reactions needed to support
418 this process are predicted to differ. While SEOs are predicted to use an acyl-CoA hydrolase as the
419 terminal enzyme of reverse β-oxidation, LEOs are predicted to use a CoAT which conserves
420 energy by activating C2 to acetyl-CoA, which would otherwise require ATP. In the previous
421 metatranscriptomic analyses of these community members,[15, 39] we found that both SEOs and
422 LEOs expressed genes for both acyl-CoA hydrolases and CoAT. The specificity of these predicted
423 chain elongating enzymes for different substrates, however, is not known. The iFermGuilds789
424 model predictions reinforce the potential importance of terminal enzymes in reverse β-oxidation
425 in determining the final product length; specifically, it predicts that all C4 is produced using CoAT
426 and all MCFAs are using acyl-CoA hydrolase.

427 Furthermore, SEOs are predicted by iFermGuilds789 to maintain redox balance by
428 producing H_2 with Hyd1, while LEOs are predicted to recycle all H_2 internally using both Hyd1
429 and ECH. Previous metatranscriptomic analyses showed that both LEOs and SEOs expressed
430 genes for both Hyd1 and ECH.[15] This agrees with iFermGuilds789 predictions for LEOs, but not
431 with SEOs that are predicted to only use Hyd1. When constraining iFermGuilds789 to require

432 SEOs to use ECH, there are no changes in interspecies metabolite transfer (**Supplementary Data**
433 **File 5),** but SEOs are predicted to generate additional ATP via ATP synthase while reducing flux
434 through other energy-conserving pathways to achieve the same growth rate (**Supplementary Data**
435 **File 5**). These findings suggest that SEOs and LEOs have multiple methods for conserving energy
436 and maintaining redox balance and this metabolic flexibility can maintain the same overall
437 transformation of complex substrates to products. This metabolic flexibility may reflect ecological
438 advantages for fermenting organisms, but are also challenges for the accurate predictions of MCFA
439 production, which may strongly depend on the specific organisms present in the microbial
440 community.
441

442 **DISCUSSION**

443 Under anaerobic conditions, a diverse set of mixed microbial communities catalyze the
444 transformation of organic substrates to a variety of products. For over a century, anaerobic
445 microbial communities have been harnessed to recover energy, in the form of methane-rich biogas,
446 from complex organic matter in anaerobic digestion.[2] Some fermentation metabolisms are well
447 understood, and are used by genetically tractable pure cultures that can be analyzed, modeled and
448 improved for use in industrial applications.[41, 42] In contrast, the activities of mixed microbial
449 communities have been harder to analyze, predict or redirect to improve substrate utilization or
450 product yields. The ability to harness the metabolic activity of mixed microbial communities to
451 produce alternative fermentation products is limited since metabolic and kinetic models that
452 accurately predict the distribution of fermentation products are only starting to emerge.[11-13] To
453 address this knowledge gap, we constructed metabolic models of mixed culture fermentations with
454 the objective of gaining greater insights into how to increase MCFA production by mixed

455 microbial communities. We focused modeling the production of C6 and C8, linear medium-length
456 MCFAs of industrial value that can be produced from a variety of organic residues.[3] To build and
457 test these models we took advantage of prior metagenomic, metatranscripotomic, and metabolomic
458 analyses of a mixed community that produces C6 and C8 fatty acids from a lignocellulosic refinery
459 residue.[15, 39]

460 The predictions of our models generated several hypotheses that should be explored with
461 further analyses. First, the single unit iFermCell215 predicted production of MCFAs from ethanol
462 as a sole source of electrons and carbon (**Fig. 4A-D**), and this is supported by recent experiments
463 with mixed cultures.[33] More work is needed to determine bioreactor conditions that promote
464 elongation of ethanol as a sole substrate or with low C2:EtOH ratios. The iFermCell215
465 simulations predict a threshold of 0.56 mol C2 per mol ethanol below which C8 production benefits
466 from C2 co-utilization, and above which C8 production starts to decrease and to be replaced by
467 production of shorter acids. This shift in product profile was predicted by iFermCell215 to be
468 associated with reversal in the direction of enzymatic activities related to energy conserving
469 metabolism, an observation that still lacks experimental verification.

470 Second, even though lactate is a known substrate for chain elongation,[9] iFermGuilds789
471 simulation results suggest that, in this mixed community, lactate may be converted to C4 instead
472 of C6 or C8 **(Fig. 7)**. Although at lower C2 to lactate ratios iFermCells215 simulations predicted
473 that this ratio was not as influential on determining the product distribution as the C2:EtOH ratio,
474 it also predicted that excess C2 could result in C4 production instead of C6 or C8 from both
475 intermediate metabolites (**Fig. 4**). The iFermGuilds789 simulation predicted co-utilization of C2
476 and lactate at a molar ratio of 1.0 (**Fig. 8B**), which according to the iFermCell215 simulations (**Fig.
477 4**) this is already an excess C2 condition predicted to yield C4 as the main fermentation product.

478 These results have the important implication that efforts directed to construct synthetic
479 communities or control self-assembled communities for production of MCFA should focus on
480 minimizing C2 production and accumulation.

481 Third, iFermCell215 modeling results suggest that controlling product formation from
482 lactate may not be possible by controlling C2-to-lactate ratios (**Fig. 4E-G**), whereas C2:EtOH
483 ratios suggests very specific targets to maximize C8 production (**Fig. 4A-D**). This suggests that
484 for MCFA production from complex organic substrates, in addition to targeting the reduction of
485 C2 production, efforts should also be directed to creating conditions in which ethanol instead of
486 lactate is the key intermediate metabolite for MCFA production. However, this may not be
487 achievable in self-assembled communities, and since simulations predict that the products of
488 lactate fermentation are dependent on specific metabolic features of LEOs (**Fig S8**), it is important
489 to continue the discovery of the natural diversity of lactate fermentation metabolisms, which
490 eventually could help identify preferred LEOs for the development of novel strategies to promote
491 MCFA production when lactate is the key intermediate metabolite.

492 Fourth, iFermGuilds789 predicts that enzyme affinity plays a major role in determining the
493 final products of reverse β-oxidation and mixed culture fermentation (**Fig 6D**). Additional work is
494 needed to assess the affinity of reverse β-oxidation, acyl-CoA hydrolase, and CoA transferase
495 enzymes for different chain lengths to identify ideal microorganisms or engineer microorganisms
496 to improve MCFA production. Past studies have identified and engineered acyl-ACP thioesterase
497 enzymes (the terminal enzyme of fatty acid synthesis via the fatty acid biosynthesis pathway) with
498 high affinity for C8.[43-45] Similar studies should be performed with acyl-CoA hydrolase and acyl-
499 CoA transferase enzymes to characterize and improve the terminal enzymes of reverse β-
500 oxidation.

501 Since many microbes in natural and self-assembled fermentation communities remain
502 uncultured, our knowledge and ability to predict important metabolic networks relies largely on
503 information from metagenomics, metatranscriptomics, and metabolomics analyses to build the
504 functional activities in each guild.[15, 46] We show in this work that metabolic modeling offers
505 additional insight that is not gained by these analyses, but that is critical in the evaluation of the
506 microbial communities and in developing hypothesis of how to direct metabolism towards a
507 specific set of fermentation products. From prior analyses,[15] we anticipated that both SEOs and
508 LEOs participated in MCFA production, but the iFermGuilds789 simulation of interspecies
509 metabolite flux (**Fig. 7**) predicted that to achieve the observed growth rate of the microbial
510 community, C6 and C8 were produced solely by SEOs, while LEOs were responsible for
511 production of C4. This difference between predictions based on metatranscriptomic and simulation
512 analyses can be tracked to the importance of co-utilization of fermentation products (i.e., C2 and
513 lactate), which is easier to explore using metabolic models.

514 iFermCell215 and iFermGuilds789 are examples of first-generation predictive models of a
515 mixed culture fermentation. Their utility can be improved by obtaining additional knowledge of
516 metabolic pathways involved in fermentation through studying individual microbes within each
517 functional guild. Further effort is needed to simulate poorly appreciated conditions that shape
518 community structure, such as the toxicity of fermentation products on the specific guilds or the
519 contribution(s) of low abundance microbes to the overall activity of the mixed culture
520 fermentation. Additionally, computational methods to predict the abundance of specific guilds,
521 rather than relying on relative abundances as a model input, are needed. Efforts to assemble such
522 models have begun for the human gut microbiome,[47, 48] and as these approaches are applied to

523 mixed microbial fermentations, their benefits to the production of MCFA and other chemicals of
524 societal importance will continue to improve.

525 **ACKNOWLEDGMENTS**

526 This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER
527 Office of Science DE-SC0018409). M.J.S. was supported by the National Science Foundation
528 Graduate Research Fellowship Program under grant No. DGE-1256259 and by the University of
529 Vermont College of Engineering and Mathematical Sciences. J.J.H. was supported by the National
530 Institute of Food and Agriculture, U.S. Department of Agriculture, under award number 2016-
531 67012-24709. The authors extend our gratitude to Robert Aidan Campbell for assistance with
532 Python coding.

533 **REFERENCES**

534 1. McGovern, P. E.; Zhang, J.; Tang, J.; Zhang, Z.; Hall, G. R.; Moreau, R. A.; Nunez, A.;
535 Butrym, E. D.; Richards, M. P.; Wang, C. S.; Cheng, G.; Zhao, Z.; Wang, C., Fermented beverages
536 of pre- and proto-historic China. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, (51), 17593-8.

537 2. Bond, T.; Templeton, M. R., History and future of domestic biogas plants in the developing
538 world. *Energy for Sustainable Development* **2011,** *15*, (4), 347-354.

539 3. Angenent, L. T.; Richter, H.; Buckel, W.; Spirito, C. M.; Steinbusch, K. J. J.; Plugge, C.
540 M.; Strik, D.; Grootscholten, T. I. M.; Buisman, C. J. N.; Hamelers, H. V. M., Chain Elongation
541 with Reactor Microbiomes: Open-Culture Biotechnology To Produce Biochemicals. *Environ. Sci.*
542 *Technol.* **2016,** *50*, (6), 2796-2810.

543 4. Sarria, S.; Kruyer, N. S.; Peralta-Yahya, P., Microbial synthesis of medium-chain
544 chemicals from renewables. *Nat. Biotechnol.* **2017,** *35*, (12), 1158-1166.

545 5. Scarborough, M. J.; Lynch, G.; Dickson, M.; McGee, M.; Donohue, T. J.; Noguera, D. R.,
546 Increasing the economic value of lignocellulosic stillage through medium-chain fatty acid
547 production. *Biotechnology for Biofuels* **2018,** *11*, (1), 200.

548 6. Nelson, R.; Peterson, D.; Karp, E.; Beckham, G.; Salvachúa, D., Mixed Carboxylic Acid
549 Production by Megasphaera elsdenii from Glucose and Lignocellulosic Hydrolysate. *Fermentation*
550 **2017,** *3*, (1), 10.

551 7. Kucek, L. A.; Xu, J. J.; Nguyen, M.; Angenent, L. T., Waste Conversion into n-Caprylate
552 and n-Caproate: Resource Recovery from Wine Lees Using Anaerobic Reactor Microbiomes and
553 In-line Extraction. *Front. Microbiol.* **2016,** *7*.

554 8. Zhu, X.; Zhou, Y.; Wang, Y.; Wu, T.; Li, X.; Li, D.; Tao, Y., Production of high-
555 concentration n-caproic acid from lactate through fermentation using a newly isolated
556 Ruminococcaceae bacterium CPB6. *Biotechnol Biofuels* **2017,** *10*, 102.

557 9. Kucek, L. A.; Nguyen, M.; Angenent, L. T., Conversion of L-lactate into n-caproate by a
558 continuously fed reactor microbiome. *Water Research* **2016,** *93*, 163-171.

559 10. Coma, M.; Vilchez-Vargas, R.; Roume, H.; Jauregui, R.; Pieper, D. H.; Rabaey, K.,
560 Product Diversity Linked to Substrate Usage in Chain Elongation by Mixed-Culture Fermentation.
561 *Environ. Sci. Technol.* **2016,** *50*, (12), 6467-76.

562 11. Rodriguez, J.; Kleerebezem, R.; Lema, J. M.; van Loosdrecht, M. C. M., Modeling product
563 formation in anaerobic mixed culture fermentations. *Biotechnol. Bioeng.* **2006,** *93*, (3), 592-606.

564 12. Zhang, F.; Zhang, Y.; Chen, M.; van Loosdrecht, M. C. M.; Zeng, R. J., A modified
565 metabolic model for mixed culture fermentation with energy conserving electron bifurcation
566 reaction and metabolite transport energy. *Biotechnol. Bioeng.* **2013,** *110*, (7), 1884-1894.

567 13. Regueira, A.; Gonzalez-Cabaleiro, R.; Ofiteru, I. D.; Rodriguez, J.; Lema, J. M., Electron
568 bifurcation mechanism and homoacetogenesis explain products yields in mixed culture anaerobic
569 fermentations. *Water Res.* **2018,** *141*, 349-356.

570 14. Regueira, A.; Lema, J. M.; Carballa, M.; Mauricio-Iglesias, M., Metabolic modeling for
571 predicting VFA production from protein-rich substrates by mixed-culture fermentation. *Biotechnol*
572 *Bioeng* **2020,** *117*, (1), 73-84.

573 15. Scarborough, M. J.; Lawson, C. E.; Hamilton, J. J.; Donohue, T. J.; Noguera, D. R.,
574 Metatranscriptomic and Thermodynamic Insights into Medium-Chain Fatty Acid Production
575 Using an Anaerobic Microbiome. *mSystems* **2018,** *3*, (6).

576 16. Covert, M. W.; Schilling, C. H.; Famili, I.; Edwards, J. S.; Goryanin, I. I.; Selkov, E.;
577 Palsson, B. O., Metabolic modeling of microbial strains in silico. *Trends Biochem. Sci.* **2001,** *26*,
578 (3), 179-186.

579 17. King, Z. A.; Lu, J.; Drager, A.; Miller, P.; Federowicz, S.; Lerman, J. A.; Ebrahim, A.;
580 Palsson, B. O.; Lewis, N. E., BiGG Models: A platform for integrating, standardizing and sharing
581 genome-scale models. *Nucleic Acids Res.* **2016,** *44*, (D1), D515-22.

582 18. Henry, C. S.; Zinner, J. F.; Cohoon, M. P.; Stevens, R. L., iBsu1103: a new genome-scale
583 metabolic model of Bacillus subtilis based on SEED annotations. *Genome Biol* **2009,** *10*, (6), R69.

584 19. Oh, Y. K.; Palsson, B. O.; Park, S. M.; Schilling, C. H.; Mahadevan, R., Genome-scale
585 reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping
586 and gene essentiality data. *J Biol Chem* **2007,** *282*, (39), 28791-9.

587 20. Dauner, M.; Sauer, U., Stoichiometric growth model for riboflavin-producing Bacillus
588 subtilis. *Biotechnol Bioeng* **2001,** *76*, (2), 132-43.

589 21. Bordbar, A.; Monk, J. M.; King, Z. A.; Palsson, B. O., Constraint-based models predict
590 metabolic and associated cellular functions. *Nat. Rev. Genet.* **2014,** *15*, (2), 107-20.

591 22. Megchelenbrink, W.; Huynen, M.; Marchiori, E., optGpSampler: an improved tool for
592 uniformly sampling the solution-space of genome-scale metabolic networks. *PLoS One* **2014,** *9*,
593 (2), e86587.

594 23. Lewis, N. E.; Hixson, K. K.; Conrad, T. M.; Lerman, J. A.; Charusanti, P.; Polpitiya, A.
595 D.; Adkins, J. N.; Schramm, G.; Purvine, S. O.; Lopez-Ferrer, D.; Weitz, K. K.; Eils, R.; Konig,
596 R.; Smith, R. D.; Palsson, B. O., Omic data from evolved E. coli are consistent with computed
597 optimal growth from genome-scale models. *Mol Syst Biol* **2010,** *6*, 390.

598 24. Ebrahim, A.; Lerman, J. A.; Palsson, B. O.; Hyduke, D. R., COBRApy: COnstraints-Based
599 Reconstruction and Analysis for Python. *BMC Syst. Biol.* **2013,** *7*, 74.

600 25. Arkin, A. P.; Cottingham, R. W.; Henry, C. S.; Harris, N. L.; Stevens, R. L.; Maslov, S.;
601 Dehal, P.; Ware, D.; Perez, F.; Canon, S.; Sneddon, M. W.; Henderson, M. L.; Riehl, W. J.;
602 Murphy-Olson, D.; Chan, S. Y.; Kamimura, R. T.; Kumari, S.; Drake, M. M.; Brettin, T. S.; Glass,
603 E. M.; Chivian, D.; Gunter, D.; Weston, D. J.; Allen, B. H.; Baumohl, J.; Best, A. A.; Bowen, B.;
604 Brenner, S. E.; Bun, C. C.; Chandonia, J. M.; Chia, J. M.; Colasanti, R.; Conrad, N.; Davis, J. J.;
605 Davison, B. H.; DeJongh, M.; Devoid, S.; Dietrich, E.; Dubchak, I.; Edirisinghe, J. N.; Fang, G.;
606 Faria, J. P.; Frybarger, P. M.; Gerlach, W.; Gerstein, M.; Greiner, A.; Gurtowski, J.; Haun, H. L.;

He, F.; Jain, R.; Joachimiak, M. P.; Keegan, K. P.; Kondo, S.; Kumar, V.; Land, M. L.; Meyer, F.; Mills, M.; Novichkov, P. S.; Oh, T.; Olsen, G. J.; Olson, R.; Parrello, B.; Pasternak, S.; Pearson, E.; Poon, S. S.; Price, G. A.; Ramakrishnan, S.; Ranjan, P.; Ronald, P. C.; Schatz, M. C.; Seaver, S. M. D.; Shukla, M.; Sutormin, R. A.; Syed, M. H.; Thomason, J.; Tintle, N. L.; Wang, D.; Xia, F.; Yoo, H.; Yoo, S.; Yu, D., KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat. Biotechnol.* **2018,** *36*, (7), 566-569.

26. Thauer, R. K.; Jungermann, K.; Decker, K., Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* **1977,** *41*, (1), 100-80.

27. Alberty, R. A., *Thermodynamics of biochemical reactions*. John Wiley & Sons, Inc. : Hoboken, New Jersey, 2003.

28. Han, W. H.; He, P. J.; Shao, L. M.; Lu, F., Road to full bioconversion of biowaste to biochemicals centering on chain elongation: A mini review. *J Environ Sci-China* **2019,** *86*, 50-64.

29. Spirito, C. M.; Richter, H.; Rabaey, K.; Stams, A. J.; Angenent, L. T., Chain elongation in anaerobic reactor microbiomes to recover resources from waste. *Curr Opin Biotechnol* **2014,** *27*, 115-22.

30. Schoberth, S.; Gottschalk, G., Considerations on the energy metabolism of Clostridium kluyveri. *Arch Mikrobiol* **1969,** *65*, (4), 318-28.

31. Thauer, R. K.; Jungermann, K.; Henninger, H.; Wenning, J.; Decker, K., The energy metabolism of Clostridium kluyveri. *Eur J Biochem* **1968,** *4*, (2), 173-80.

32. Yin, Y.; Zhang, Y.; Karakashev, D. B.; Wang, J.; Angelidaki, I., Biological caproate production by Clostridium kluyveri from ethanol and acetate as carbon sources. *Bioresour Technol* **2017,** *241*, 638-644.

33. Spirito, C. M.; Marzilli, A. M.; Angenent, L. T., Higher Substrate Ratios of Ethanol to Acetate Steered Chain Elongation toward n-Caprylate in a Bioreactor with Product Extraction. *Environ Sci Technol* **2018,** *52*, (22), 13438-13447.

34. Roghair, M.; Hoogstad, T.; Strik, D.; Plugge, C. M.; Timmers, P. H. A.; Weusthuis, R. A.; Bruins, M. E.; Buisman, C. J. N., Controlling Ethanol Use in Chain Elongation by CO2 Loading Rate. *Environ Sci Technol* **2018,** *52*, (3), 1496-1505.

635 35. Weimer, P. J.; Moen, G. N., Quantitative analysis of growth and volatile fatty acid
636 production by the anaerobic ruminal bacterium Megasphaera elsdenii T81. *Appl Microbiol*
637 *Biotechnol* **2013,** *97*, (9), 4075-81.

638 36. Zhu, X.; Tao, Y.; Liang, C.; Li, X.; Wei, N.; Zhang, W.; Zhou, Y.; Yang, Y.; Bo, T., The
639 synthesis of n-caproate from lactate: a new efficient process for medium-chain carboxylates
640 production. *Sci Rep* **2015,** *5*, 14360.

641 37. Andersen, S. J.; De Groof, V.; Khor, W. C.; Roume, H.; Props, R.; Coma, M.; Rabaey, K.,
642 A Clostridium Group IV Species Dominates and Suppresses a Mixed Culture Fermentation by
643 Tolerance to Medium Chain Fatty Acids Products. *Front Bioeng Biotechnol* **2017,** *5*, 8.

644 38. Seedorf, H.; Fricke, W. F.; Veith, B.; Bruggemann, H.; Liesegang, H.; Strittmatter, A.;
645 Miethke, M.; Buckel, W.; Hinderberger, J.; Li, F.; Hagemeier, C.; Thauer, R. K.; Gottschalk, G.,
646 The genome of Clostridium kluyveri, a strict anaerobe with unique metabolic features. *Proc Natl*
647 *Acad Sci U S A* **2008,** *105*, (6), 2128-33.

648 39. Scarborough, M. J.; Myers, K. S.; Donohue, T. J.; Noguera, D. R., Medium-Chain Fatty
649 Acid Synthesis by "Candidatus Weimeria bifida" gen. nov., sp. nov., and "Candidatus
650 Pseudoramibacter fermentans" sp. nov. *Appl Environ Microbiol* **2020,** *86*, (3).

651 40. Hedderich, R.; Forzi, L., Energy-converting [NiFe] hydrogenases: More than just H-2
652 activation. *J Mol Microb Biotech* **2005,** *10*, (2-4), 92-104.

653 41. Juturu, V.; Wu, J. C., Microbial production of lactic acid: the latest development. *Crit. Rev.*
654 *Biotechnol.* **2016,** *36*, (6), 967-977.

655 42. Buschke, N.; Schafer, R.; Becker, J.; Wittmann, C., Metabolic engineering of industrial
656 platform microorganisms for biorefinery applications--optimization of substrate spectrum and
657 process robustness by rational and evolutive strategies. *Bioresour. Technol.* **2013,** *135*, 544-54.

658 43. Sarria, S.; Bartholow, T. G.; Verga, A.; Burkart, M. D.; Peralta-Yahya, P., Matching
659 Protein Interfaces for Improved Medium-Chain Fatty Acid Production. *ACS Synth Biol* **2018,** *7*,
660 (5), 1179-1187.

661 44. Hernandez Lozada, N. J.; Lai, R. Y.; Simmons, T. R.; Thomas, K. A.; Chowdhury, R.;
662 Maranas, C. D.; Pfleger, B. F., Highly Active C8-Acyl-ACP Thioesterase Variant Isolated by a
663 Synthetic Selection Strategy. *ACS Synth Biol* **2018,** *7*, (9), 2205-2215.

664 45. Deng, X.; Chen, L.; Hei, M.; Liu, T.; Feng, Y.; Yang, G. Y., Structure-guided reshaping of
665 the acyl binding pocket of 'TesA thioesterase enhances octanoic acid production in E. coli. *Metab*
666 *Eng* **2020,** *61*, 24-32.

667 46. Han, W.; He, P.; Shao, L.; Lu, F., Metabolic Interactions of a Chain Elongation
668 Microbiome. *Appl. Environ. Microbiol.* **2018**.

669 47. Zomorrodi, A. R.; Maranas, C. D., OptCom: a multi-level optimization framework for the
670 metabolic modeling and analysis of microbial communities. *PLoS Comput. Biol.* **2012,** *8*, (2),
671 e1002363.

672 48. Kumar, M.; Ji, B.; Zengler, K.; Nielsen, J., Modelling approaches for studying the
673 microbiome. *Nat Microbiol* **2019,** *4*, (8), 1253-1267.
674

-

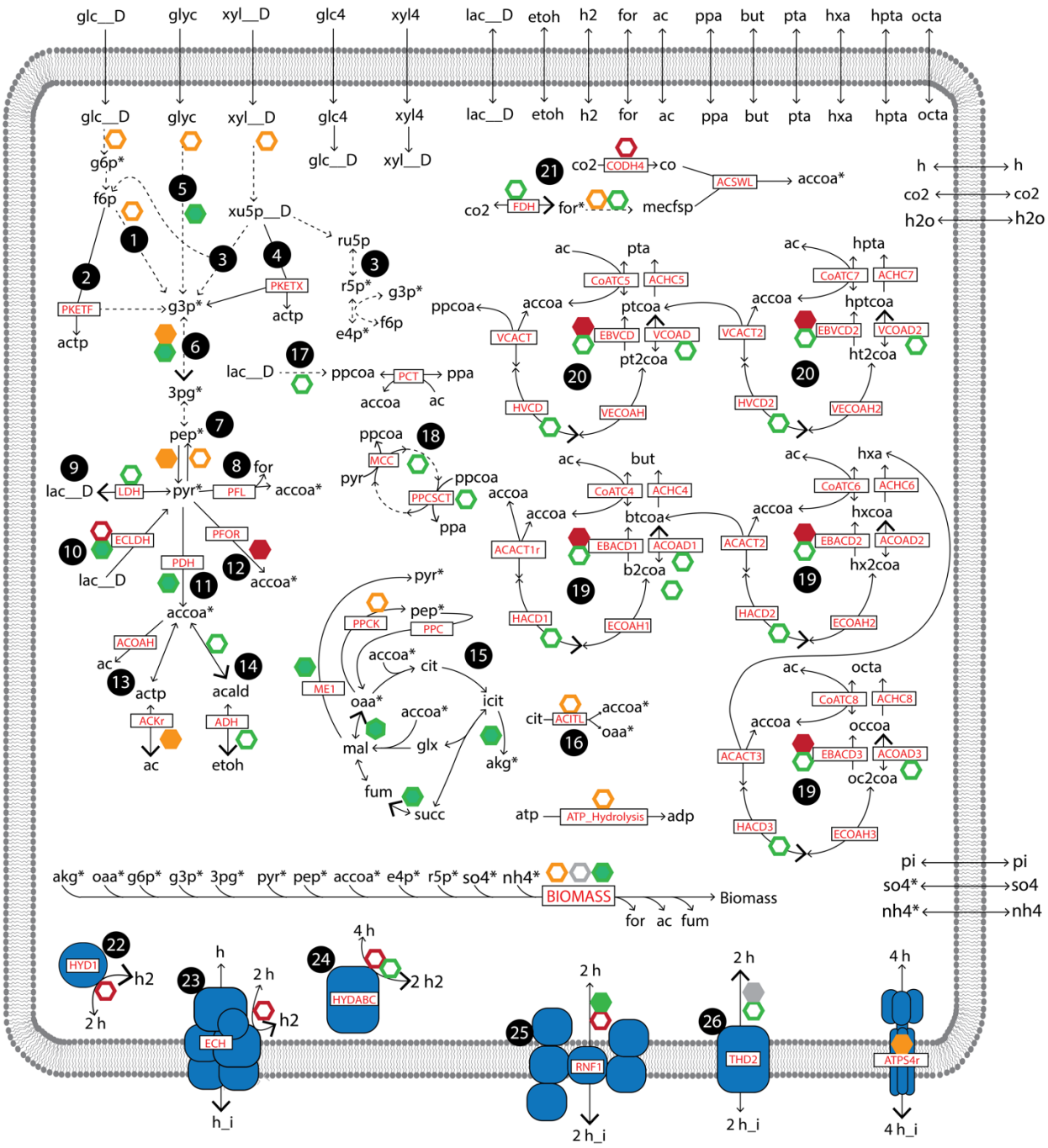
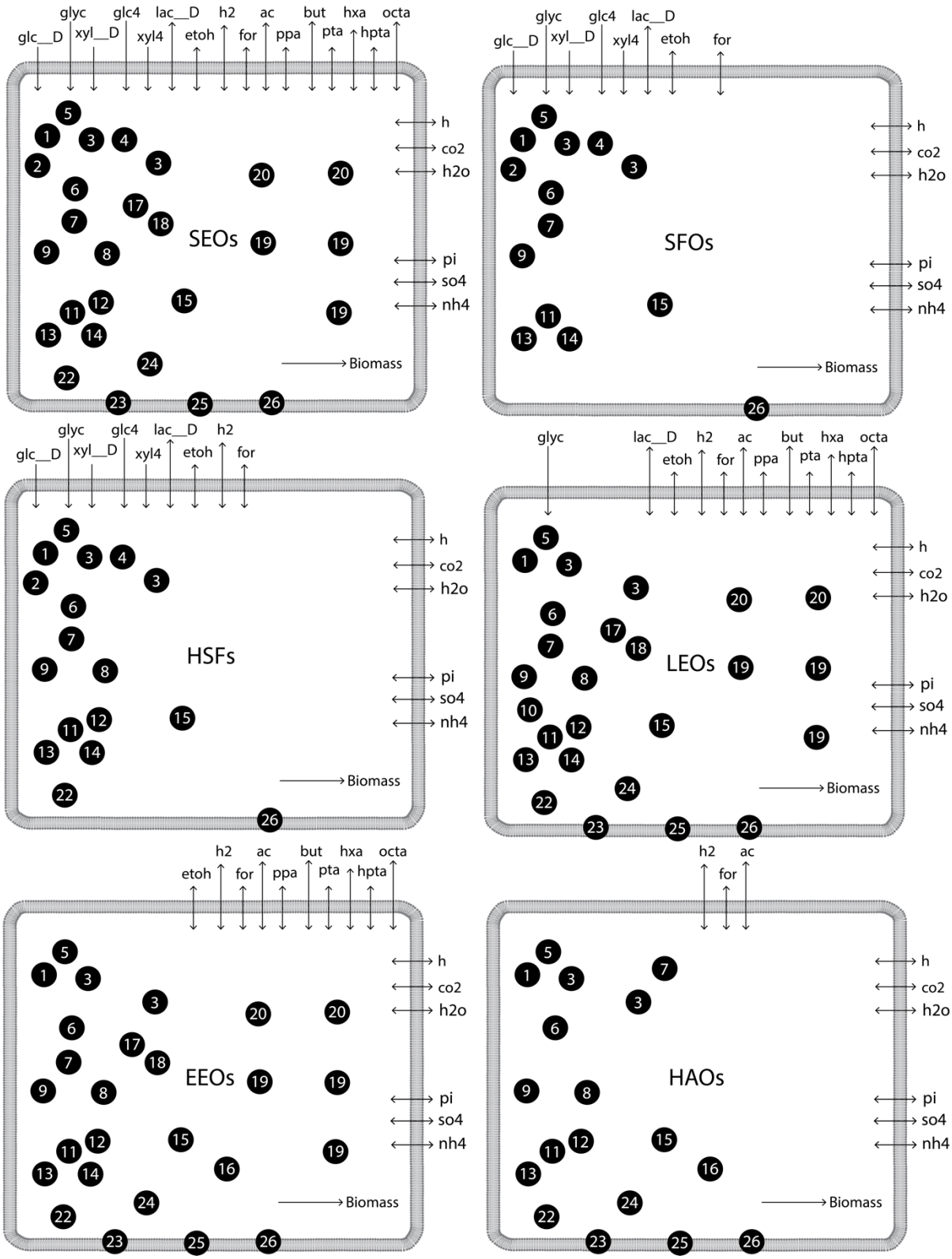


Figure 1. Summary of metabolic networks in iFermCell215. Metabolites with an asterisk are required for biomass production. Open hexagons indicate consumption of ATP (orange), NADH (green), NADPH (grey), and reduced ferredoxin (red). Closed hexagons indicate production of these compounds and consumption of ADP (orange), NAD+ (green), NADP+ (grey), and oxidized

681 ferredoxin (red). For reversible reactions, larger arrowheads indicate the reaction directionality
682 corresponding with consumption and/ or production of ATP, NADH, NADPH and reduced
683 ferredoxin.. Dashed lines indicate multiple reactions. Key reactions are indicated according to the
684 abbreviation used in iFermCell215 (**Supplementary File 3**). Pathways are identified by number:
685 (1) upper glycolysis; (2) phosphoketolase metabolism of fructose; (3) non-oxidative pentose
686 phosphate pathway; (4) phosphoketolase metabolism of xylulose; (5) glycerol utilization; (6) lower
687 glycolysis; (7) gluconeogenesis; (8) pyruvate formate lyase; (9) lactate metabolism; (10) pyruvate
688 flavodoxin oxidoreductase; (11) pyruvate dehydrogenase; (12) acetate metabolism; (13) ethanol
689 metabolism; (14) incomplete citric acid cycle with glyoxylate shunt; (15) anaplerotic metabolism
690 of phosphoenolpyruvate, pyruvate, and oxaloacetate; (16) ATP citrate lyase; (17) propionate
691 production via the acryloyl-CoA pathway; (18) propionate production via the methylmalonyl-CoA
692 pathway; (19) even-chain reverse β-oxidation; (20) odd-chain reverse β-oxidation; (21)
693 homoacetogenesis via the Wood Ljungdahl pathway; (22) ferredoxin-dependent hydrogenase; (23)
694 proton-translocating hydrogenase; (24) electron confurcating hydrogenase; (25) proton-
695 translocating ferredoxin: NAD^+ oxidoreductase; (26) NAD(P) transhydrogenase. For clarity, not
696 all reactions or metabolites are shown. More information on the reactions and metabolites
697 contained in iFermCell215 is provided in **Supplementary File 3**.


glc__D
glyc
xyl__D
glc4
xyl4
lac__D
etoh
h2
for
ac
ppa
but
pta
hxa
hpta
octa
h
co2
h2o
pi
so4
nh4
Biomass
SEOs
SFOs
HSFs
LEOs
EEOs
HAOs

698

699 **Figure 2.** Summary of metabolic networks for six functional guilds contained in iFermGuilds789.
700 Refer to **Fig. 1** for details on pathways included in the iFermGuilds789 model. More information
701 on the reactions and metabolites contained in the iFermGuilds789 model is provided in
702 **Supplementary File 4**.

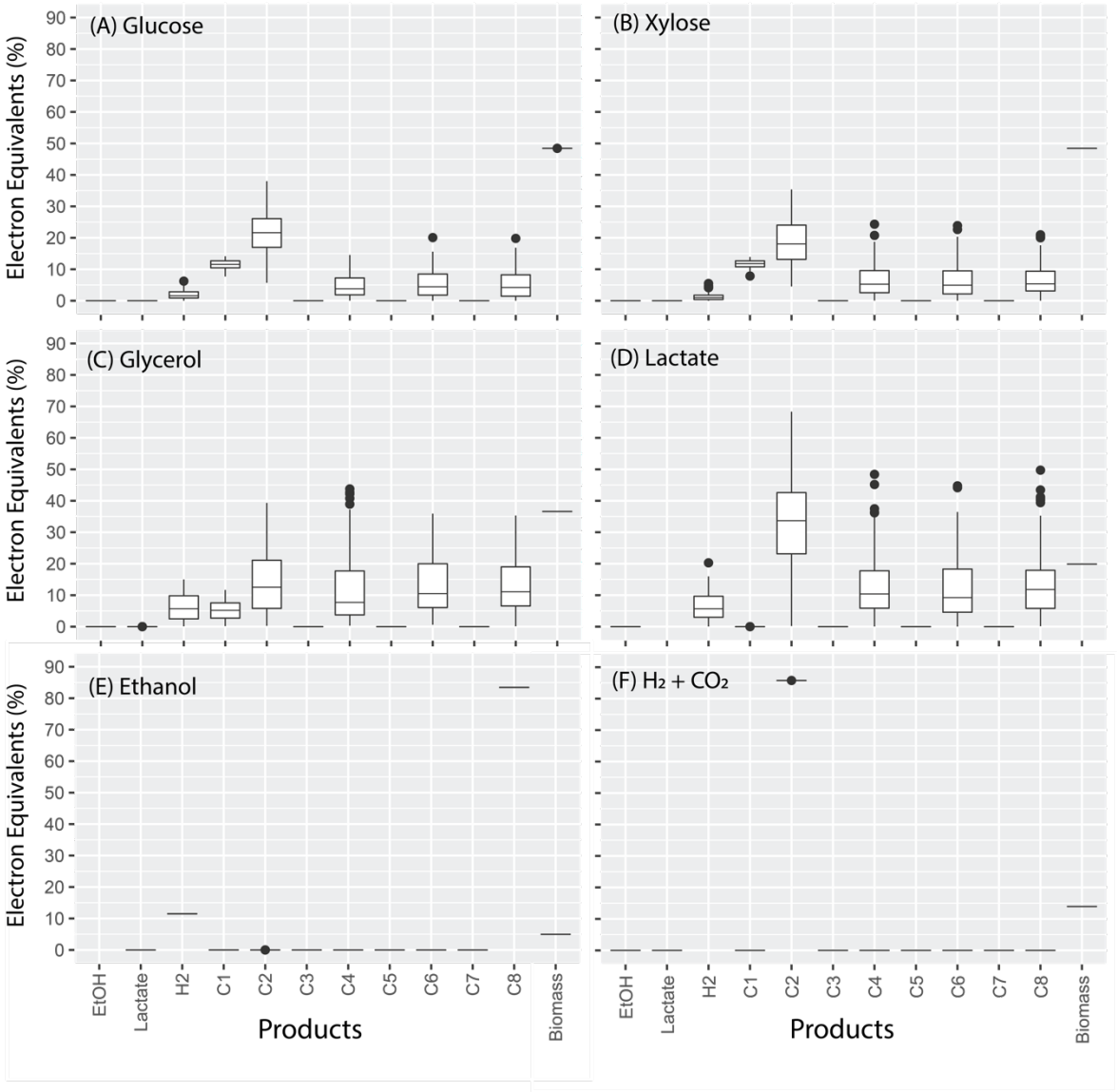
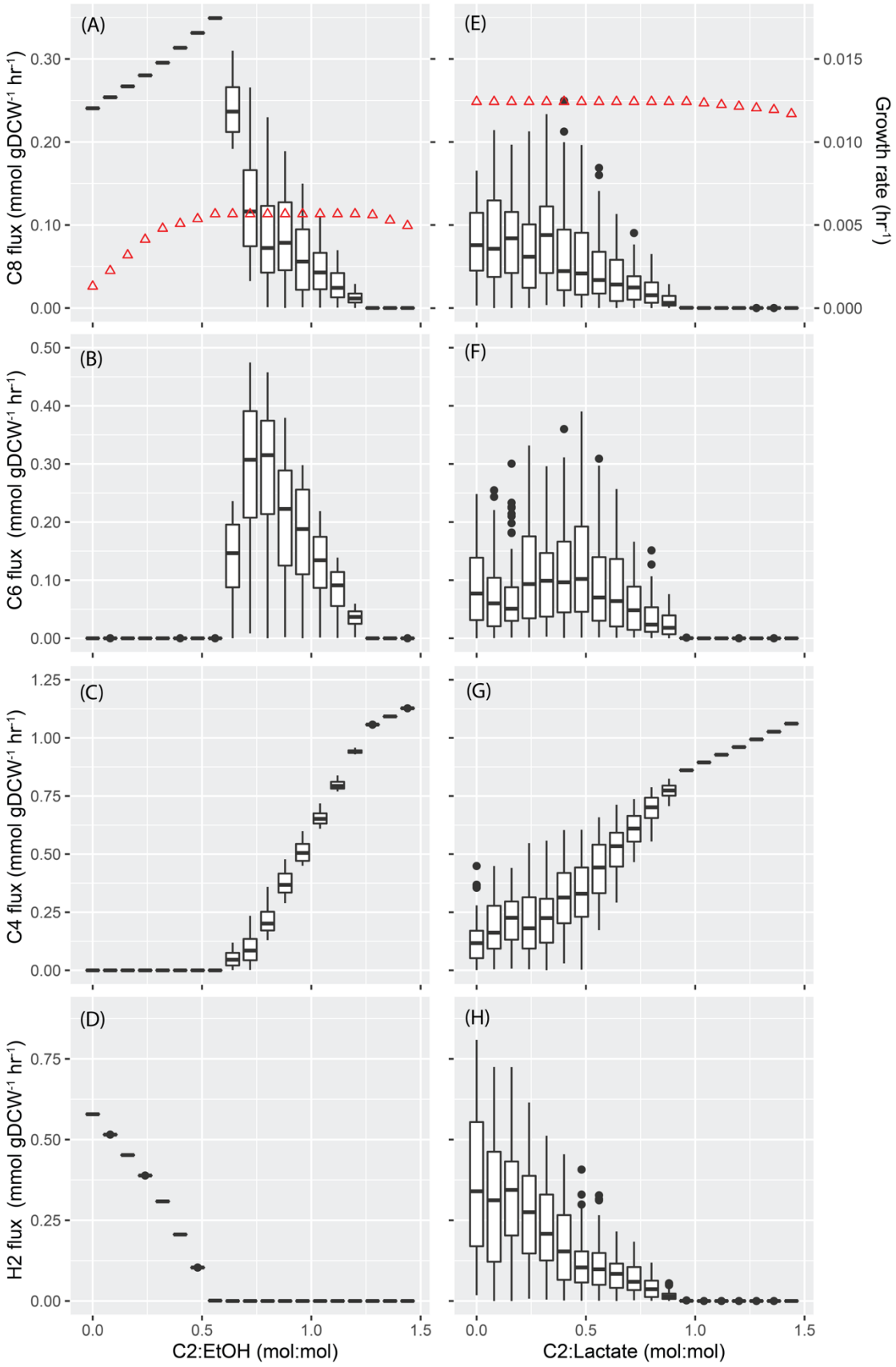
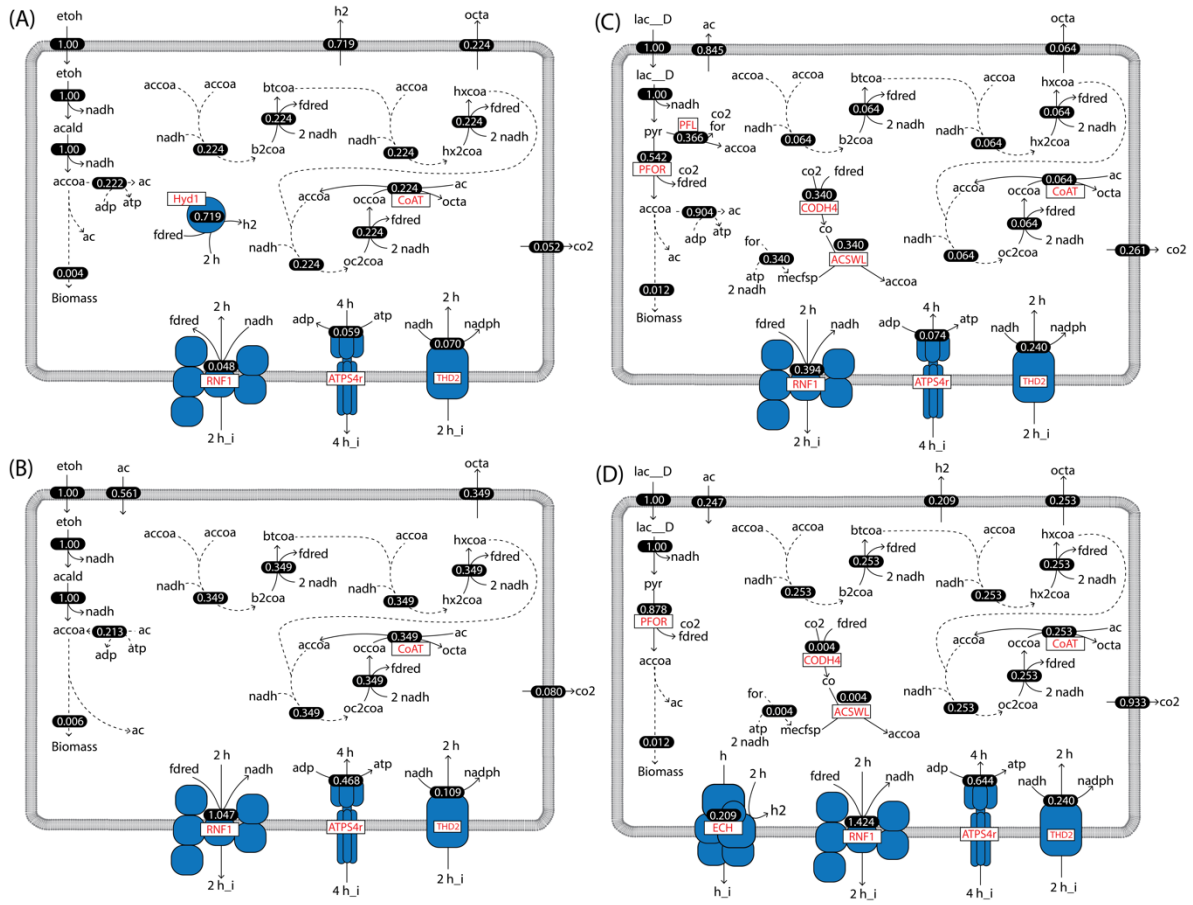


Figure 3. Flux sampling results for iFermCell215 with objective function to optimize biomass production when fed substrates indicated within the plots: (A) glucose, (B) xylose, (C) glycerol, (D) lactate, (E) ethanol, (F) H_2 and CO_2 (with H_2 as the limiting substrate). Box plots show the results of 101 simulations. Median values are indicated by a horizontal line within the box, and the top and bottom of the box indicate the 25th and 75th percentile, respectively. Whiskers indicate values within 1.5 times the interquartile range and closed circles indicate values outside of this range.


(A)
(E)
(B)
(F)
(C)
(G)
(D)
(H)
C8 flux (mmol gDCW-1 hr-1)
C6 flux (mmol gDCW-1 hr-1)
C4 flux (mmol gDCW-1 hr-1)
H2 flux (mmol gDCW-1 hr-1)
Growth rate (hr-1)
C2:EtOH (mol:mol)
C2:Lactate (mol:mol)

711

712 **Figure 4.** Flux sampling results for iFermCell215 when consuming ethanol and acetate (A-D) or
713 lactate and acetate (E-H) at different molar ratios of acetate-to-ethanol (C2:EtOH) or acetate-to-
714 lactate (C2:Lactate). Results are shown for fluxes of octanoate (C8), hexanoate (C4), butyrate (C4)
715 and H_2. These results are shown as box plots, as described previously (Figure 3), representing the
716 results of 101 simulations for each modeled ratio. The predicted maximum growth rate (hr^{-1}) is
717 shown for ethanol (A) and lactate (E) as red triangles plotted on the secondary y-axis.
718
719



720

721 **Figure 5.** Predicted pathways of ethanol utilization by iFermCell215 using pFBA when (A)
722 ethanol is the sole substrate; (B) ethanol and acetate are co-substrates; (C) lactate is a sole substrate;
723 and (D) lactate and acetate are co-substrates. Flux values (shown as white text) have units of mmol
724 gDCW^{-1} hr^{-1}. For both scenarios, the model was constrained to uptake 1 mmol gDCW^{-1} hr^{-1} of
725 ethanol or 1 mmol gDCW^{-1} hr^{-1} of lactate. For scenarios with acetate co-utilization, acetate was
726 provided in excess. Key reactions indicated by red text include a ferredoxin hydrogenase (Hyd1),
727 a CoA transferase (CoAT), a proton-translocating ferredoxin: NAD+ oxidoreductase (RNF1), ATP
728 synthase (ATPS4r), NADH:NADP transhydrogenase (THD2), and energy-conserving
729 hydrogenase (ECH). Metabolite abbreviations are and are defined in **Supplementary Data File**
730 **2**.

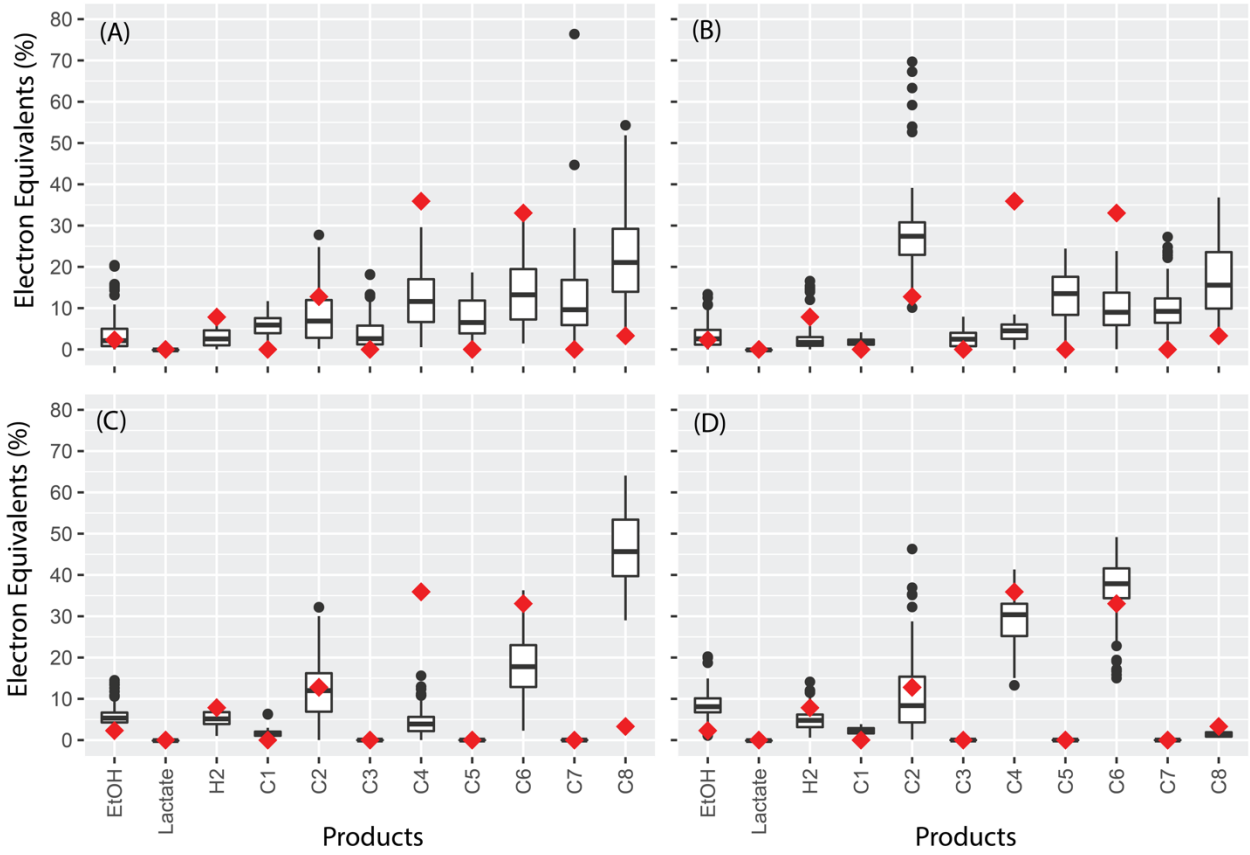


Figure 6. Predicted and observed product formation for a mixed culture fermentation bioreactor fed lignocellulosic biorefinery residues. (A) predictions with iFermCell215; (B) predictions with iFermGuilds789; (C) predictions with iFermGuilds789 with reactions knocked-out based on metagenomic analyses of the microbial community; (D) predictions with iFermGuilds789 with additional constraints for maximum fluxes through acyl-CoA and CoAT reactions for C6 and C8 production in SEOs and LEOs

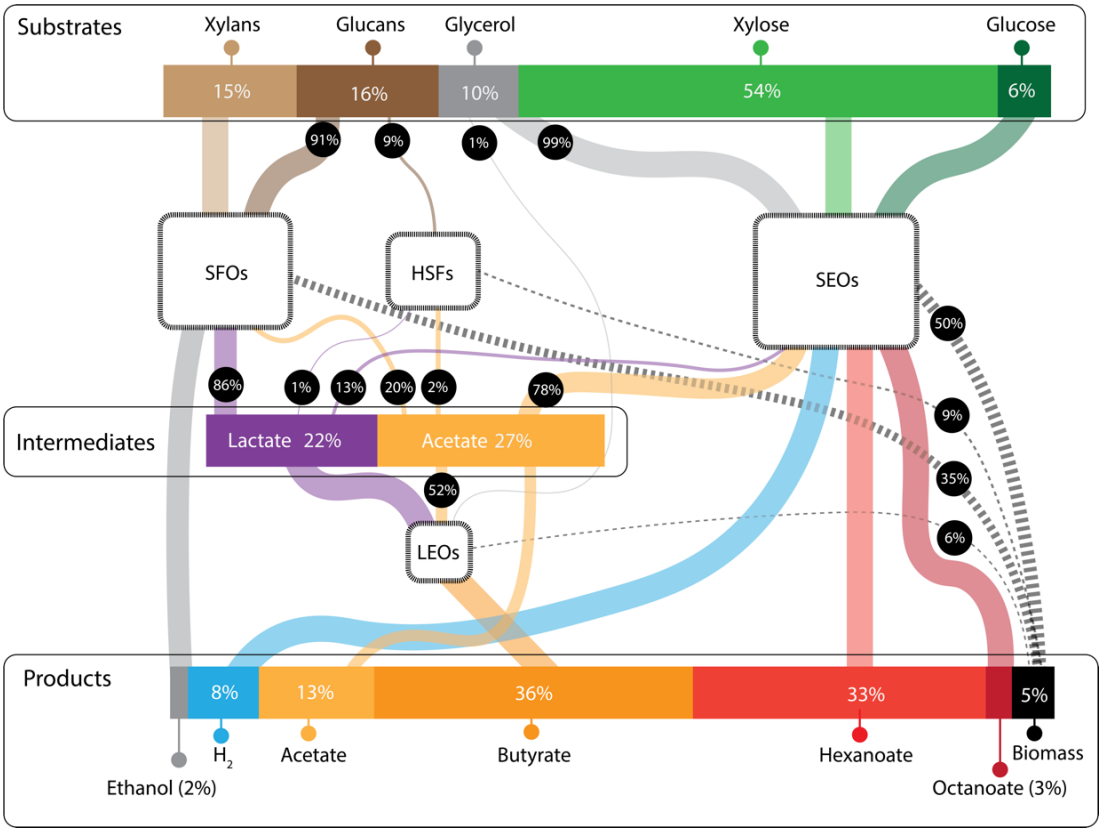


Figure 7. Predicted substrate flow through a mixed culture fermentation producing a mixture of MCFA from a complex feedstock using iFermGuilds789. Values were obtained using pFBA while constraining the products according to observed bioreactor performance. The relative abundance of each functional guild was assumed based on reactor DNA abundance. The relative amount of eeq in the substrates, products, and intermediates are also provided. Line width indicates the relative amount of each substrate consumed and product created by each of the functional guilds. H_2 was not measured, so for these simulations it was assumed that the difference between eeq in the substrate and eeq in the measured soluble and insoluble products represented H_2 production. In this simulation, 22% of the eeq in conversion residue was predicted to be directed to lactate as an intermediate and 27% to acetate.

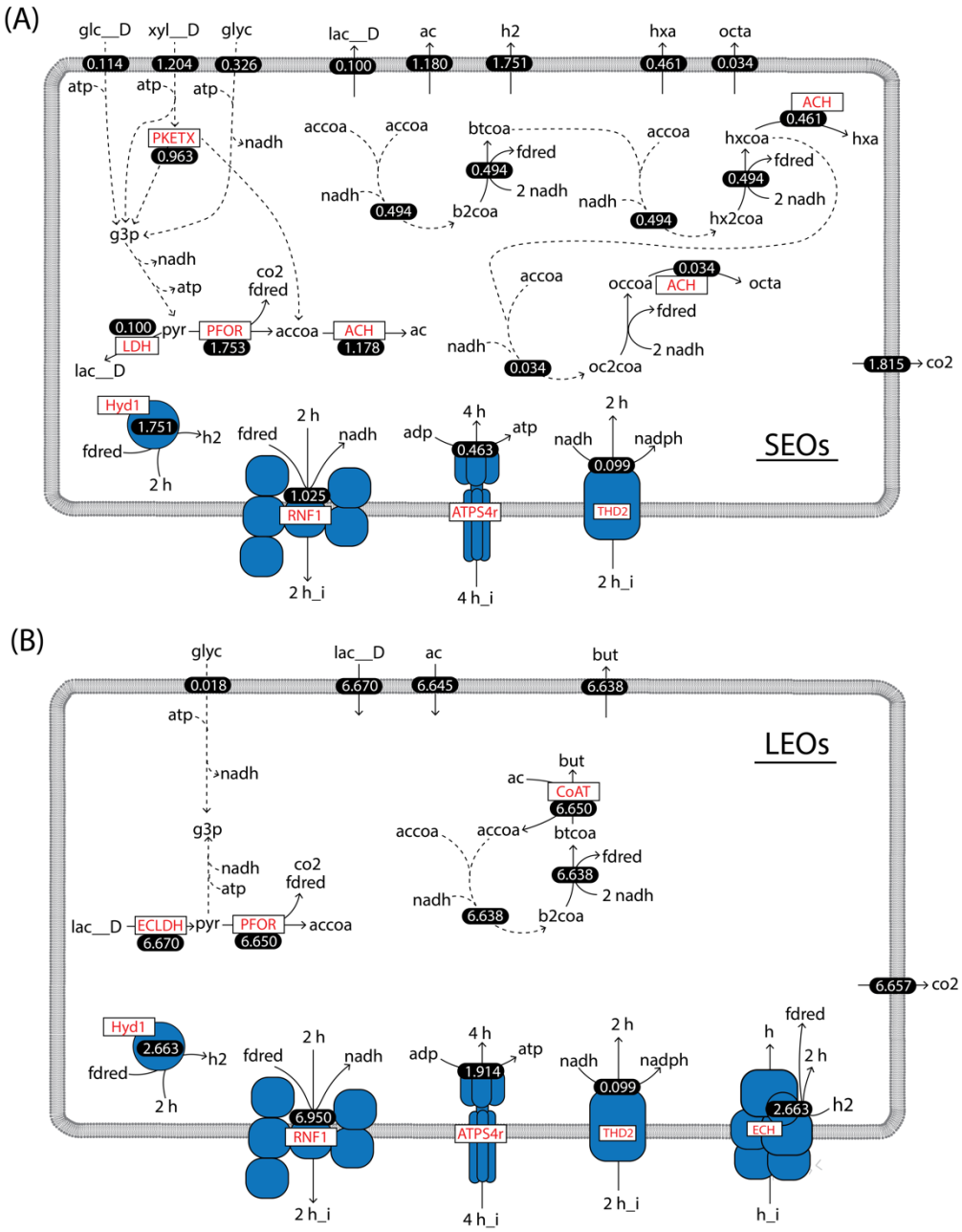


Figure 8. Predicted chain elongation metabolism by SEOs and LEOs using iFermGuilds789. Predicted fluxes values are shown as white numbers and have units of mmol gDCW^{-1} hr^{-1}. (A) SEOs are predicted to produce all of the hexanoate (hxa) and octanoate (octa) in the community and use an acyl-CoA hydrolase (ACH) reaction as the terminal step of reverse β-oxidation. (B) LEOs are predicted to produce all of the butyrate in the community while using a CoA transferase

758 (CoAT). LEOs are predicted to recycle all of the H_2 that they produce via an energy conserving

759 hydrogenase.