A gap-filling algorithm for prediction of metabolic interactions in microbial communities

Dafni Giannari[1], Cleo Hanchen Ho[2], Radhakrishnan Mahadevan[1,3*]

1 Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ontario, Canada
2 Viome Inc, San Diego, California, United States of America
3 The Institute of Biomaterials & Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

* krishna.mahadevan@utoronto.ca

Abstract

The study of microbial communities and their interactions has attracted the interest of the scientific community, because of their potential for applications in biotechnology, ecology and medicine. The complexity of interspecies interactions, which are key for the macroscopic behavior of microbial communities, cannot be studied easily experimentally. For this reason, the modeling of microbial communities has begun to leverage the knowledge of established constraint-based methods, which have long been used for studying and analyzing the microbial metabolism of individual species based on genome-scale metabolic reconstructions of microorganisms. A main problem of genome-scale metabolic reconstructions is that they usually contain metabolic gaps due to genome misannotations and unknown enzyme functions. This problem is traditionally solved by using gap-filling algorithms that add biochemical reactions from external databases to the metabolic reconstruction, in order to restore model growth. However, gap-filling algorithms could evolve by taking into account metabolic interactions among species that coexist in microbial communities. In this work, a gap-filling method that resolves metabolic gaps at the community level was developed. The efficacy of the algorithm was tested by analyzing its ability to resolve metabolic gaps on a synthetic community of auxotrophic *Escherichia coli* strains. Subsequently, the algorithm was applied to resolve metabolic gaps and predict metabolic interactions in a community of *Bifidobacterium adolescentis* and *Faecalibacterium prausnitzii*, two species present in the human gut microbiota, and in an experimentally studied community of *Dehalobacter* and *Bacteroidales* species of the ACT-3 community. The community gap-filling method can facilitate the improvement of metabolic models and the identification of metabolic interactions that are difficult to identify experimentally in microbial communities.

Author summary

Microbes compose the most abundant form of life on our planet and they are almost never found in isolation, as they live in close association with one another and with other organisms. The metabolic capacity of individual microbial species dictates their ways of interaction with other species as well as with their environment. The metabolic interactions among microorganisms have been recognised as the driving force for the emergence of the properties of microbial communities. For this reason, understanding

the effect of microbial metabolism on interspecies metabolic interactions is detrimental for the study of microbial communities. This study can be benefited by metabolic modeling and the insights offered by constraint-based methods which have been developed for interrogating metabolic models. In this paper, we present an algorithm that predicts cooperative and competitive metabolic interactions between species, while it resolves metabolic gaps in their metabolic models in a computationally efficient way. We use our community gap-filling algorithm to study microbial communities with interesting environmental and health-related applications.

Introduction

Microorganisms form the most abundant group of living organisms on our planet. In nature microorganisms do not live in isolation, but in close association with one another, forming microbial communities. The value of microbial communities for biotechnology, the environment and human health has been profound due to their potential for use in industrial bioprocesses for the production of valuable chemicals [1], their role in biogeochemical cycles [2], and their effects on human health through the human microbiome [3].

The study of microbiomes has been extensively limited to the taxonomic classification of species and their correlation with different phenotypes. However, such correlations do not offer any mechanistic explanation on how the interactions among microbes and their environment form the observed phenotypes of the ecosystem [4]. A way to elucidate metabolic interactions in microbial communities comes from the use of a set of mathematical and computational techniques, called constraint-based methods, that make use of genome-scale metabolic models [5,6]. Constraint-based methods have been used extensively for the study of the metabolic functions of individual microorganisms as well as for strain design in metabolic engineering [7]. In the context of microbial communities, methods like SteadyCom [8], OptCom [9], d-OptCom [10], DMMM (Dynamic Multispecies Metabolic Modeling) [11], and COMETS (Computation Of Microbial Ecosystems in Time and Space) [12] give the opportunity to evaluate growth rates and metabolic interactions of community members under various conditions.

Genome-scale metabolic models (GSMMs) can be reconstructed automatically from isolated genomes or pangenomes of specific organisms with a variety of tools [13], like ModelSEED [14] and Kbase [15], that create gene - protein - reaction (GPR) associations. However, genomes are often fragmented and contain misannotated genes [16], while the databases with information about enzyme functions and biochemical reactions used in the reconstruction process are not highly curated [13]. These problems lead to the creation of models with metabolic gaps, mass and charge imbalances and thermodynamic infeasibilities. Metabolic gaps are solved with the use of gap-filling methods, which are an indispensable part of the reconstruction process of organism-specific metabolic models [17]. The first published gap-filling algorithm was GapFill [18–20], which was formulated as a Mixed Integer Linear Programming (MILP) problem that identified dead-end metabolites and added reactions from the MetaCyc database [21] in the metabolic network. Some platforms for metabolic network reconstruction and curation, like gapseq [22] and AMMEDEUS [23], make use of more computationally efficient gap-filling algorithms formulated as Linear Programming (LP) problems. Other methods [24–26], including gapseq [22] and CarveMe [27], take into account genomic or taxonomic information in order to decide which biochemical reactions to add to the metabolic network. Other notable methods are OMNI [28] and GrowMatch [29], that try to maximize the consistency of the model with experimentally observed fluxes and growth rates respectively, while OptFill [30] is a method formulated

to simultaneously solve metabolic gaps and thermodynamically infeasible cycles. Overall, most of the gap-filling methods to date are constraint-based methods that resolve metabolic gaps by adding biochemical reactions from a reference database, like ModelSEED [31], MetaCyc [21], KEGG [32], or BiGG [33], to the metabolic reconstruction of a specific organism.

Curated GSMMs of individual organisms can be combined in order to create a community model. However, GSMMs of organisms that naturally live in complex microbial communities, cannot be easily curated individually, and they often do not realistically represent the organism's metabolic potential after the gap-filling process is completed. This problem is mainly created because of the restricted amount of physiological information that can be collected experimentally for members of complex microbial communities. Microorganisms cannot be easily cultivated individually due to their complex metabolic interdependencies with other community members. This lack of physiological information is especially acute in the case of metagenomes from the environment or from enrichments. In this paper, we propose a gap-filling algorithm that resolves metabolic gaps in microbial communities, while considering metabolic interactions in the community. The proposed community gap-filling method combines incomplete metabolic reconstructions of microorganisms that are known to coexist in microbial communities and permits them to interact metabolically during the gap-filling process. Therefore, community gap-filling can be a useful method not only for resolving metabolic gaps of GSMMs, but also for predicting non-intuitive metabolic interdependencies in microbial communities. Our algorithm also takes advantage of a formulation that ensures decreased solution times for the community gap-filling problem. Considering the unexplored metabolic potential of many natural microbial communities, as well as the difficulty to produce highly-curated metabolic reconstructions for microorganisms, we believe that this method will benefit the study of complex microbial communities.

The described community gap-filling method was applied to three independent case studies that demonstrate its ability to restore growth in metabolic models and predict both cooperative and competitive metabolic interactions between them by adding the minimum possible number of biochemical reactions from a reference database to the models. The three case studies involve a community of two *Escherichia coli* strains created for testing purposes, and the more complex communities of *Bifidobacterium adolescentis* with *Faecalibacterium prausnitzii*, and *Dehalobacter* with *Bacteroidales* species, which have interesting applications.

First, we demonstrate that our method successfully restores growth in a synthetic community comprised of two auxotrophic *Escherichia coli* strains: an obligatory glucose consumer and an obligatory acetate consumer. This community represents the well-known phenomenon of acetate cross-feeding that emerges among *E. coli* strains which grow in homogeneous environments containing glucose as the sole carbon source [34, 35].

The community gap-filling method was also used to study the codependent growth of *Bifidobacterium adolescentis* and *Faecalibacterium prausnitzii*, two well-known bacterial members of the human gut microbiome, the most diverse group of microorganisms found in the human body [36, 37]. One of the main functions of the intestinal microbiota is the anaerobic saccharolytic fermentation of complex carbohydrates such as dietary fibers [38], which leads to the production of short chain fatty acids (SCFAs) like acetate, propionate and butyrate. SCFAs are known for their various beneficial effects on the mammalian gut and metabolism [39, 40], and in particular butyrate is known for its ability to function as an energy source for the epithelial cells of the gut and works against oxidative stress and inflammation [41]. A major butyrate producer in the gut of healthy adults is the bacterium *Faecalibacterium prausnitzii* [42, 43]. The presence of *F.*

prausnitzii has been connected with anti-inflammatory effects in mice [44], while its abundance is significantly decreased in some diseases of the gastrointestinal tract [45–47]. For these reasons, it is believed to have the ability to function as a probiotic and as a marker for inflammatory bowel diseases and colon cancer [48]. *F. prausnitzii* has been identified as a commensal, strictly anaerobic bacterium of the human gut that performs saccharolytic fermentation. It can utilize a variety of carbon sources from glucose, fructose, and fructo-oligosaccharides to complex molecules such as pectin and N-acetylglucosamine [49], and its typical fermentation products are butyrate, formate, lactate, and acetate [50]. *F. prausnitzii* is known for developing both competitive and syntrophic interactions with other species that live in the gut, as it competes for common carbon sources, but also has the ability to consume acetate produced by other species and convert it to butyrate. These interactions have been observed in cocultures of *F. prausnitzii* with bacterial species of the genus Bifidobacterium [51, 52]. Bifidobacterial species belong to the phylum Actinobacteria, and are commensal, obligate anaerobes of the human gut, that perform saccharolytic fermentation. They can metabolize a variety of carbohydrates [53], and their most common fermentation products are acetate, formate, lactate, succinate, and ethanol [54–56]. The cogrowth of the species *Faecalibacterium prausnitzii* and *Bifidobacterium adolescentis* has been studied both experimentally [57, 58] and computationally [59], and it has been shown that their metabolic interaction has the potential to enhance butyrate production [57]. For the community of *F. prausnitzii* and *B. adolescentis*, the community gap-filling method managed to predict metabolic interactions and SCFA production that were previously reported [57].

Finally, the community gap-filling method was used to replicate experimental results for the interaction between *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50, which are the two main members of the ACT-3 community. The ACT-3 community is a microbial community found in soils and underground waters that are usually contaminated with chlorinated compounds. The study of this community has garnered scientific interest in recent years, as it has the ability to degrade chlorinated environmental pollutants [60–62]. The ACT-3 community is mainly composed of bacterial species that belong in the genera Dehalobacter, with relative abundance up to 80%, Bacteroides, and Clostridium [63]. The dechlorinating ability of the ACT-3 community mainly comes from the ability of *Dehalobacter* species to perform dehalogenation, i.e., anaerobic respiration with the use of chloroform or other chlorinated compounds as electron acceptors [64–68]. In this study, we use our metabolic reconstructions for *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50, in order to simulate experimental results for the cogrowth of the two most abundant species of the ACT-3 community [63] and improve our models.

Materials and methods

The community gap-filling method

The general principle behind the proposed method is demonstrated in Fig 1. Our algorithm creates a compartment for each organism of the simulated microbial community, by combining the biochemical reactions of the organism's metabolic model with those that come from a reference database. In this study, we created a curated database using the biochemical reactions available in BiGG [33]. While continuing to use the BiGG nomenclature, we removed all the biomass equations, reactions found only in eukaryotic organisms, and reactions with mass imbalances. Then, the organism compartments were allowed to exchange metabolites with one another and with their environment through a common metabolite pool where extracellular compounds can

accumulate or deplete. In this way, each organism can synthesize biomass only by using its intracellular metabolites, and metabolites defined in the growth medium or produced by the other organisms of the community.

Fig 1. Graphical representation of the community gap-filling method. The metabolic reconstructions of two individual organisms (blue and orange respectively) are allowed to exchange metabolites and interact with the environment through a common metabolite pool (green). **a)** The algorithm adds biochemical reactions from a reference database (dark purple) to the community model. **b)** The community gap-filling algorithm is an MILP problem with the objective to add the minimum number of database reactions to the community model in order to restore biomass production in the individual metabolic models, while it satisfies some constrains for the reaction fluxes. The fate of each database reaction is controlled by a binary variable, which takes the value of 1 if the database reaction is added to the community model, and 0 otherwise.

As we can see in Fig 1, our community model consists of the organism compartments and an exchange compartment which is the common metabolite pool. This community model is used to formulate an MILP problem that predicts the minimum number of reactions that need to be added from the reference database to the community model in order to achieve a user-determined growth rate for the organisms that make up the community. The mathematical formulation of our community gap-filling algorithm for a microbial community of N organisms is:

$$\text{minimize} \sum_{n \in N} \sum_{j \in J^n_{\text{Database}}} y^n_j \tag{1}$$

subject to

$$
\left[
\begin{array}{ll}
\sum_{j \in J^n} S^n_{ij} \, v^n_j = 0, & \forall \, i \in I^n \\[2mm]
lb^n_j \le v^n_j \le ub^n_j, & \forall \, j \in J^n_{\text{Model}} \\[1mm]
y^n_j \, lb^n_j \le v^n_j \le y^n_j \, ub^n_j, & \forall \, j \in J^n_{\text{Database}} \\[1mm]
v^n_{\text{Biomass}} \ge v^{n,\,\text{min}}_{\text{Biomass}} & \\[1mm]
y^n_j \in \{0, 1\}, & \forall \, j \in J^n_{\text{Database}}
\end{array}
\right] \; \forall \, n \in N
\quad
\begin{array}{l}
(2) \\[4mm]
(3) \\[1mm]
(4) \\[1mm]
(5) \\[1mm]
(6)
\end{array}
$$

$$
- v^c_{\text{ex(i)}} + \sum_{n \in N} v^n_{\text{ex(i)}} = 0, \qquad\qquad \forall \, i \in I^c \, , \tag{7}
$$

where I^n and J^n represent the number of metabolites and reactions, respectively, in the nth microorganism compartment, while I^c represents the exchanged metabolites in the common metabolite pool.

Eq (1) is the objective function of the MILP problem which minimizes the total number of biochemical reactions that are added from the database to the metabolic models of the organisms composing the community. Eq (2) - Eq (7) are the constraints of the optimization problem. Specifically, Eq (2) and Eq (7) show that the metabolite pools belonging to each organism compartment and to the common exchange compartment respectively, are at steady state. Eq (3) shows the constrained fluxes of the reactions originally belonging to the organism models. The lower and upper bounds for the model reactions are selected in order to represent the growth of the organism under the desired conditions. Eq (5) represents the constraints for the fluxes through the biomass reactions of the organism models. The minimum threshold for the growth rate of each

organism is defined based on existing information for the growth of the microorganism under the simulated conditions. Eq (4) implies that the fluxes of the reactions coming from the database are constrained between a lower and an upper bound, but a binary variable controls whether each database reaction will be added to the corresponding model. The binary variables, which are also used in the objective function, are defined in Eq (6). Each binary variable takes the value 1 if the database reaction is added to the community model, and the value 0 otherwise. It is noted that all the reactions from the database are initially considered reversible. Then, the minimum and maximum possible fluxes for these reactions are calculated from Flux Variability Analysis (FVA) [69] when the organism compartment is made. The FVA results are used as constraints for the database reaction fluxes, which also gives the opportunity to remove the reactions with both minimum and maximum flux zero. Using this approach, the solution space and time of the final gap-filling problem is reduced.

The simulations were performed on a machine with two 22-core Intel® Xeon® Gold 6238 @ 2.10GHz and 768GB of RAM, running CentOS 7. For the generation of the results presented in this paper, the MILP problem was formulated and solved in MATLAB 2017b with CPLEX Interactive Optimizer 12.8 and CobraToolbox 3, and alternative solutions for the problem were calculated with an upper time limit of two hours. The alternative solutions are used in each of the presented case studies in order to identify reactions added to the metabolic models of the participating microorganisms as well as metabolic interactions that appear more frequently in the solutions.

Sources of metabolic models

Core *E. coli* community

S1 Appendix explains how we used the *E. coli* core model from BiGG [70] for the creation of two *E. coli* strains, one with the ability to use glucose as a substrate and another with the ability to use only acetate excreted from the first strain. S1 Table and S2 Table contain information about reactions that were deleted from the core model and constraints for the exchange reactions that were implemented in order to create the models of the *E. coli* glucose utilizer and the *E. coli* acetate utilizer respectively. In this study, the selected lower bounds for the growth rates of the *E. coli* glucose utilizer and the *E. coli* acetate utilizer, were 0.9 and 0.09 h^{-1}, respectively, based on existing information about acetate cross-feeding between *E. coli* strains [34, 71, 72].

Bifidobacterium adolescentis and *Faecalibacterium prausnitzii* community

For our simulations, we downloaded the metabolic models for the strains *Bifidobacterium adolescentis* ATCC 15703 and *Faecalibacterium prausnitzii* A2-165 from the Virtual Metabolic Human Database [73]. The models are part of the AGORA version 1.03 [74], a collection of metabolic reconstructions of 773 members of the human gut microbiota. Both models are known to contain gaps in the pathways for the biosynthesis of some amino acids and vitamin B, as well as in their central carbon metabolism [73]. In this study we simulate the ability of the two models to use glucose as a substrate and exchange acetate and amino acids under anaerobic conditions. The applied constraints for the exchange reactions of the two models can be seen in S3 Table and S4 Table respectively. The lower bounds for the growth rates of the *B. adolescentis* ATCC 15703 and the *F. prausnitzii* A2-165, were set to 0.6 and 0.4 h^{-1}, respectively. The constraints used for the exchange reactions of the organic acids, the amino acids, and the growth rates are based on experimental observations of *B. adolescentis* and *F. prausnitzii* strains cogrowth [57], and the study of the metabolic potential of the strain *F. prausnitzii* A2-165 [50].

ACT-3 community

For this case study, we reconstructed the metabolic models of the two most abundant species of the ACT-3 community, *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50. S2 Appendix contains information about the reconstruction process for the two models. We allowed *Dehalobacter* sp. CF to consume chloroform and *Bacteroidales* sp. CF50 to consume lactate, while the two models grow anaerobically in the same medium and are allowed to exchange organic acids and amino acids. Based on the experimental study for the coculture [63], after 18 days, the community consumes 0.5 mM of chloroform and 0.75 mM of lactate, while the total cell density is around $6 \cdot 10^7$ cells·mL^{-1}, and the relative abundance of the *Dehalobacter* in the community is 66% on average. This information was used in order to calculate the uptake rates of chloroform and lactate by the community as shown in S5 Table. The reaction constraints for the two models are shown in S6 Table and S7 Table. For our simulations, we used 0.1 and 0.01 d^{-1} as lower bounds for the growth rates of the *Dehalobacter* sp. CF and the *Bacteroidales* sp. CF50, respectively. All the constraints applied for the exchange reactions and the growth rates are representing experimental data from the cogrowth of the *Dehalobacter* and *Bacteroidales* species [63].

Results

We used our community gap-filling method in order to add biochemical reactions to metabolic models of organisms that contain incomplete pathways. Since the method simulates the cogrowth of organisms that are known to coexist in microbial communities, it is used for identifying both expected and unexpected metabolic interactions in the community. In this paper, we tested the potential of the method on an artificial *E. coli* community, and then, we applied it to two communities with significance for the health of the human gut and for the bioremediation of soil and water, respectively. For each microbial community, we discuss the best solution calculated by the algorithm, i.e. the solution that adds the minimum number of biochemical reactions to the models, as well as the patterns emerging from the best alternative solutions suggested by the algorithm. The MATLAB files with all the solutions calculated by the community gap-filling for each one of the studied microbial communities can be found in the Supplementary Material. A summary of the characteristics of the used metabolic models before and after the application of the community gap-filling can be seen in Table 1. The SBML files of all the used models are available on our GitHub.

Table 1. Characteristics of the metabolic models used in the study before and after the application of the community gap-filling method. The algorithm simulated the aerobic growth of a community of an *E. coli* glucose utilizer and an *E. coli* acetate utilizer on glucose, the anaerobic growth of a community of the strains *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165 on glucose, and the anaerobic growth of a community of the species *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50 on media with lactate and chloroform.

Model	Before community gap-filling			After community gap-filling		
	metabolites	reactions	growth rate*	metabolites	reactions	growth rate*
E. coli glucose utilizer	72	85	1.15 h^{-1}	72	86	0.9 h^{-1}
E. coli acetate utilizer	72	83	0 h^{-1}	72	85	0.09 h^{-1}
B. adolescentis ATCC 15703	991	1129	0 h^{-1}	991	1134	0.6 h^{-1}
F. prausnitzii A2-165	1047	1225	0 h^{-1}	1055	1233	0.4 h^{-1}
Dehalobacter sp. CF	978	1087	infeasible	978	1088	0.1 d^{-1}
Bacteroidales sp. CF50	964	967	0 d^{-1}	965	969	0.01 d^{-1}

*in the community conditions

Core *E. coli* community

In the beginning, we tested if our community gap-filling method can restore growth in an artificial microbial community of an *E. coli* glucose utilizer and an *E. coli* acetate utilizer strain, that is allowed to grow aerobically on glucose. Our community was made after the deletion of 12 reactions from the original models of *E. coli* core metabolism (Fig 2) as discussed in S1 Appendix. As expected, the community gap-filling algorithm, instead of adding back all the knockouts that we made, suggested the addition of the minimum possible number of reactions in the community in order to restore its function. The first solution calculated by the algorithm suggested the addition of one reaction to the *E. coli* glucose utilizer model and two reactions to the *E. coli* acetate utilizer model (Fig 2). More specifically, the pyruvate oxidase (POX) reaction was added to the model of the *E. coli* glucose utilizer to produce acetate from pyruvate. The additions to the model of the *E. coli* acetate utilizer include citrate synthase (CS), which was one of the initial knockouts, and fructose-6-phosphate utilizing phosphoketolase (PKETF), which converts acetyl-phosphate to fructose-6-phosphate. Apart from the reactions that were added from the database to the models, the community gap-filling algorithm suggested the activation of reactions that were already present in the models, but did not carry flux in the FBA solution of the original core model. More specifically, in order to sustain growth in the glucose utilizer the reactions of 6-phosphogluconate dehydratase (EDD) and 2-dehydro-3-deoxy-phosphogluconate aldolase (EDA) were activated to produce pyruvate from gluconate 6-phosphate, while pyruvate dehydrogenase (PDH) was also activated to convert pyruvate to acetyl-CoA. In the acetate utilizer the preexisting reactions of acetyl-CoA synthetase (ACS) and acetate kinase (ACKr) were activated in order to convert acetate to acetyl-CoA and acetyl-phosphate respectively, and malic enzyme (ME1) converted malate to pyruvate. The fluxes for all the reactions of the community, as well as a summative table of the reactions that were added from the database to the community can be found in S8 Table. A post gap-filling FBA for the two models shows that all the biomass precursors are replenished in both models. However, both of the models grow at suboptimal rates compared to the original core model (S1 Appendix and Table 1). For the glucose utilizer, the suboptimal growth rate of 0.9 h^{-1} was due to the export of acetate that led to a reduction in resources directed towards growth, whereas for the acetate utilizer the growth rate of 0.09 h^{-1} was caused by the use of reduced acetate as the only substrate.

Fig 2. Graphical representation of the core *E. coli* community after the application of the community gap-filling method. The central carbon metabolism of an *E. coli* core model (blue rectangular: Glycolysis, pink square: Pentose Phosphate Pathway, yellow circle: TCA cycle) was used in order to create two E. coli strains: one that consumes glucose (left) and one that consumes acetate (right), after the deletion of the reactions marked with red crosses. The best solution of the community gap-filling algorithm predicted the addition of the reactions represented by continuous blue arrows and the activation of the existing reactions represented by dashed blue arrows, in order to restore biomass production in the two models. The dashed black arrows show the exchange reactions for glucose and acetate in the community. The metabolites in bold represent biomass precursors. The numbered deleted reactions are: (1) PGM, (2) MALS, (3) SUCOAS, (4) PFL, (5) PTAr, (6) GLCpts, (7) CS. The numbered added and activated reactions are: (1) POX, (2) CS, (3) PKETF, (4) EDD and EDA, (5) PDH, (6) ACS, (7) ACKr, (8) ME1.

The fluxes from the ten best solutions calculated by the community gap-filling algorithm for the exchange reactions of the community (S9 Table) show that all the solutions predicted the ability of the *E. coli* glucose utilizer model to uptake glucose

and produce acetate that is used by the *E. coli* acetate utilizer model for aerobic or anaerobic growth. We can see that the acetate utilizing strain either shares oxygen with the glucose utilizing strain or ends up growing anaerobically due to oxygen depletion by the glucose utilizing strain in different solutions. Moreover, all the reactions that were added from the database to the community in the ten best solutions (S10 Table), carry realistic fluxes with values ranging from -10 to 15 mmol·gDW^{-1}·h^{-1}.

We further used our artificial *E. coli* community in order to estimate the computational efficiency of our method. One major innovation of our community gap-filling method is the use of FVA before the solution of the main MILP problem. FVA is performed in order to identify the minimum and maximum fluxes that are possible for every database reaction in each organism compartment of the community. The fluxes calculated by FVA are used as constraints for the database reactions before building the community model. As a result, the solution space of the community gap-filling MILP problem is smaller and therefore, it can be solved faster. Many of the gap-filling methods to date solve MILP problems, but have not made an effort to reduce the solution space. From a computational perspective, the search of optimality can be tedious for MILP problems, and it is known that smaller feasible solution spaces can accelerate the solution of the optimization problem. Regarding the community gap-filling algorithm, the solving time increases when bigger metabolic models and databases are used and more importantly when the metabolic models of more organisms are added to the problem. As we can see from S1 Fig, the solving time of the MILP problem after the application of FVA is increasingly smaller as more organisms are added in the community, compared to the problem that was formulated immediately without the use of FVA. These results suggest that the FVA preprocessing step improves the solution time up to 84% (from 28 min to 4.5 min) for a microbial community of two organisms.

Bifidobacterium adolescentis and *Faecalibacterium prausnitzii* community

In this case study, we applied the community gap-filling method to a community made from the models of the strains *Bifidobacterium adolescentis* ATCC 15703 and *Faecalibacterium prausnitzii* A2-165. The interaction between these two species has been studied experimentally and is considered beneficial since it promotes the production of the SCFA butyrate, which has antinflammatory effects in the gut [57]. Here, we simulated the anaerobic cogrowth of *B. adolescentis* and *F. prausnitzii* with glucose as the sole carbon source. None of the models can synthesize biomass in the simulated conditions, and the two models are allowed to produce SCFAs and exchange amino acids. The model of *F. prausnitzii* was permitted to either uptake or excrete acetate.

The community gap-filling algorithm restored community growth by predicting both competitive and cooperative metabolic interactions and adding database reactions to the models. The first calculated solution added five reactions to the model of *B. adolescentis* ATCC 15703 and eight reactions to the model of *F. prausnitzii* A2-165 (Table 2 and S11 Table). The reactions added to the models of *B. adolescentis* and *F. prausnitzii* participated in the biosynthesis of cofactors and amino acids necessary for growth. The first solution predicted that *B. adolescentis* and *F. prausnitzii* share the available glucose from the community medium, while *F. prausnitzii* uses part of the acetate exported from *B. adolescentis* (Fig 3). Moreover, both models export CO_2, *B. adolescentis* exports lactate and formate in addition to acetate, and *F. prausnitzii* exports butyrate, lactate and formate (Fig 3). This correlates with experimental observations [57]. Interestingly the model of *B. adolescentis* ATCC 15703 uses the bifid shunt pathway, which is centered around the key enzyme frutose-6-phosphate

erythrose-4-phosphate-lyase (F6PE4PL) in our model. The bifid shunt pathway is 334
unique in Bifidobacteria species that commonly use it for the efficient conversion of 335
carbohydrates, like glucose, to acetate and lactate [54, 55]. The biosynthesis of butyrate 336
from the model of *F. prausnitzii* A2-165 happens through the reaction of 337
butyryl-CoA:acetate CoA-transferase (BTCOAACCOAT) that converts acetate to 338
butyrate. Beyond the expected behavior of the community regarding the use of carbon 339
and energy sources, and the production of SCFAs, the best solution also predicts the 340
exchange of numerous amino acids between the two community members (Fig 3). After 341
adding the reactions suggested by the community gap-filling algorithm, FBA for the two 342
models shows that the growth is restored, with growth rates of 0.6 h^{-1} and 0.4 h^{-1} under 343
the simulated conditions, for the models of *B. adolescentis* and *F. prausnitzii* 344
respectively. 345

Table 2. Reactions added from the database to the community model according to the best solution calculated by the community gap-filling algorithm. The reactions (1) - (5) are added to the metabolic model of *B. adolescentis* ATCC 15703, while the reactions (1) and (6) - (12) are added to *F. prausnitzii* A2-165.

	Reaction Formula	Reaction (BiGG ID)
(1)	2 H$^+$[c] + SO$_4$$^{2-}$[c] → H$_2$S[c] + 2 O$_2$[c]	H2SO
(2)	L-Homocysteine[c] + 5-Methyltetrahydrofolate[c] → L-Methionine[c] + THF[c]	METS_1
(3)	CMP[c] + PPi[c] + H$^+$[c] → CTP[c] + H$_2$O[c]	NTPP4
(4)	3-Carboxy-4-methyl-2-oxopentanoate[c] + H$^+$[c] → 4-Methyl-2-oxopentanoate[c] + CO$_2$[c]	OMCDC
(5)	L-Glutamate[c] + PLP[c] + Pi[c] + H$^+$[c] + 3 H$_2$O[c] → L-Glutamine[c] + D-Ribulose-5-phosphate[c] + Dihydroxyacetone phosphate[c]	PDBL_3
(6)	SAM[c] + 2-Octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol[c] → SAH[c] + UQ8[c] + H$^+$[c]	ABAH
(7)	trans-Octaprenyl-diphosphate[c] + 4-Hydroxybenzoate[c] → 3-Octaprenyl-4-hydroxybenzoate[c] + PPi[c]	HBZOPT
(8)	SAM[c] + 2-Octaprenyl-6-hydroxyphenol[c] → SAH[c] + 2-Octaprenyl-6-methoxyphenol[c] + H$^+$[c]	OHPHM
(9)	SAM[c] + 2-Octaprenyl-6-methoxy-1,4-benzoquinol[c] → SAH[c] + 2-Octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol[c] + H$^+$[c]	OMBZLM
(10)	3-Octaprenyl-4-hydroxybenzoate[c] + H$^+$[c] → 2-Octaprenylphenol[c] + CO$_2$[c]	OPHBDC
(11)	2 SAM[c] + Fe^{2+}[c] + NAD[c] + Uroporphyrinogen-3[c] → 2 SAH[c] + Siroheme[c] + NADH[c] + 5 H$^+$[c]	SHS1
(12)	Hydroxymethylbilane[c] → Uroporphyrinogen-3[c] + H$_2$O[c]	UPP3S

Fig 3. Graphical representation of the community of *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165 after the application of the community gap-filling method. The best solution calculated by the community gap-filling algorithm predicted that the metabolic models of the strains *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165 share the available Glucose from the common medium and produce lactate (LAC), formate (FOR), and butyrate (BUT), while they exchange acetate (AC), and amino acids (3 letter code). The non-dashed arrows represent intracellular metabolite flow and the dashed arrows represent the exchange reactions of SCFAs and amino acids. The thickness of the dashed arrows represents the relative order of magnitude of the calculated fluxes for the exchange reactions.

We also evaluated the ten best alternative solutions calculated by the community 346
gap-filling algorithm and found several interesting interactions. First, in almost all the 347
solutions (S12 Table), the models of *B. adolescentis* and *F. prausnitzii* shared the 348
glucose provided in the community medium. Also, four out of ten solutions predicted 349
that *F. prausnitzii* consumes the acetate produced by *B. adolescentis*, while the rest of 350

the solutions predicted that *F. prausnitzii* consumes glucose from the media and produces acetate (S2 Fig). This behavior is consistent with previous studies that show the ability of *F. prausnitzii* A2-165 to function both as a consumer and a producer of acetate depending on the conditions [50]. However, to our current knowledge, acetate production from *F. prausnitzii* in the presence of *B. adolescentis* has not been observed experimentally. For this reason, more detailed research into the conditions that could affect the acetate production in cocultures of *F. prausnitzii* with *B. adolescentis* could yield interesting results.

We can see that in most of the solutions in which *F. prausnitzii* consumes acetate, it also produces CO_2 (S12 Table, S2 Fig, Solutions 1, 3, 5), while it does not produce CO_2 when it produces acetate (S12 Table, S2 Fig, Solutions 4, 6, 8, 9, 10). This observation is in agreement with existing information about the acetate consumption and production from *F. prausnitzii* A2-165 [50]. It is noted that in most of the solutions, CO_2 production from *F. prausnitzii* is accompanied by CO_2 consumption from *B. adolescentis*, while in the rest of the cases the reverse CO_2 crossfeeding is observed. The exchange of CO_2 between the two models could be an indication that *F. prausnitzii* and *B. adolescentis* can use CO_2 fixation to cover their carbon needs especially in environments with limited amounts of carbohydrates. This possibility is supported by an experimental study that shows that the presence of CO_2 in the medium promotes the anoxic growth of *B. adolescentis* species [75]. However, it is unclear if CO_2 fixation that leads to acetogenesis can be performed through the Wood–Ljungdahl pathway in *B. adolescentis* and in *F. prausnitzii*, as it happens in other species of the gut microbiome [76, 77].

Furthermore, the production of SCFAs from the models complies with what is expected from literature [57] being that for most of the alternative solutions *B. adolescentis* produces acetate, lactate, ethanol and formate, while *F. prausnitzii* produces butyrate, fromate and succinate (S12 Table, S2 Fig). Finally, we observed patterns in the amino acid exchanges with alanine, leucine, serine and tryptophan transferred from *B. adolescentis* to *F. prausnitzii*, and aspartate and cysteine transferred from *F. prausnitzii* to *B. adolescentis* in the majority of the alternative solutions (S12 Table, S2 Fig).

Apart from the patterns that emerged regarding the exchanged metabolites in the community, the alternative solutions to the problem also indicate a consistency in the reactions that are added from the database to the model of *F. prausnitzii* A2-165 (S13 Table). However, this is not the case for *B. adolescentis*, since the reactions added from the database to the model of *B. adolescentis* ATCC 15703 are not the same through different solutions (S13 Table). Even though most of the added reactions carry realistic fluxes, some of the reactions added to *B. adolescentis* carry unrealistically high fluxes (in the order of 10^2 mmol·gDW^{-1}·h^{-1}) and participate in thermodynamic infeasible cycles (S13 Table, Solutions 1 and 4, reaction PDBL_3).

ACT-3 community

In this case study, we simulated the cogrowth of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50 that represent the most abundant bacterial species of the dehalogenating ACT-3 community [63]. Our simulation is set up in order to represent the conditions of a previous experimental study on the interaction between *Dehalobacter* and *Bacteroidales* [63]. More specifically, our community is allowed to grow anaerobically on chloroform, that can be used by the model of *Dehalobacter* sp. CF as an energy source, and lactate, that can be used by the model of *Bacteroidales* sp. CF50 as a carbon and energy source. The two models are allowed to exchange organic acids and amino acids. We show that the growth in the community of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50 can be restored by the community gap-filling algorithm.

The first solution calculated by the community gap-filling algorithm suggested the [402] addition of one reaction to the model of *Dehalobacter* sp. CF and two reactions to the [403] model of *Bacteroidales* sp. CF50 (Table 3 and S14 Table). More specifically, the [404] reaction of calcium transport via the ABC system (CA2abc) was added in the model of [405] *Dehalobacter* sp. CF. The model uses the export of Ca^{2+} through the ABC system as a [406] proton exchange pump to generate ATP. The additions to the model of *Bacteroidales* sp. [407] CF50 include the reactions of ammonia transport (NH4ti) and anthranilate synthase [408] (ANS2) that are used in the model in order to import ammonia and use it to produce [409] pyruvate from chorismate. Moreover, the first solution also predicts the ability of [410] *Bacteroidales* sp. CF50 to uptake lactate from the medium and produce hydrogen, CO_2, [411] acetate, and malate that are used by *Dehalobacter* sp. CF. In parallel, *Dehalobacter* sp. [412] CF consumes the hydrogen produced by *Bacteroidales* sp. CF50 and chloroform from [413] the medium, and produces chlorine and dichloromethane as it performs dechlorination. [414] In addition, *Dehalobacter* sp. CF provides pyruvate to *Bacteroidales* sp. CF50. Overall, [415] we see that the gap-filling algorithm was able to predict the experimentally observed [416] behavior of the community [63], and it also predicted the exchange of amino acids [417] between the two community members, which remained elusive in the experimental study. [418] The exchanges between the two models are depicted in Fig 4. A post gap-filling analysis [419] shows that all biomass precursors are replenished in both models, and FBA calculated a [420] growth rate of 0.1 d^{-1} for *Dehalobacter* sp. CF, and 0.01 d^{-1} for *Bacteroidales* sp. CF50. [421]

Table 3. Reactions added from the database to the community model according to the best solution calculated by the community gap-filling algorithm. Reaction (1) is added to the metabolic model of *Dehalobacter* sp. CF, while reactions (2) and (3) are added to *Bacteroidales* sp. CF50.

	Reaction Formula	Reaction (BiGG ID)
(1)	$Ca_{2+}[c] + Pi[c] + H^+[c] + ADP[c] \rightarrow Ca_{2+}[e] + H_2O[c] + ATP[c]$	CA2abc
(2)	$NH^{4+}[e] \rightarrow NH^{4+}[c]$	NH4ti
(3)	$NH^{4+}[c] + Chorismate[c] \rightarrow H^+[c] + H_2O[c] + Anthranilate[c] + Pyruvate[c]$	ANS2

Fig 4. Graphical representation of the community of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50 after the application of the community gap-filling method. The best solution calculated by the community gap-filling algorithm predicted that the metabolic model of *Dehalobacter* sp. CF consumes chloroform (CF), while the metabolic model of *Bacteroidales* sp. CF50 consumes lactate (LAC) from the common medium. *Dehalobacter* sp. CF consumes H_2 and CO_2 produced by *Bacteroidales* sp. CF50, and uses part of the consumed H_2 in order to respire chloroform (CF) to chlorine (Cl) and dichloromethane (DCM). The two models exchange acetate (AC), malate (MAL) and pyruvate (PYR) as expected from experimental studies. The algorithm also predicted the exchange of specific amino acids (3 letter code) between the models. The non-dashed arrows represent intracellular metabolite flow, while the dashed arrows represent exchange reactions. The thickness of the dashed arrows represents the relative order of magnitude of the calculated fluxes for the exchange reactions.

An analysis of the ten best solutions calculated by the community gap-filling [422] algorithm shows that the model of *Bacteroidales* sp. CF50 consistently uptakes lactate, [423] the only carbon and energy source in the medium, and ferments it into CO_2, hydrogen, [424] acetate, malate, fumarate and succinate, while the model of *Dehalobacter* sp. CF [425] consistently performs dechlorination by reducing chloroform and producing [426] dichloromethane and chlorine (S15 Table). Moreover, all the alternative solutions [427] predicted the ability of *Dehalobacter* sp. CF to consume part of the acetate and all the [428]

malate and hydrogen produced by *Bacteroidales* sp. CF50, while providing   429
Bacteroidales sp. CF50 with pyruvate (S15 Table, S3 Fig).   430

In most of the solutions *Dehalobacter* sp. CF consumes part of the CO_2 produced by   431
Bacteroidales sp. CF50 (S15 Table, S3 Fig, Solutions 1 - 5, 8, 10). This exchange could   432
indicate that *Dehalobacter* sp. CF performs CO_2 fixation, which is possible since the   433
species is known to possess the Wood–Ljungdahl pathway [63]. Furthermore, we   434
observed patterns in the amino acid exchanges between the two community members,   435
with arginine, asparagine, isoleucine, lysine, phenylalanine, threonine, and tyrosine   436
transferred from *Bacteroidales* sp. CF50 to *Dehalobacter* sp. CF, and alanine and   437
glutamate transferred from *Dehalobacter* sp. CF to *Bacteroidales* sp. CF50 in the   438
majority of the alternative solutions (S15 Table, S3 Fig).   439

More information about the quality of the solutions calculated by the community   440
gap-filling algorithm can be retrieved by closer inspection of the reactions that were   441
added from the database to the models in the ten best solutions. We can see that in   442
most of the solutions (S16 Table, Solutions 1 - 6, 8, 10), even though the reactions   443
added to the model of *Bacteroidales* sp. CF50 have realistic fluxes, those added to the   444
model of *Dehalobacter* sp. CF carry unrealistically high fluxes (close to the infinite   445
reaction bounds). For example, the already discussed reaction of calcium transport via   446
the ABC system (CA2abc), which is used for ATP production in the model, has   447
unrealistically high fluxes and forms a cycle with the reaction calcium transport in/out   448
via proton antiporter (CAt4) that already exists in the *Dehalobacter* sp. CF model (S16   449
Table, Solutions 1, 2, 3).   450

Discussion   451

Our results demonstrate the ability of the community gap-filling method to resolve   452
metabolic gaps and restore growth in metabolic models by adding the minimum possible   453
number of biochemical reactions from a reference database to the models and activating   454
already existing reactions of the models, while the models are allowed to interact   455
metabolically. In the first test case study, the algorithm restored community growth and   456
predicted the exchange of acetate between the two *E. coli* strains. For the community of   457
B. adolescentis and *F. prausnitzii* in the second case study, the algorithm calculated the 458
competitive consumption of glucose present in the common medium, the syntrophic   459
exchange of acetate, CO_2 and amino acids, and the production of SCFAs. The   460
algorithm also identified the exchange of organic acids, amino acids, CO_2 and H_2 for the 461
community of *Dehalobacter* and *Bacteroidales*. In all three case studies, the metabolic   462
exchanges were not strictly forced with constraints, but they emerged from the   463
algorithm. We also showed that the implementation of FVA in our algorithm resulted in   464
significantly decreased solution time of the MILP community gap-filling problem.   465

In general, the community gap-filling method could facilitate the study of complex   466
communities where the existing information for the involved species is incomplete. A   467
method like this can be used at the last steps of the metabolic reconstruction process,   468
and it can also become part of the efforts for the creation of a microbiome modeling   469
toolbox [78]. However, like every other constraint-based method, our community   470
gap-filling method has some weaknesses. One of the most prominent challenges is that   471
the performance of the algorithm depends highly on the quality of the input models and   472
databases being used. More specifically, the absence of some important reactions from   473
the models can lead to poor predictive power, while unconstrained reaction   474
directionality, especially for exchange reactions, can affect the predicted metabolic   475
interactions and give birth to flux-carrying cycles between the community members.   476
The algorithm performs better with highly curated models, but even then it shows   477
preference over specific pathways. This bias comes from the fact that the process of   478

manual curation and gap-filling for a model is based on the use of specific media and requires knowledge about the growth of the organism under specific conditions. In addition, the composition and the quality of the database reactions is important. Ideally the database should contain balanced biochemical reactions that represent a diversity of metabolic functions observed in prokaryotic organisms, or any type of organisms that represents the models being used. Regarding the solutions of the algorithm that predict reactions with unrealistically big fluxes and reactions that form thermodynamically infeasible cycles, they can either be excluded completely from the analysis of results as they are not trustworthy, or they could be avoided by further constraining the reactions of the metabolic models based on thermodynamic information [79,80]. From the perspective of computational efficiency, the total running time of the algorithm can be further decreased with the use of newer and faster parallel implementations of FVA, such as VFFVA [81].

Conclusion

In this work, we presented an algorithm that performs gap-filling while permitting metabolic interactions in microbial communities. Metabolites produced from one organism can either be released to the environment or used by another member of the community. Therefore, one of the key characteristics of the community gap-filling method is that it attempts to identify if interspecies exchanges of metabolites can compensate for insufficient metabolic capabilities of the community members, instead of just adding biochemical reactions from reference databases to the metabolic models of the community members. Our results showed that the algorithm can successfully predict both cooperative and competitive metabolic interactions in microbial communities and comply with experimental measurements and observations. With the ability to predict elusive metabolic interactions, the community gap-filling method can be used to generate hypotheses about potential metabolic relationships among community members. This direction is useful not only for the improvement of metabolic models, but also for understanding functions at the community level and providing valuable information for selective media composition that can enable the isolation of pure cultures for microorganisms.

The community gap-filling method can be further applied on communities with more than two members, even though such a scale-up could probably require a more computationally efficient reformulation. A way to use the community gap-filling method for the study of microbial communities could be the following. Applying the community gap-filling method for the simulation of a microbial community on different media can generate plenty of data on the function of the community in different environments. Evaluating these results to identify the reactions added to the models, as well as the metabolic interactions between the models, can offer an insight on which reactions are necessary for the metabolic reconstructions and which metabolic interdependencies are prevalent in the community. Such predictions can be further tested experimentally and can be used for improving gene-annotations by identifying gene functions, not based on sequence similarities, but rather on a growth restoration strategy.

Supporting information

S1 Appendix.   Creation of glucose and acetate utilizer strains from the *E. coli* core model.

S2 Appendix.   Reconstruction and individual curation of the metabolic

models of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50.   525

S1 File. **core_E.coli_community_solutions.mat.** Alternative solutions for the   526
core *E. coli* community.   527

S2 File. **gut-microbiome_community_solutions.mat.** Alternative solutions for   528
the community of *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165.   529

S3 File. **ACT-3_community_solutions.mat.** Alternative solutions for the   530
community of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50.   531

S1 Table. **Deleted and Exchange Reactions for *E. coli* glucose utilizer.**   532

S2 Table. **Deleted and Exchange Reactions for *E. coli* acetate utilizer.**   533

S3 Table. **Exchange Reactions for *B. adolescentis* ATCC 15703.**   534

S4 Table. **Exchange Reactions for *F. prausnitzii* A2-165.**   535

S5 Table. **Uptake Rates for ACT-3 community.**   536

S6 Table. **Constrained and Exchange Reactions for *Dehalobacter* sp. CF.**   537

S7 Table. **Exchange Reactions for *Bacteroidales* sp. CF50.**   538

S8 Table. **Reaction Fluxes from the best solution for the core *E. coli*   539
community and Reactions Added to the models.**   540

S9 Table. **Fluxes of Exchange Reactions from the ten best solutions for   541
the core *E. coli* community.**   542

S10 Table. **Fluxes of Added Database Reactions from the ten best   543
solutions for the core *E. coli* community.**   544

S11 Table. **Reaction Fluxes from the best solution for the community of   545
B. adolescentis ATCC 15703 and *F. prausnitzii* A2-165 and Reactions   546
Added to the models.**   547

S12 Table. **Fluxes of Exchange Reactions from the ten best solutions for   548
the community of *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165.**   549

S13 Table. **Fluxes of Added Database Reactions from the ten best   550
solutions for the community of *B. adolescentis* ATCC 15703 and *F.   551
prausnitzii* A2-165.**   552

S14 Table. **Reaction Fluxes from the best solution for the community of   553
Dehalobacter sp. CF and *Bacteroidales* sp. CF50 and Reactions Added to   554
the models.**   555

S15 Table. **Fluxes of Exchange Reactions from the ten best solutions for** ₅₅₆
the community of *Dehalobacter* **sp. CF and** *Bacteroidales* **sp. CF50.** ₅₅₇

S16 Table. **Fluxes of Added Database Reactions from the ten best** ₅₅₈
solutions for the community of *Dehalobacter* **sp. CF and** *Bacteroidales* **sp.** ₅₅₉
CF50. ₅₆₀

S1 Fig. **Time needed for the solution of the MILP problem with and** ₅₆₁
without the prior use of FVA for the reduction of the solution space, ₅₆₂
depending on the number of organisms in the community. As the number of ₅₆₃
organisms that make up the microbial community increases, the community gap-filling ₅₆₄
algorithm needs increasingly more time to solve the MILP problem that restores growth ₅₆₅
in the community. However, the increase in solving time is significantly reduced with ₅₆₆
the use of our community gap-filling method that reduces the solution space of the ₅₆₇
MILP problem by performing FVA for each organism compartment of the community ₅₆₈
before formulating the optimization problem. (The presented measurements were ₅₆₉
performed with the models from the toy *E. coli* community.) ₅₇₀

S2 Fig. **Fluxes of Exchange Reactions from the ten best solutions for the** ₅₇₁
community of *B. adolescentis* **ATCC 15703 and** *F. prausnitzii* **A2-165. a)** ₅₇₂
Fluxes of exchange reactions for the SCFAs acetate, butyrate, ethanol, lactate, formate, ₅₇₃
and succinate, and CO_2. **b)** Fluxes of exchange reactions for the amino acids alanine, ₅₇₄
aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, leucine, ₅₇₅
phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. ₅₇₆

S3 Fig. **Fluxes of Exchange Reactions from the ten best solutions for the** ₅₇₇
community of *Dehalobacter* **sp. CF and** *Bacteroidales* **sp. CF50. a)** Fluxes ₅₇₈
of exchange reactions for the organic acids acetate, malate, and pyruvate, CO_2 and H_2. ₅₇₉
b) Fluxes of exchange reactions for the amino acids alanine, arginine, asparagine, ₅₈₀
glutamate, isoleucine, lysine, phenylalanine, threonine, and tyrosine. ₅₈₁

Availability ₅₈₂

The code used to perform this study can be found on GitHub: ₅₈₃
`https://github.com/LMSE/Community_Gap-Filling`. ₅₈₄

Acknowledgments ₅₈₅

References

1. Jiang LL, Zhou JJ, Quan CS, Xiu ZL. Advances in Industrial Microbiome Based on Microbial Consortium for Biorefinery. Bioresources and Bioprocessing. 2017;4(1). doi:10.1186/s40643-017-0141-0.

2. Falkowski PG, Fenchel T, Delong EF. The Microbial Engines That Drive Earth's Biogeochemical Cycles. Science. 2008;320(5879):1034–1039. doi:10.1126/science.1153213.

3. Wang B, Yao M, Lv L, Ling Z, Li L. The Human Microbiota in Health and Disease. Engineering. 2017;3(1):71–82. doi:10.1016/J.ENG.2017.01.008.

4. Konopka A. What Is Microbial Community Ecology? The ISME Journal. 2009;3(11):1223–1230. doi:10.1038/ismej.2009.88.

5. Ang KS, Lakshmanan M, Lee NR, Lee DY. Metabolic Modeling of Microbial Community Interactions for Health, Environmental and Biotechnological Applications. Current Genomics. 2018;19(8):712–722. doi:10.2174/1389202919666180911144055.

6. Mahadevan R, Henson MA. Genome-Based Modeling and Design of Metabolic Interactions in Microbial Communities. Computational and Structural Biotechnology Journal. 2012;3(4):e201210008. doi:10.5936/csbj.201210008.

7. O'Brien EJ, Monk JM, Palsson BO. Using Genome-Scale Models to Predict Biological Capabilities. Cell. 2015;161(5):971–987. doi:10.1016/j.cell.2015.05.019.

8. Chan SHJ, Simons MN, Maranas CD. SteadyCom: Predicting Microbial Abundances While Ensuring Community Stability. PLOS Computational Biology. 2017;13(5):e1005539. doi:10.1371/journal.pcbi.1005539.

9. Zomorrodi AR, Maranas CD. OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. PLOS Computational Biology. 2012;8(2):e1002363. doi:10.1371/journal.pcbi.1002363.

10. Zomorrodi AR, Islam MM, Maranas CD. D-OptCom: Dynamic Multi-Level and Multi-Objective Metabolic Modeling of Microbial Communities. ACS Synthetic Biology. 2014;3(4):247–257. doi:10.1021/sb4001307.

11. Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, et al. Genome-Scale Dynamic Modeling of the Competition between Rhodoferax and Geobacter in Anoxic Subsurface Environments. The ISME Journal. 2011;5(2):305–316. doi:10.1038/ismej.2010.117.

12. Dukovski I, Bajić D, Chacón JM, Quintin M, Vila JC, Sulheim S, et al. Computation Of Microbial Ecosystems in Time and Space (COMETS): An Open Source Collaborative Platform for Modeling Ecosystems Metabolism. arXiv:200901734 [q-bio]. 2020;.

13. Mendoza SN, Olivier BG, Molenaar D, Teusink B. A Systematic Assessment of Current Genome-Scale Metabolic Reconstruction Tools. Genome Biology. 2019;20(1):158. doi:10.1186/s13059-019-1769-1.

14. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. High-Throughput Generation, Optimization and Analysis of Genome-Scale Metabolic Models. Nature Biotechnology. 2010;28(9):977–982. doi:10.1038/nbt.1672.

15. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, et al. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nature Biotechnology. 2018;36(7):566–569. doi:10.1038/nbt.4163.

16. Gerlt JA. How to Find "Missing" Genes. Chemistry & Biology. 2003;10(12):1141–1142. doi:10.1016/j.chembiol.2003.12.008.

17. Thiele I, Palsson BØ. A Protocol for Generating a High-Quality Genome-Scale Metabolic Reconstruction. Nature Protocols. 2010;5(1):93–121. doi:10.1038/nprot.2009.203.

18. Satish Kumar V, Dasika MS, Maranas CD. Optimization Based Automated Curation of Metabolic Reconstructions. BMC Bioinformatics. 2007;8(1):212. doi:10.1186/1471-2105-8-212.

19. Thiele I, Vlassis N, Fleming RMT. fastGapFill: Efficient Gap Filling in Metabolic Networks. Bioinformatics. 2014;30(17):2529–2531. doi:10.1093/bioinformatics/btu321.

20. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, et al. Creation and Analysis of Biochemical Constraint-Based Models Using the COBRA Toolbox v.3.0. Nature Protocols. 2019;14(3):639–702. doi:10.1038/s41596-018-0098-2.

21. Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, et al. The MetaCyc Database of Metabolic Pathways and Enzymes - a 2019 Update. Nucleic Acids Research. 2020;48(D1):D445–D453. doi:10.1093/nar/gkz862.

22. Zimmermann J, Kaleta C, Waschina S. Gapseq: Informed Prediction of Bacterial Metabolic Pathways and Reconstruction of Accurate Metabolic Models. bioRxiv. 2020; p. 2020.03.20.000737. doi:10.1101/2020.03.20.000737.

23. Medlock GL, Papin JA. Guiding the Refinement of Biochemical Knowledgebases with Ensembles of Metabolic Networks and Machine Learning. Cell Systems. 2020;10(1):109–119.e3. doi:10.1016/j.cels.2019.11.006.

24. Vitkin E, Shlomi T. MIRAGE: A Functional Genomics-Based Approach for Metabolic Network Model Reconstruction and Its Application to Cyanobacteria Networks. Genome Biology. 2012;13(11):R111. doi:10.1186/gb-2012-13-11-r111.

25. Benedict MN, Mundy MB, Henry CS, Chia N, Price ND. Likelihood-Based Gene Annotations for Gap Filling and Quality Assessment in Genome-Scale Metabolic Models. PLOS Computational Biology. 2014;10(10):e1003882. doi:10.1371/journal.pcbi.1003882.

26. Ong WK, Midford PE, Karp PD. Taxonomic Weighting Improves the Accuracy of a Gap-Filling Algorithm for Metabolic Models. Bioinformatics. 2020;36(6):1823–1830. doi:10.1093/bioinformatics/btz813.

27. Machado D, Andrejev S, Tramontano M, Patil KR. Fast Automated Reconstruction of Genome-Scale Metabolic Models for Microbial Species and Communities. Nucleic Acids Research. 2018;46(15):7542–7553. doi:10.1093/nar/gky537.

28. Herrgård MJ, Fong SS, Palsson BØ. Identification of Genome-Scale Metabolic Network Models Using Experimentally Measured Flux Profiles. PLOS Computational Biology. 2006;2(7):e72. doi:10.1371/journal.pcbi.0020072.

29. Kumar VS, Maranas CD. GrowMatch: An Automated Method for Reconciling In Silico/In Vivo Growth Predictions. PLOS Computational Biology. 2009;5(3):e1000308. doi:10.1371/journal.pcbi.1000308.

30. Schroeder WL, Saha R. OptFill: A Tool for Infeasible Cycle-Free Gapfilling of Stoichiometric Metabolic Models. iScience. 2020;23(1):100783. doi:10.1016/j.isci.2019.100783.

31. Seaver SMD, Liu F, Zhang Q, Jeffryes J, Faria JP, Edirisinghe JN, et al. The ModelSEED Database for the Integration of Metabolic Annotations and the Reconstruction, Comparison, and Analysis of Metabolic Models for Plants, Fungi, and Microbes. bioRxiv. 2020; p. 2020.03.31.018663. doi:10.1101/2020.03.31.018663.

32. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research. 2000;28(1):27–30.

33. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, et al. BiGG Models: A Platform for Integrating, Standardizing and Sharing Genome-Scale Models. Nucleic Acids Research. 2016;44(D1):D515–D522. doi:10.1093/nar/gkv1049.

34. Helling RB, Vargas CN, Adams J. Evolution of Escherichia Coli during Growth in a Constant Environment. Genetics. 1987;116(3):349–358.

35. Treves DS, Manning S, Adams J. Repeated Evolution of an Acetate-Crossfeeding Polymorphism in Long-Term Populations of Escherichia Coli. Molecular Biology and Evolution. 1998;15(7):789–797. doi:10.1093/oxfordjournals.molbev.a025984.

36. Cresci GA, Bawden E. Gut Microbiome. Nutrition in Clinical Practice. 2015;30(6):734–746. doi:10.1177/0884533615609899.

37. Rajilić-Stojanović M, de Vos WM. The First 1000 Cultured Species of the Human Gastrointestinal Microbiota. FEMS microbiology reviews. 2014;38(5):996–1047. doi:10.1111/1574-6976.12075.

38. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A Human Gut Microbial Gene Catalogue Established by Metagenomic Sequencing. Nature. 2010;464(7285):59–65. doi:10.1038/nature08821.

39. Macfarlane GT, Macfarlane S. Fermentation in the Human Large Intestine: Its Physiologic Consequences and the Potential Contribution of Prebiotics. Journal of Clinical Gastroenterology. 2011;45 Suppl:S120–127. doi:10.1097/MCG.0b013e31822fecfe.

40. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, et al. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. Frontiers in Immunology. 2019;10. doi:10.3389/fimmu.2019.00277.

41. Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, Brummer RJ. Review Article: The Role of Butyrate on Colonic Function. Alimentary Pharmacology & Therapeutics. 2008;27(2):104–119. doi:10.1111/j.1365-2036.2007.03562.x.

42. Martín R, Bermúdez-Humarán LG, Langella P. Searching for the Bacterial Effector: The Example of the Multi-Skilled Commensal Bacterium Faecalibacterium Prausnitzii. Frontiers in Microbiology. 2018;9. doi:10.3389/fmicb.2018.00346.

43. Miquel S, Martín R, Bridonneau C, Robert V, Sokol H, Bermúdez-Humarán LG, et al. Ecology and Metabolism of the Beneficial Intestinal Commensal Bacterium Faecalibacterium Prausnitzii. Gut Microbes. 2014;5(2):146–151. doi:10.4161/gmic.27651.

44. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux JJ, et al. Faecalibacterium Prausnitzii Is an Anti-Inflammatory Commensal Bacterium Identified by Gut Microbiota Analysis of Crohn Disease Patients. Proceedings of the National Academy of Sciences. 2008;105(43):16731–16736. doi:10.1073/pnas.0804812105.

45. Sokol H Dr, Seksik P, Rigottier-Gois L, Lay C, Lepage P, Podglajen I, et al. Specificities of the Fecal Microbiota in Inflammatory Bowel Disease. Inflammatory Bowel Diseases. 2006;12(2):106–111. doi:10.1097/01.MIB.0000200323.38139.c6.

46. Kang S, Denman SE, Morrison M, Yu Z, Dore J, Leclerc M, et al. Dysbiosis of Fecal Microbiota in Crohn's Disease Patients as Revealed by a Custom Phylogenetic Microarray. Inflammatory Bowel Diseases. 2010;16(12):2034–2042. doi:10.1002/ibd.21319.

47. Chen W, Liu F, Ling Z, Tong X, Xiang C. Human Intestinal Lumen and Mucosa-Associated Microbiota in Patients with Colorectal Cancer. PLOS ONE. 2012;7(6):e39743. doi:10.1371/journal.pone.0039743.

48. Lopez-Siles M, Duncan SH, Garcia-Gil LJ, Martinez-Medina M. Faecalibacterium Prausnitzii : From Microbiology to Diagnostics and Prognostics. The ISME Journal. 2017;11(4):841–852. doi:10.1038/ismej.2016.176.

49. Lopez-Siles M, Khan TM, Duncan SH, Harmsen HJM, Garcia-Gil LJ, Flint HJ. Cultured Representatives of Two Major Phylogroups of Human Colonic Faecalibacterium Prausnitzii Can Utilize Pectin, Uronic Acids, and Host-Derived Substrates for Growth. Applied and Environmental Microbiology. 2012;78(2):420–428. doi:10.1128/AEM.06858-11.

50. Heinken A, Khan MT, Paglia G, Rodionov DA, Harmsen HJM, Thiele I. Functional Metabolic Map of Faecalibacterium Prausnitzii, a Beneficial Human Gut Microbe. Journal of Bacteriology. 2014;196(18):3289–3302. doi:10.1128/JB.01780-14.

51. Rivière A, Selak M, Lantin D, Leroy F, De Vuyst L. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. Frontiers in Microbiology. 2016;7. doi:10.3389/fmicb.2016.00979.

52. De Vuyst L, Leroy F. Cross-Feeding between Bifidobacteria and Butyrate-Producing Colon Bacteria Explains Bifdobacterial Competitiveness, Butyrate Production, and Gas Production. International Journal of Food Microbiology. 2011;149(1):73–80. doi:10.1016/j.ijfoodmicro.2011.03.003.

53. Rossi M, Corradini C, Amaretti A, Nicolini M, Pompei A, Zanoni S, et al. Fermentation of Fructooligosaccharides and Inulin by Bifidobacteria: A Comparative Study of Pure and Fecal Cultures. Applied and Environmental Microbiology. 2005;71(10):6150–6158. doi:10.1128/AEM.71.10.6150-6158.2005.

54. Pokusaeva K, Fitzgerald GF, van Sinderen D. Carbohydrate Metabolism in Bifidobacteria. Genes & Nutrition. 2011;6(3):285–306. doi:10.1007/s12263-010-0206-6.

55. O'Callaghan A, van Sinderen D. Bifidobacteria and Their Role as Members of the Human Gut Microbiota. Frontiers in Microbiology. 2016;7. doi:10.3389/fmicb.2016.00925.

56. Devika NT, Raman K. Deciphering the Metabolic Capabilities of Bifidobacteria Using Genome-Scale Metabolic Models. Scientific Reports. 2019;9(1):18222. doi:10.1038/s41598-019-54696-9.

57. Rios-Covian D, Gueimonde M, Duncan SH, Flint HJ, de los Reyes-Gavilan CG. Enhanced Butyrate Formation by Cross-Feeding between Faecalibacterium Prausnitzii and Bifidobacterium Adolescentis. FEMS Microbiology Letters. 2015;362(21). doi:10.1093/femsle/fnv176.

58. Moens F, Weckx S, De Vuyst L. Bifidobacterial Inulin-Type Fructan Degradation Capacity Determines Cross-Feeding Interactions between Bifidobacteria and Faecalibacterium Prausnitzii. International Journal of Food Microbiology. 2016;231:76–85. doi:10.1016/j.ijfoodmicro.2016.05.015.

59. El-Semman IE, Karlsson FH, Shoaie S, Nookaew I, Soliman TH, Nielsen J. Genome-Scale Metabolic Reconstructions of Bifidobacterium Adolescentis L2-32 and Faecalibacterium Prausnitzii A2-165 and Their Interaction. BMC Systems Biology. 2014;8(1):41. doi:10.1186/1752-0509-8-41.

60. Cappelletti M, Frascari D, Zannoni D, Fedi S. Microbial Degradation of Chloroform. Applied Microbiology and Biotechnology. 2012;96(6):1395–1409. doi:10.1007/s00253-012-4494-1.

61. Jugder BE, Ertan H, Bohl S, Lee M, Marquis CP, Manefield M. Organohalide Respiring Bacteria and Reductive Dehalogenases: Key Tools in Organohalide Bioremediation. Frontiers in Microbiology. 2016;7. doi:10.3389/fmicb.2016.00249.

62. Laturnus F, Haselmann KF, Borch T, Grøn C. Terrestrial Natural Sources of Trichloromethane (Chloroform, CHCl3) – An Overview. Biogeochemistry. 2002;60(2):121–139. doi:10.1023/A:1019887505651.

63. Wang PH, Correia K, Ho HC, Venayak N, Nemr K, Flick R, et al. An Interspecies Malate–Pyruvate Shuttle Reconciles Redox Imbalance in an Anaerobic Microbial Community. The ISME Journal. 2019;13(4):1042–1055. doi:10.1038/s41396-018-0333-4.

64. Grostern A, Duhamel M, Dworatzek S, Edwards EA. Chloroform Respiration to Dichloromethane by a Dehalobacter Population. Environmental Microbiology. 2010;12(4):1053–1060. doi:10.1111/j.1462-2920.2009.02150.x.

65. Justicia-Leon SD, Higgins S, Mack EE, Griffiths DR, Tang S, Edwards EA, et al. Bioaugmentation with Distinct Dehalobacter Strains Achieves Chloroform Detoxification in Microcosms. Environmental Science & Technology. 2014;48(3):1851–1858. doi:10.1021/es403582f.

66. Grostern A, Edwards EA. A 1,1,1-Trichloroethane-Degrading Anaerobic Mixed Microbial Culture Enhances Biotransformation of Mixtures of Chlorinated Ethenes and Ethanes. Applied and Environmental Microbiology. 2006;72(12):7849–7856. doi:10.1128/AEM.01269-06.

67. Justicia-Leon SD, Ritalahti KM, Mack EE, Löffler FE. Dichloromethane Fermentation by a Dehalobacter Sp. in an Enrichment Culture Derived from Pristine River Sediment. Applied and Environmental Microbiology. 2012;78(4):1288–1291. doi:10.1128/AEM.07325-11.

68. Grostern A, Edwards EA. Characterization of a Dehalobacter Coculture That Dechlorinates 1,2-Dichloroethane to Ethene and Identification of the Putative Reductive Dehalogenase Gene. Applied and Environmental Microbiology. 2009;75(9):2684–2693. doi:10.1128/AEM.02037-08.

69. Gudmundsson S, Thiele I. Computationally Efficient Flux Variability Analysis. BMC Bioinformatics. 2010;11(1):489. doi:10.1186/1471-2105-11-489.

70. Orth JD, Fleming RMT, Palsson BØ. Reconstruction and Use of Microbial Metabolic Networks: The Core Escherichia Coli Metabolic Model as an Educational Guide. EcoSal Plus. 2010;4(1). doi:10.1128/ecosalplus.10.2.1.

71. San Roman M, Wagner A. Acetate and Glycerol Are Not Uniquely Suited for the Evolution of Cross-Feeding in E. Coli. PLoS Computational Biology. 2020;16(11). doi:10.1371/journal.pcbi.1008433.

72. Wolfsberg E, Long CP, Antoniewicz MR. Metabolism in Dense Microbial Colonies: 13C Metabolic Flux Analysis of E. Coli Grown on Agar Identifies Two Distinct Cell Populations with Acetate Cross-Feeding. Metabolic Engineering. 2018;49:242–247. doi:10.1016/j.ymben.2018.08.013.

73. Noronha A, Modamio J, Jarosz Y, Guerard E, Sompairac N, Preciat G, et al. The Virtual Metabolic Human Database: Integrating Human and Gut Microbiome Metabolism with Nutrition and Disease. Nucleic Acids Research. 2019;47(D1):D614–D624. doi:10.1093/nar/gky992.

74. Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, et al. Generation of Genome-Scale Metabolic Reconstructions for 773 Members of the Human Gut Microbiota. Nature Biotechnology. 2017;35(1):81–89. doi:10.1038/nbt.3703.

75. Kawasaki S, Nagasaku M, Mimura T, Katashima H, Ijyuin S, Satoh T, et al. Effect of CO2 on Colony Development by Bifidobacterium Species. Applied and Environmental Microbiology. 2007;73(23):7796–7798. doi:10.1128/AEM.01163-07.

76. Wolin MJ, Miller TL, Yerry S, Zhang Y, Bank S, Weaver GA. Changes of Fermentation Pathways of Fecal Microbial Communities Associated with a Drug Treatment That Increases Dietary Starch in the Human Colon. Applied and Environmental Microbiology. 1999;65(7):2807–2812.

77. Ríos-Covián D, Ruas-Madiedo P, Margolles A, Gueimonde M, de los Reyes-Gavilán CG, Salazar N. Intestinal Short Chain Fatty Acids and Their Link with Diet and Human Health. Frontiers in Microbiology. 2016;7. doi:10.3389/fmicb.2016.00185.

78. Baldini F, Heinken A, Heirendt L, Magnusdottir S, Fleming RMT, Thiele I. The Microbiome Modeling Toolbox: From Microbial Interactions to Personalized Microbial Communities. Bioinformatics. 2019;35(13):2332–2334. doi:10.1093/bioinformatics/bty941.

79. Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-Based Metabolic Flux Analysis. Biophysical Journal. 2007;92(5):1792–1805. doi:10.1529/biophysj.106.093138.

80. Salvy P, Fengos G, Ataman M, Pathier T, Soh KC, Hatzimanikatis V. pyTFA and matTFA: A Python Package and a Matlab Toolbox for Thermodynamics-Based Flux Analysis. Bioinformatics. 2019;35(1):167–169. doi:10.1093/bioinformatics/bty499.

81. Guebila MB. VFFVA: Dynamic Load Balancing Enables Large-Scale Flux Variability Analysis. BMC Bioinformatics. 2020;21(1):424. doi:10.1186/s12859-020-03711-2.

May 4, 2021

Figures

Fig 1. Graphical representation of the community gap-filling method. The metabolic reconstructions of two individual organisms (blue and orange respectively) are allowed to exchange metabolites and interact with the environment through a common metabolite pool (green). **a)** The algorithm adds biochemical reactions from a reference database (dark purple) to the community model. **b)** The community gap-filling algorithm is an MILP problem with the objective to add the minimum number of database reactions to the community model in order to restore biomass production in the individual metabolic models, while it satisfies some constrains for the reaction fluxes. The fate of each database reaction is controlled by a binary variable, which takes the value of 1 if the database reaction is added to the community model, and 0 otherwise.

Fig 2. Graphical representation of the core *E. coli* community after the application of the community gap-filling method. The central carbon metabolism of an *E. coli* core model (blue rectangular: Glycolysis, pink square: Pentose Phosphate Pathway,

yellow circle: TCA cycle) was used in order to create two *E. coli* strains: one that consumes glucose (left) and one that consumes acetate (right), after the deletion of the reactions marked with red crosses. The best solution of the community gap-filling algorithm predicted the addition of the reactions represented by continuous blue arrows and the activation of the existing reactions represented by dashed blue arrows, in order to restore biomass production in the two models. The dashed black arrows show the exchange reactions for glucose and acetate in the community. The metabolites in bold represent biomass precursors. The numbered deleted reactions are: (1) PGM, (2) MALS, (3) SUCOAS, (4) PFL, (5) PTAr, (6) GLCpts, (7) CS. The numbered added and activated reactions are: (1) POX, (2) CS, (3) PKETF, (4) EDD and EDA, (5) PDH, (6) ACS, (7) ACKr, (8) ME1.

Fig 3. Graphical representation of the community of *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165 after the application of the community gap-filling method. The best solution calculated by the community gap-filling algorithm predicted that the metabolism modelled by the strains *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165 share the available Glucose from the common medium and produce lactate (LAC), formate (FOR), and butyrate (BUT), while they exchange acetate (AC), and amino acids (3 letter code). The non-dashed arrows represent intracellular metabolite flow and the dashed arrows represent the exchange reactions of SCFAs and amino acids. The thickness of the dashed arrows represents the relative order of magnitude of the calculated fluxes for the exchange reactions.

Fig 4. Graphical representation of the community of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50 after the application of the community gap-filling method. The best solution calculated by the community gap-filling algorithm predicted that the metabolic model of *Dehalobacter* sp. CF consumes chloroform (CF), while the metabolic model of *Bacteroidales* sp. CF50 consumes lactate (LAC) from the common medium. *Dehalobacter* sp. CF consumes H_2 and CO_2 produced by *Bacteroidales* sp. CF50, and uses part of the consumed H_2 in order to respire chloroform (CF) to chlorine (Cl) and dichloromethane (DCM). The two models exchange acetate (AC), malate (MAL) and pyruvate (PYR) as expected from experimental studies. The algorithm also predicted the exchange of specific amino acids (3 letter code) between the models. The non-dashed arrows represent intracellular metabolite flow, while the dashed arrows represent exchange reactions. The thickness of the dashed arrows represents the relative order of magnitude of the calculated fluxes for the exchange reactions.

S1 Fig. Time needed for the solution of the MILP problem with and without the prior use of FVA for the reduction of the solution space, depending on the number of organisms in the community. As the number of organisms that make up the

microbial community increases, the community gap-filling algorithm needs increasingly more time to solve the MILP problem that restores growth in the community. However, the increase in solving time is significantly reduced with the use of our community gap-filling method that reduces the solution space of the MILP problem by performing FVA for each organism compartment of the community before formulating the optimization problem. (The presented measurements were performed with the models from the toy *E. coli* community.)
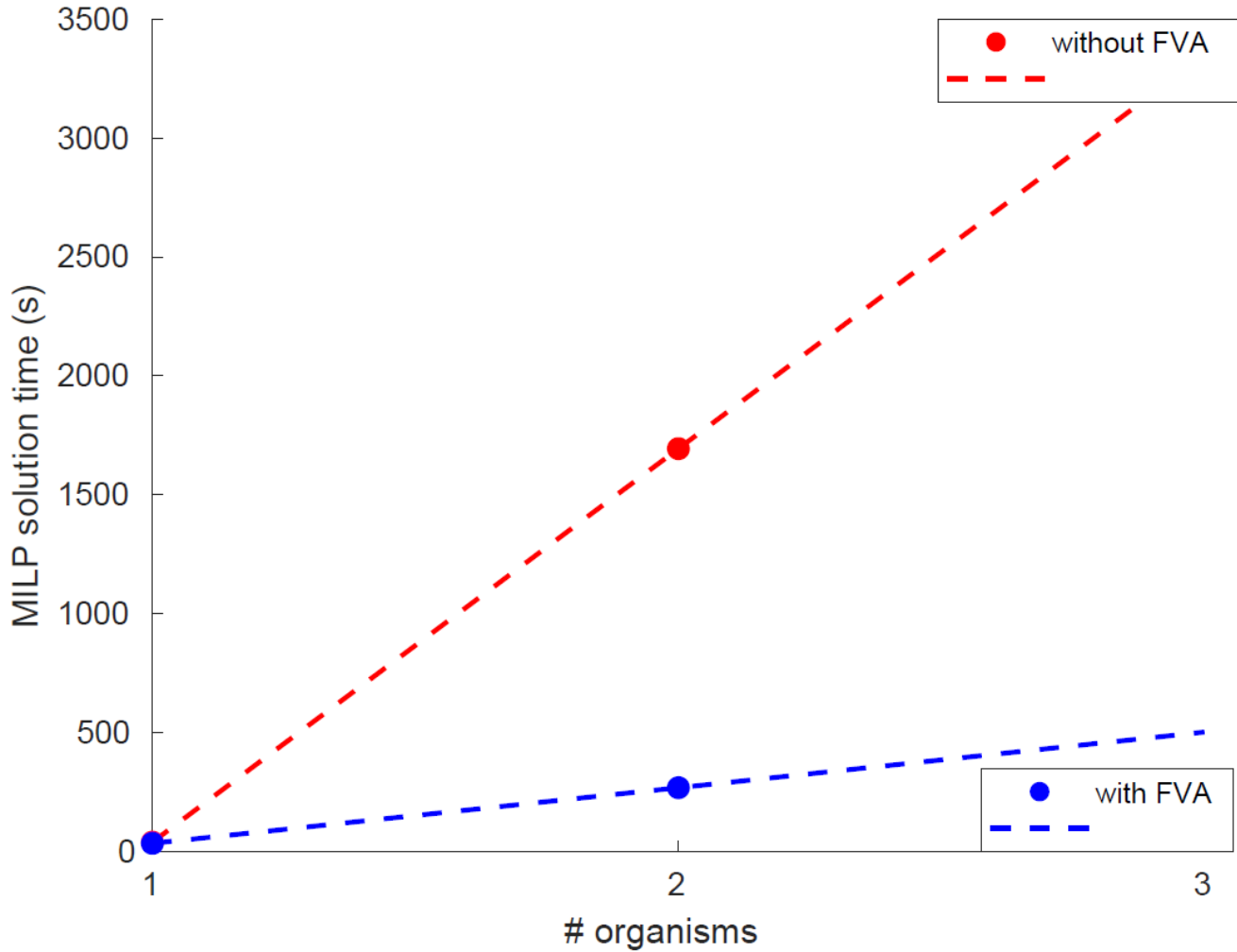
S2 Fig. Fluxes of Exchange Reactions from the ten best solutions for the community of *B. adolescentis* ATCC 15703 and *F. prausnitzii* A2-165. a) Fluxes of exchange reactions for the SCFAs acetate, butyrate, ethanol, lactate, formate, and succinate, and CO_2. **b)** Fluxes of exchange reactions for the amino acids alanine, arginine, aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, leucine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine.
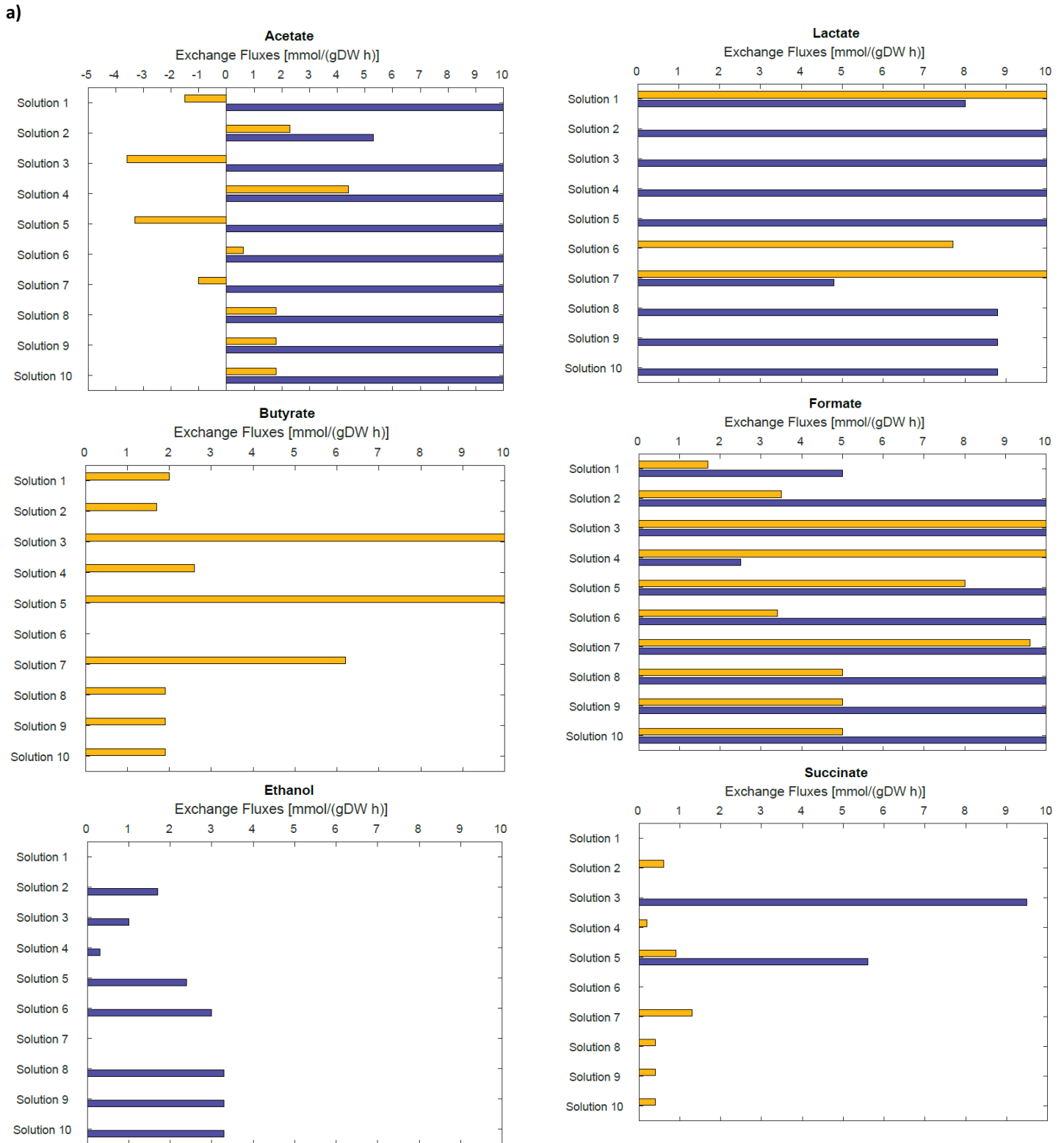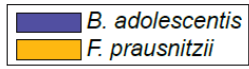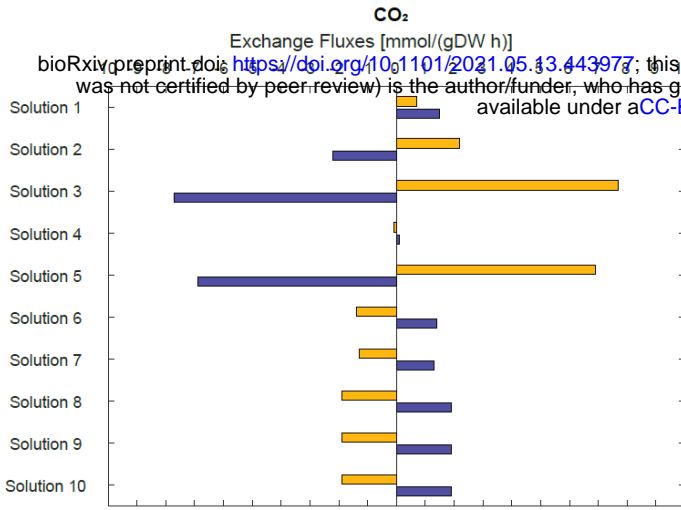
B. adolescentis
F. prausnitzii
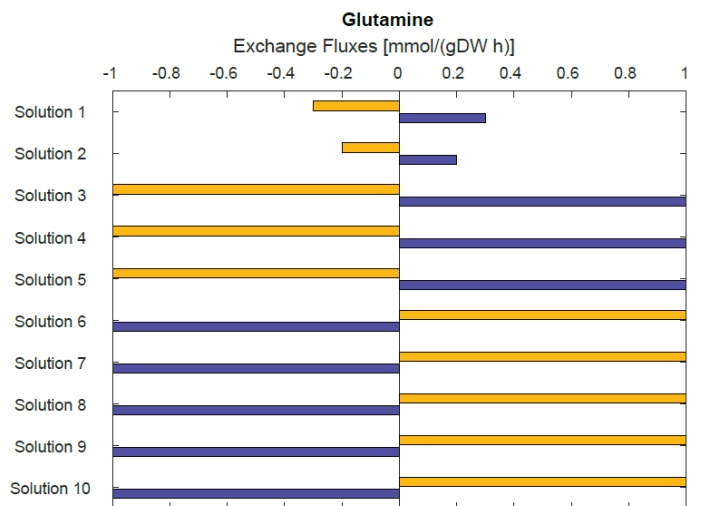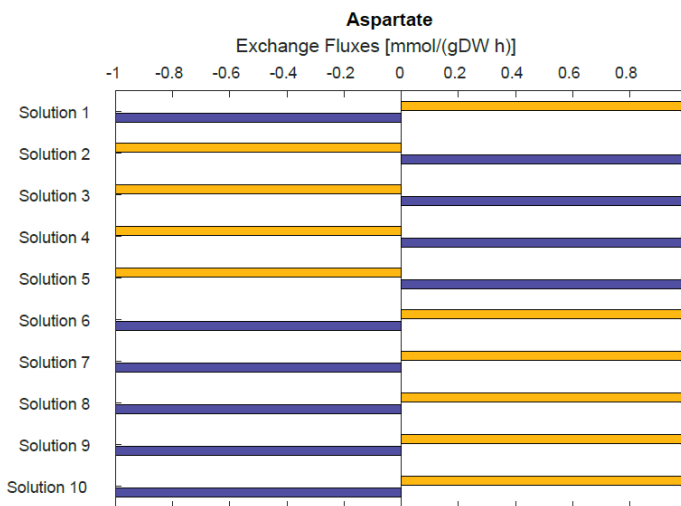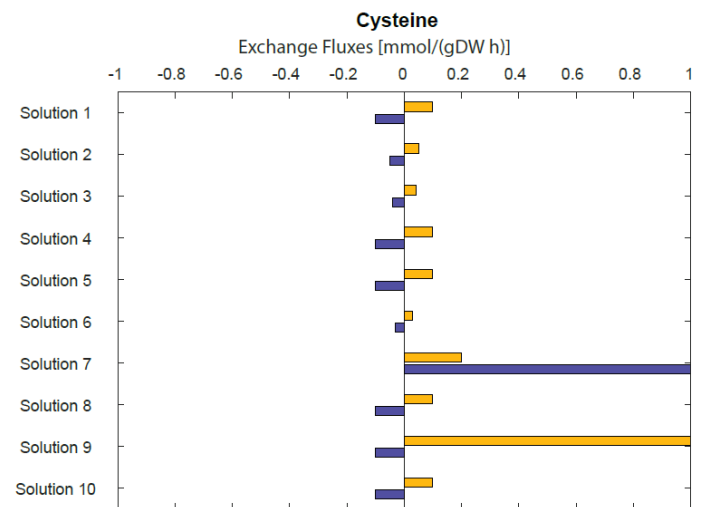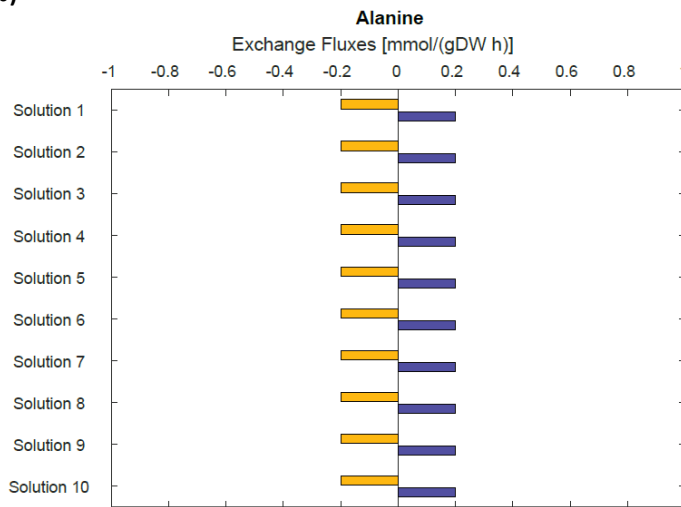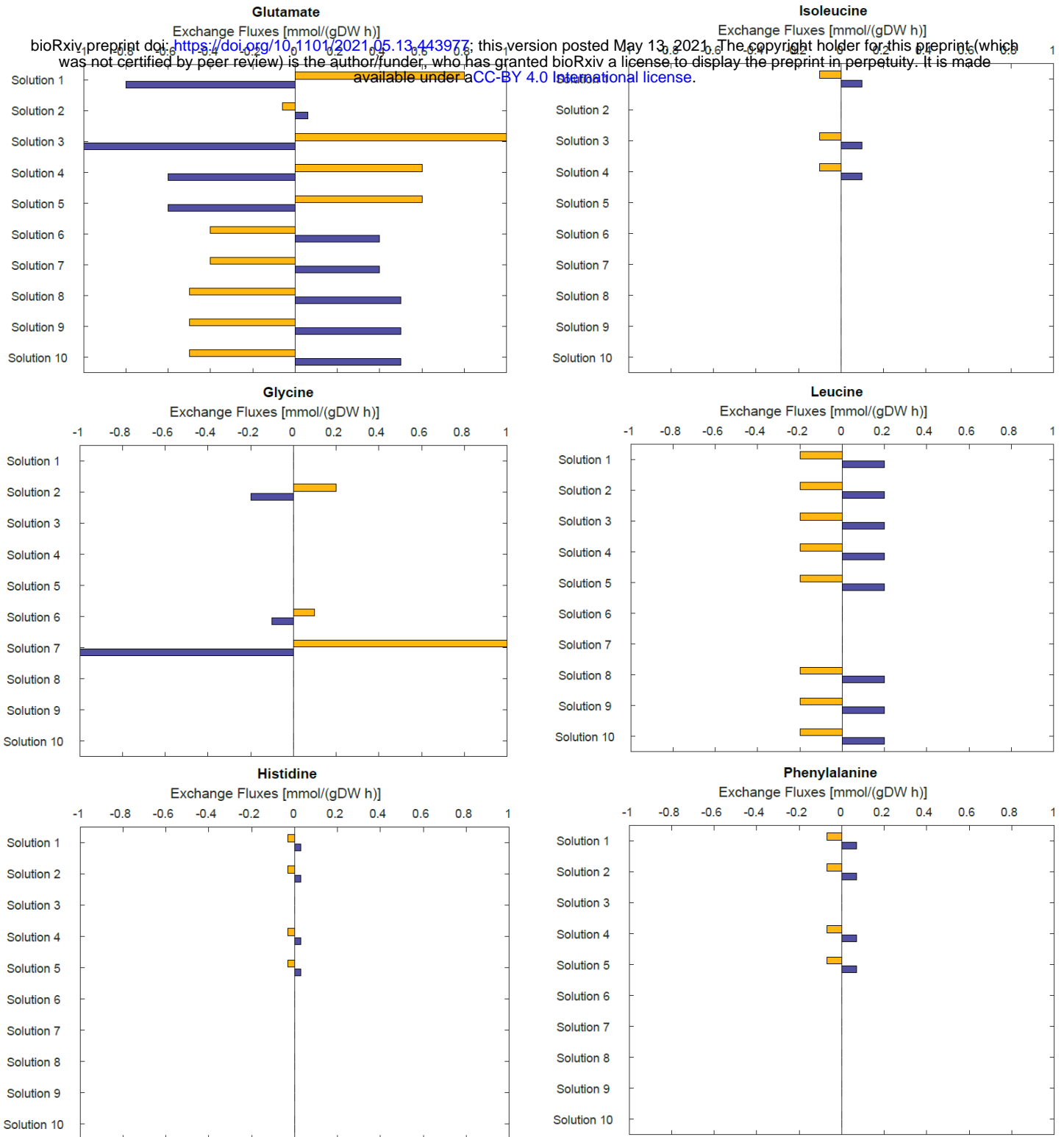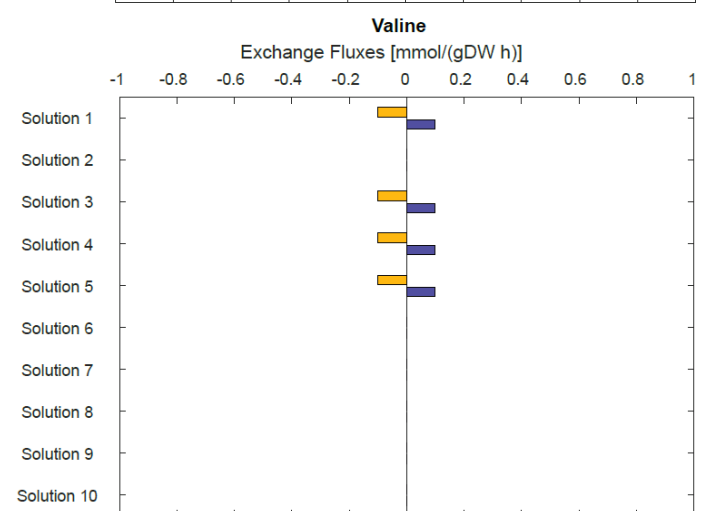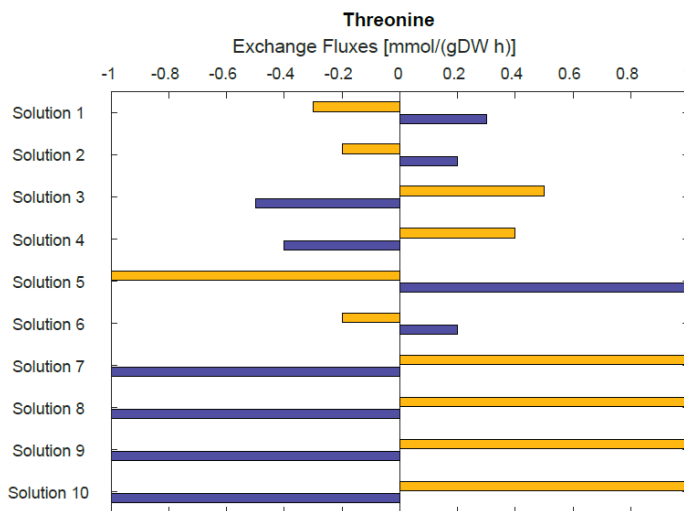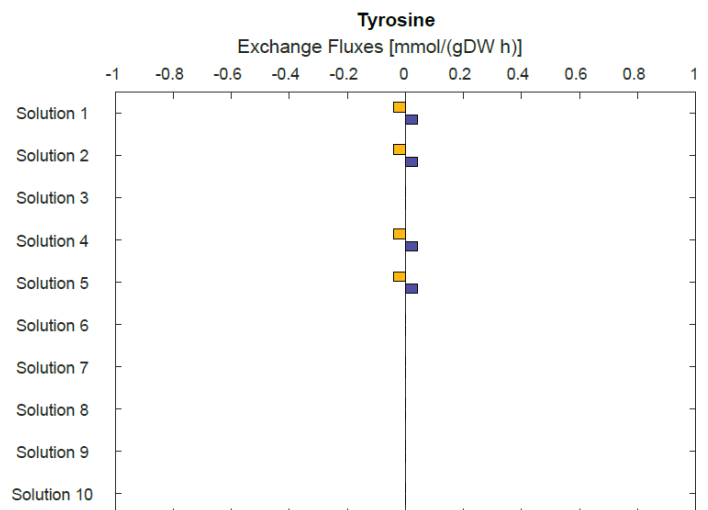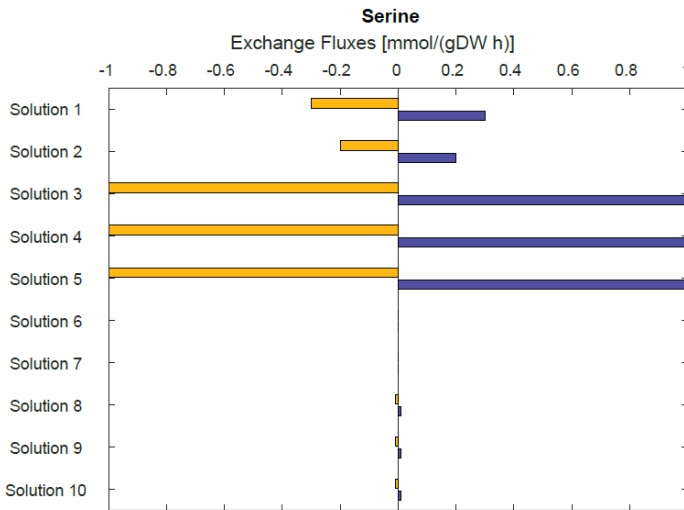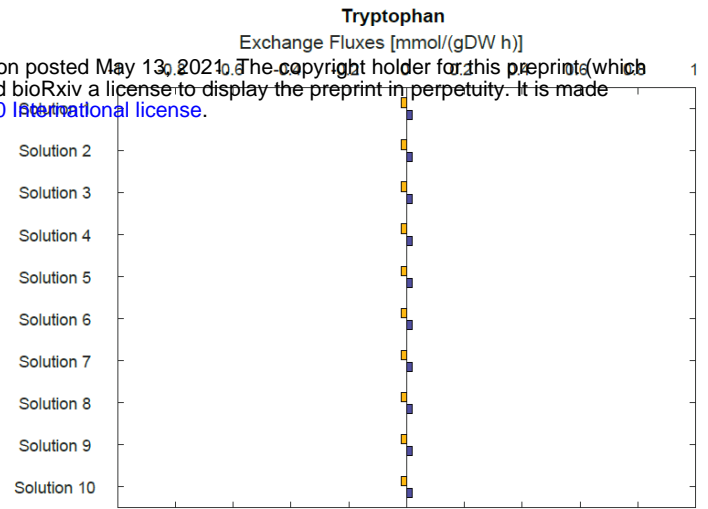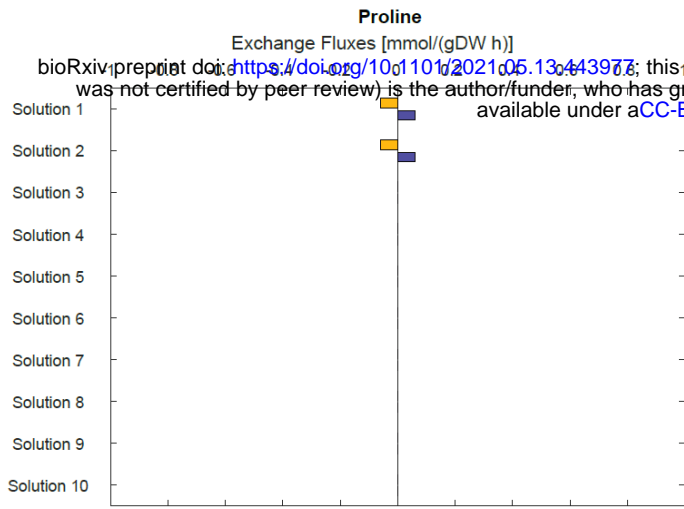
a)

S3 Fig. Fluxes of Exchange Reactions from the ten best solutions for the community of *Dehalobacter* sp. CF and *Bacteroidales* sp. CF50. a) Fluxes of exchange reactions for the organic acids acetate, malate, and pyruvate, CO_2 and H_2. **b)** Fluxes of exchange reactions for the amino acids alanine, arginine/asparagine, glutamate, isoleucine, lysine, phenylalanine, threonine, and tyrosine.

a)
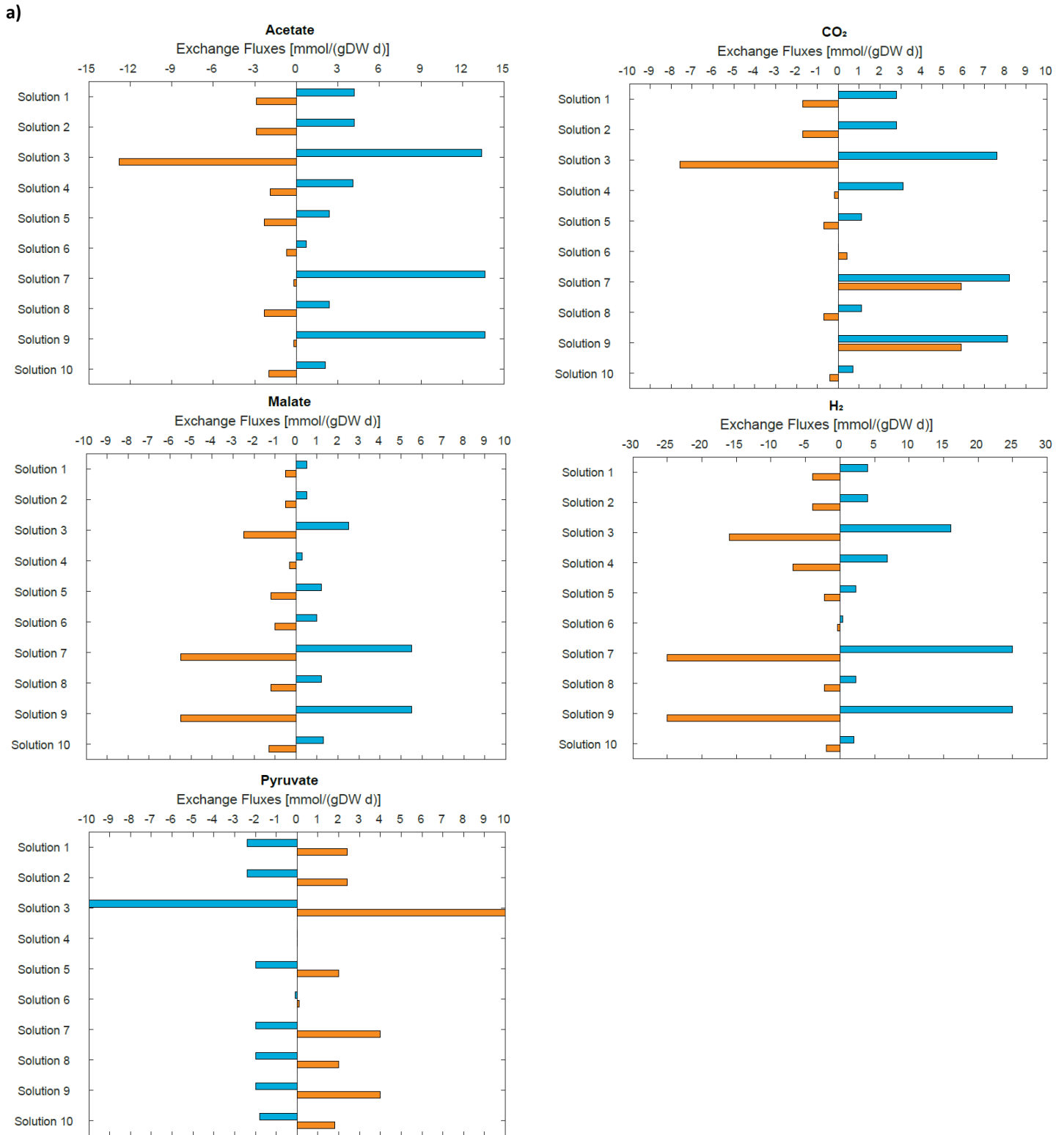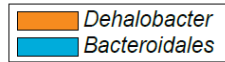
b)
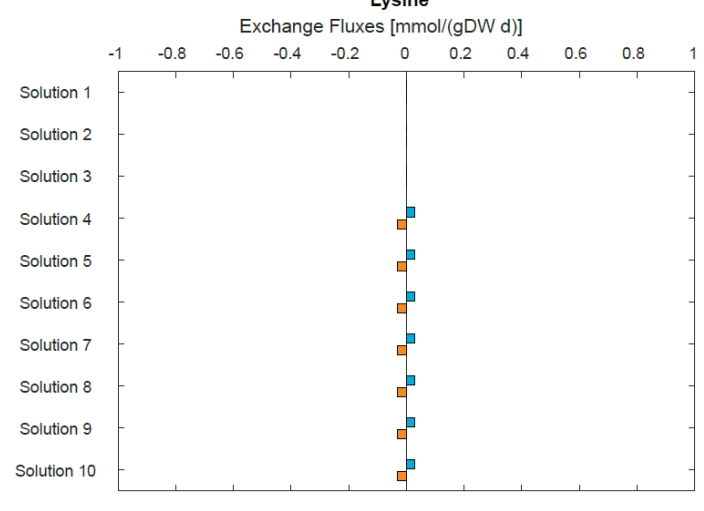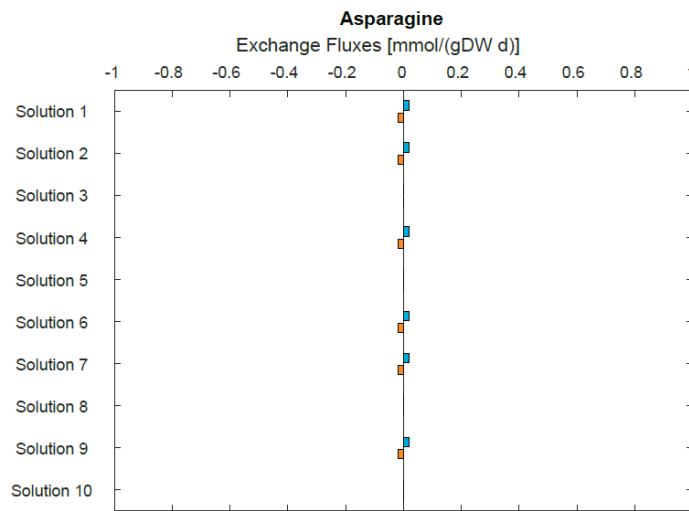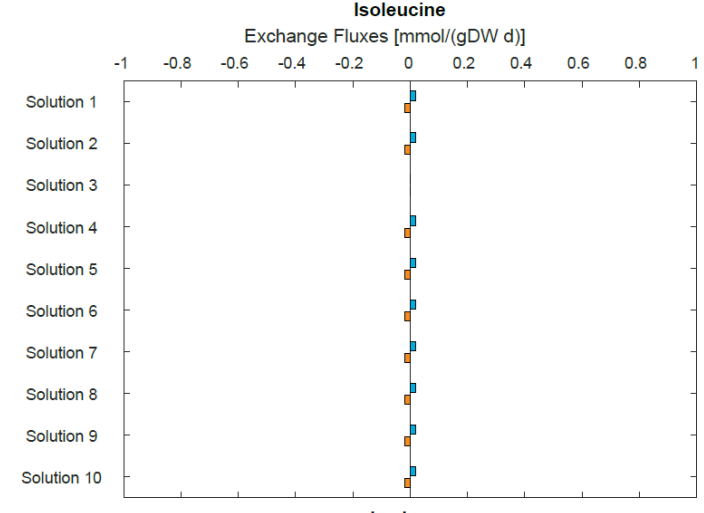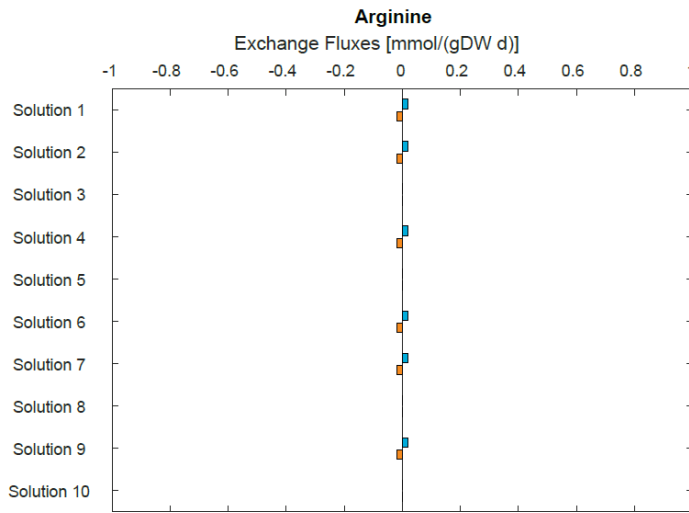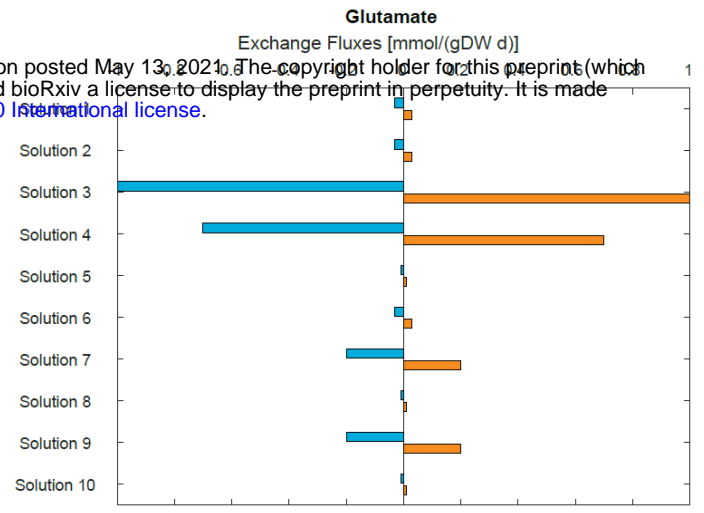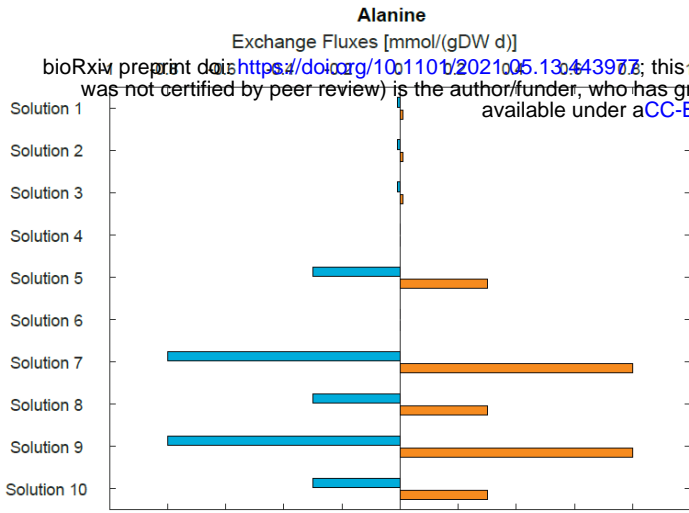
Phenylalanine

Exchange Fluxes [mmol/(gDW d)]
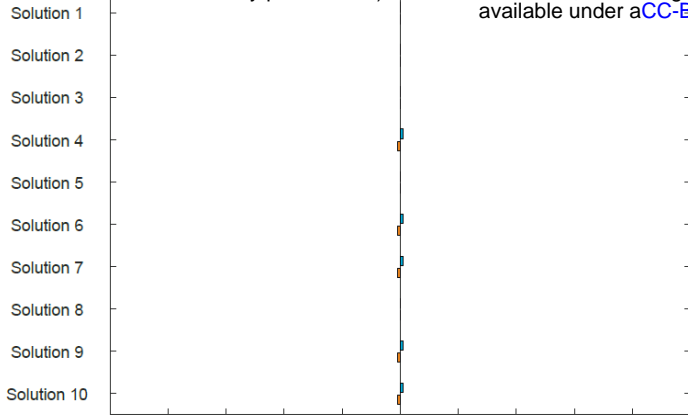
Threonine

Exchange Fluxes [mmol/(gDW d)]

Tyrosine

Exchange Fluxes [mmol/(gDW d)]