

Thermodynamic model of cross-feeding in a two-species microbial community

Tong Wang ^{1, 2, †}, Chen Liao ^{3, †}, Sergei Maslov ^{1, 2, 4, *}

1 Department of Physics, University of Illinois at Urbana-Champaign, IL 61801, USA

2 Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, IL 61801, USA

3 Memorial Sloan Kettering Cancer Center, 1275 York Avenue, NY 10065

4 Department of Bioengineering, University of Illinois at Urbana-Champaign, IL 61801, USA

† These authors contributed equally to this work.

* Correspondence: maslov@illinois.edu

Abstract

Even in the presence of the oxygen microbes often incompletely oxidize carbon substrates to produce overflow metabolites. These metabolites may in turn support the growth of other microbial species via cross-feeding interactions. Recent experiments revealed that pathways responsible for the overflow metabolism are often thermodynamically controlled. Here, we propose a thermodynamic modeling framework of cross-feeding in a two-strain microbial community. We introduce a coarse-grained single-cell model taking into account the thermodynamics of the overflow metabolism and proteome allocation constraints and use it to describe the steady state of a two-species cross-feeding microbial community in a chemostat. Our approach quantitatively describes the acetate overflow in *E. coli*, and accurately captures all carbon fluxes and growth rates under different environmental conditions. Furthermore it can correctly predict the response of a community composed of two strains of *E. coli* to changes in the environment. Our approach proposes modifications of flux and growth terms in the MacArthur's consumer-resource model accounting for the effects brought by thermodynamics of overflow pathways.

1 Introduction

At the first glance the overflow metabolism seems to be a wasteful metabolic strategy: despite being able to fully oxidize carbon sources such as glucose down to carbon dioxide, many microbial organisms including *E. coli* [1, 2] and yeast species *S. cerevisiae* [3] use incomplete fermentation pathways, which produce less ATP. Traditionally, the apparent wastefulness of the overflow metabolism was explained by proteome allocation constraints [4]. According to this hypothesis organisms allocate their proteome budget to pathways in order to maximize their yield-on-cost and thus achieve best overall growth rate. When overflow metabolites are excreted into the environment, they open up a new niche for cross-feeding species to use them as primary carbon and energy sources. In the Long Term Evolutionary Experiment [5] cross-feeding between a pair of *E. coli* strains has been repeatedly established. In this metabolic community, one of the strains evolved to consume the overflow metabolites such as acetate excreted by the other strain [5, 6].

Recently, many overflow pathways have been experimentally verified as thermodynamically controlled [7, 8, 9]. In these pathways, the production rate of the overflow metabolite may be reduced and even reversed as its extracellular concentration increases. For example, the dynamic

¹³C-metabolic flux analysis (MFA) confirmed that there is a strong bidirectional exchange of the acetate between *E. coli* and its environment [7]. Similarly, for *Clostridium ljungdahlii*, it has been also observed that the ethanol production is controlled by thermodynamics rather than by changes in enzyme expression [8]. Accumulation of the overflow metabolites in the environment poses a serious challenge to microbes producing them because it reduces their overflow flux, decreases available free energy, and ultimately lowers their growth rate. This problem is routinely solved by “cleaner” species that consume the accumulated end products. Thereby, the producer and the cleaner species establish a mutually beneficial cross-feeding relationship. However, how exactly the thermodynamically controlled overflow metabolism influences feasibility and stability of such microbial communities remains poorly understood.

Previous studies of the overflow metabolism in microbial communities used the MacArthur’s Consumer-Resource Model (CRM) [10], and dynamic Flux Balance Analysis (dFBA) [11]. While these approaches correctly capture the effects of the environmental concentration of the overflow metabolite on the growth rate of the consumer species, they ignore how it thermodynamically suppresses the growth of the producer species. dFBA is a constraint-based approach based on maximizing the growth rates of all species, while balancing fluxes through all of their intracellular metabolic reactions and transport reactions to and from the environment. As such, this approach is not sensitive to thermodynamic effects such as suppression of the growth rate due to accumulation of the overflow metabolite in the environment. Furthermore, while consumer resource models ignore all intracellular details, genome-scale models used in the dFBA approach include 1000s of intracellular metabolic reactions, making their predictions very difficult to interpret.

Here, we developed a new coarse-grained single-cell growth model taking into account the effects of thermodynamic constraints, proteome allocation, regulation of metabolic pathways on the overflow pathway. We extended this single-cell framework to model a simple microbial community consisting of two subspecies of *E. coli* cross-feeding over an overflow metabolite. We applied this model to quantitatively describe the coexistence between two subspecies of *E. coli* repeatedly evolved in the Long Term Evolutionary Experiment [12] as well as in shorter-term studies in a chemostat [6]. Finally, we propose simple mathematical expressions for all growth rates and nutrient fluxes in and out of individual species. These expressions can be used in general consumer-resource models with metabolic byproducts (see e.g. Refs. []) to accurately describe the thermodynamics of the overflow metabolism.

2 Results

2.1 Modeling framework

Motivated by the repetitive emergence of coexistence between two evolved cross-feeding *E. coli* strains in evolutionary experiments [6, 12], we aim to build a thermodynamic model of a two-strain microbial community populated by two strains evolved from the same ancestor. Due to their genetic and metabolic similarity, most kinetic parameters such as dissociation constants of pathways and the equilibrium constant of the overflow metabolism in these two strains are likely to be close to each other. Therefore, for simplicity, we first calibrate model parameters of both strains using fluxes through pathways and the growth rate based on the experimentally measured ¹³C-MFA data collected under different environmental conditions (fig. 1A) [7]. To calibrate strain-specific uptake rates of the primary resource and the overflow metabolite, we perform a fine-tuned calibration of the maximal rates of resource consumption to better match their OD growth curves in a particular growth medium (fig. 1B and 1C) [6].

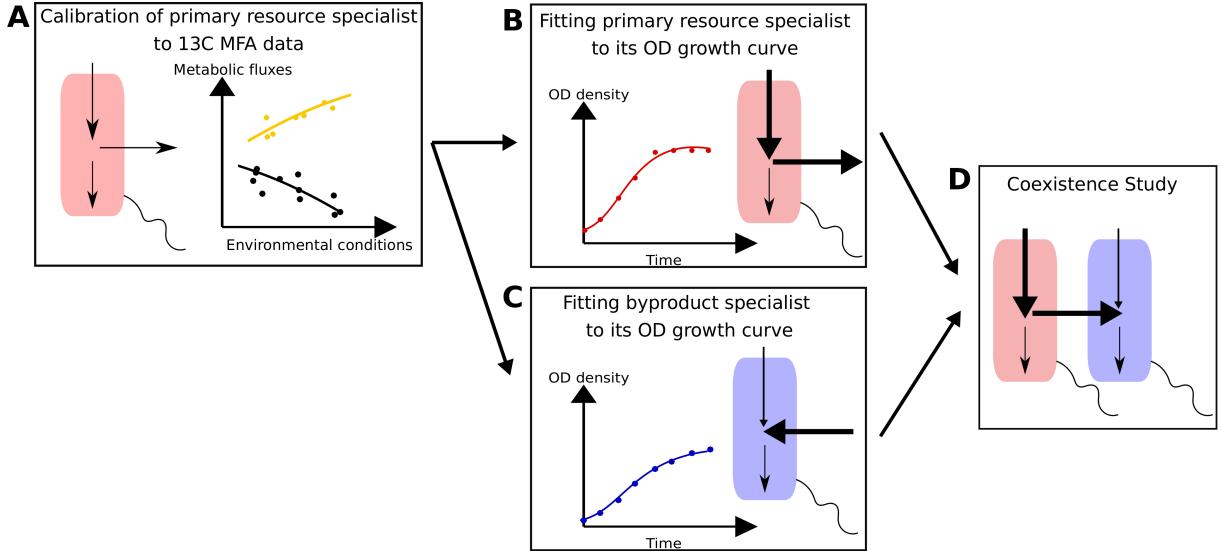


Figure 1: Step-by-step guide for building a thermodynamic model of cross-feeding in a two-strain microbial community. (A) Rough calibration of the kinetic parameters of both *E. coli* strains using the experimental data from ^{13}C -MFA (metabolic flux analysis). (B) Fine calibration of the uptake rates of the primary resource specialist to have a better match with its OD growth curve. (C) Fine calibration of the uptake rates of the byproduct specialist to have a better match with its OD growth curve. (D) The thermodynamic model of the microbial community with two strains (the primary resource specialist and the byproduct specialist) in a chemostat environment. Prediction of the coexistence state and its properties (strain abundances and metabolic fluxes).

Finally, we construct a computational model (fig. 1D) in which these two well-calibrated strains are co-cultured in a chemostat environment to predict whether they can coexist under different nutrient concentrations and dilution rates.

2.2 Model of the primary resource specialist

In fig. 2A, for each cell of the primary resource specialist, four coarse-grained metabolic fluxes linked by a precursor (such as pyruvate in *E. coli*) are modeled: the glycolysis flux J_g , the TCA cycle flux J_{tca} , the overflow metabolism flux J_o , and the biomass production flux J_b . The total available proteome is divided into two parts: ribosomes R and metabolic enzymes E . Since many metabolic pathways are regulated by global metabolic regulators such as Cra and Crp [13], the metabolic fluxes of the glycolysis J_g and the TCA cycle J_{tca} are assumed to be proportional to the proteome fraction E of metabolic enzymes. For comparison, the biomass production flux J_b is proportional to the amount of ribosomes R and the concentration of the precursor metabolite because the biomass is synthesized from precursors by ribosomes. The overflow flux J_o is assumed to be not influenced by the regulation due to the constitutive expression of the enzymes responsible for the overflow Pta-AckA pathway [7]. To take into account the fact that the growth of cells comes from the balance between the carbon flux J_c and the energy flux J_e , we assume the growth rate to be given by $g = \min(g_c, g_e)$ where g_b is the growth rate when it is limited by the carbon flux $J_c = J_b$ [SM: can we avoid using another term J_b if it is simply the same as J_c ? Also, we should clarify that n_g , n_{tca} , and n_o quantify the relative contributions of the energy generated pathways to the net energy. Where do we get the values of these parameters? We should refer to the relevant section of the SI. The letter e in J_e is very small and can be

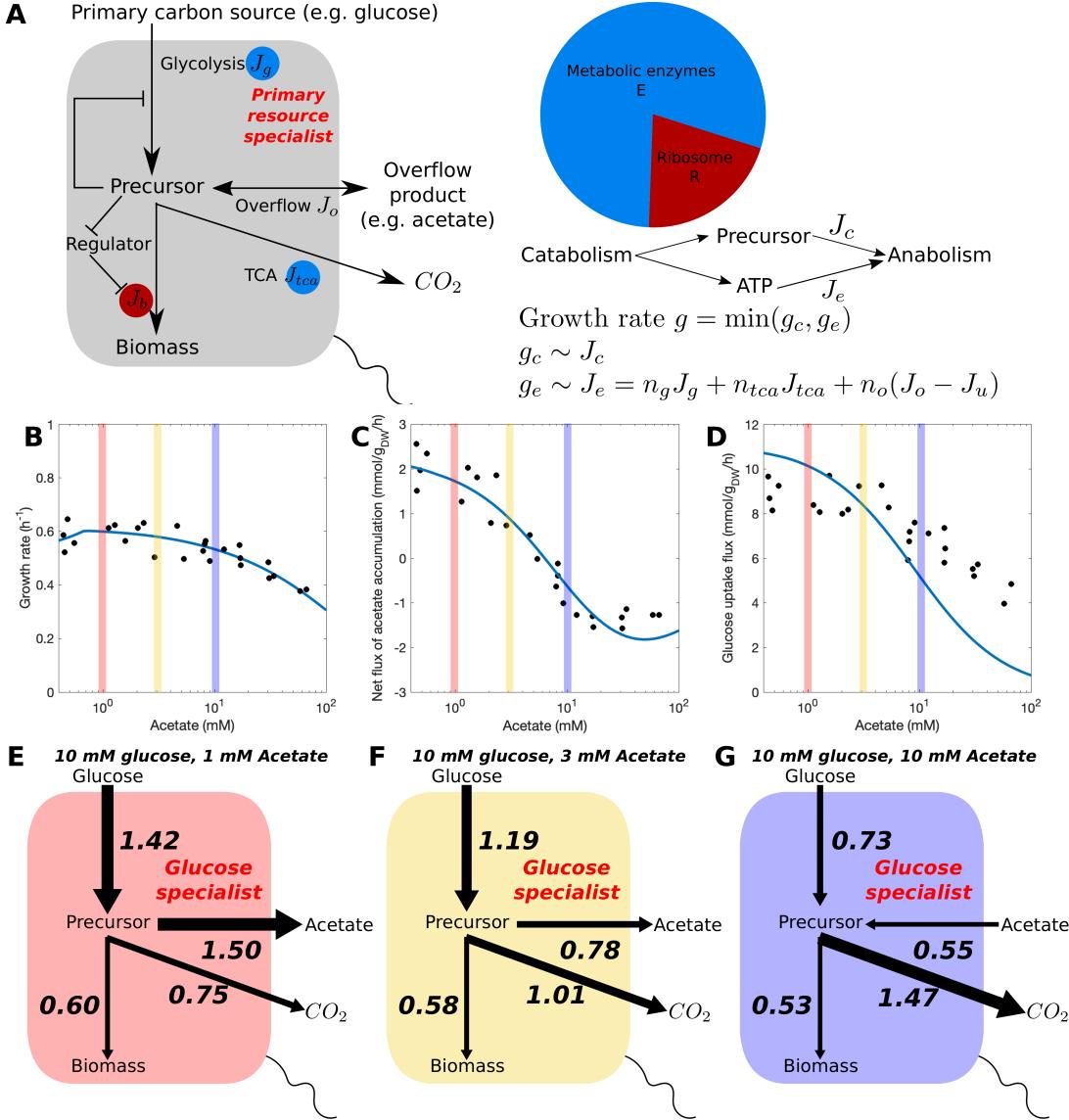


Figure 2: Modeling of the primary resource specialist and its calibration to the experimentally measured data from the ^{13}C metabolic flux analysis [7]. (A) The primary resource specialist is modeled as a single-cell organism with four coarse-grained metabolic pathways (glycolysis, overflow pathway, TCA cycle, and biomass synthesis) linked by a universal precursor such as pyruvate. Proteome allocation, regulation, and thermodynamic constraints on the overflow pathway are incorporated. The proteome comprises two parts: metabolic enzymes that regulate the glycolysis and the TCA cycle and ribosomes that control the biomass synthesis. The growth is considered as a balance of two fluxes: carbon precursor flux and energy flux. (B) The growth rates, (C) The net acetate accumulation, and (D) The glucose uptake flux of the primary resource specialist under different extracellular acetate concentrations (blue line: fitted model result; dots: experimental results from [7]). The balance of metabolic fluxes of the primary resource specialist when the extracellular acetate concentration is (E) 1 mM, (F) 3 mM, and (G) 10 mM.

easily confused with c . Perhaps we should replace it with $E.$, while g_e is the growth rate when it is limited by the energy flux $J_e = n_g J_g + n_{tca} J_{tca} + n_o J_o$. The total ATP flux is a sum of that through three ATP-generating pathways: glycolysis J_g , TCA cycle J_{tca} , and overflow metabolism

J_o weighted by the number of ATPs produced by each pathway. All metabolic flux terms are assumed to follow the Michaelis-Menten kinetics, except for modifications made to account for thermodynamic constraints on the overflow metabolism J_o identical to previously proposed thermodynamic models [14, 15]. More descriptions of the flux terms can be found in the Methods and Supplementary Information.

To demonstrate the utility of the single-cell modeling framework laid out above, we fitted the kinetic parameters such as maximum rates of pathways, dissociation constants, and the equilibrium constant of the overflow metabolism using nutrient consumption fluxes and growth rates of *Escherichia coli* K-12 MG1655 measured using ^{13}C -MFA technique under a wide range of acetate concentrations [7]. Figs. 2B-D show the best-fitted growth rates, fluxes of acetate accumulation [SM: replace accumulation to excretion everywhere in the text since it is not clear whether the acetate is accumulated extra- or intra-cellularly], and glucose uptake rates when the acetate concentration changes from 0.3 mM to 50 mM. Model-derived growth rates and fluxes (solid lines in 2B-D) agree with experimentally measured results (dots in 2B-D). Because of the thermodynamic nature of the overflow metabolism, the flux of the acetate accumulation reduces as the extracellular acetate concentration increases. In fig. 2C, the model captures a complete reversion from acetate production to consumption when the acetate concentration is above 8 mM, which has been experimentally observed [7]. As the acetate concentration increases, the reduced overflow metabolism results in the synthesis of more precursors from the acetate. Furthermore, it leads to a reduction of the glucose uptake rate shown in fig. 2D because the reduced overflow metabolism contributes to the accumulation of precursors and thus inhibits the glycolysis. The co-utilization of both glucose and acetate leads to a slight reduction in the growth rate as the acetate concentration increases (fig. 2B).

Figs. 2E-G dissect the metabolic fluxes under 3 acetate concentrations: 1, 3 and 10 mM. For 1 mM acetate, microbial cells consume glucose and the secreted acetate flux is 106% [SM: We need to explain how it can be above 100 percent. I understand it, but the breeders might be confused] of the glucose uptake flux. As the acetate concentration increases to 3 mM, the ratio between these two fluxes reduces significantly to 66%, although the overall metabolic pattern is similar to the previous one. As the concentration is further increased to 10 mM, *E. coli* starts consuming acetate instead of excreting it. In this regime, cells co-utilize two carbon sources (glucose and acetate). Though it may seem wasteful to spend ATP to uptake the acetate, further degradation of precursors through the TCA cycle generates sufficient ATP to maintain the growth rate of cells. Historically, extended versions of CRMs taking into account cross-feeding interactions, always assumed constant byproduct fraction — the ratio between secreted carbon sources and consumed carbon sources was assumed to be a fixed value [10, 16, 17, 18, 19]. However, this study demonstrates that taking into account the thermodynamic/reversible nature of the overflow metabolism results in byproduct fraction depending on the extracellular concentration of the overflow metabolite. Therefore, to better capture the ecological dynamics, using concentration-dependent byproduct fraction in models might be necessary.

2.3 Model of the byproduct specialist

The above model and its calibration were made for the wild-type *E. coli* K-12 strain. However, in evolutionary experiments, sub-strains using distinct metabolic strategies have repeatedly emerged. For example, in a glucose-limited culture initiated with a single clone of *E. coli* K-12, two evolved strains (CV101 and CV103) can establish a stable coexistence for thousands of generations because of their differential consumption abilities towards primary resource glucose and overflow metabolite acetate [6]. More specifically, three observations have been derived based on the

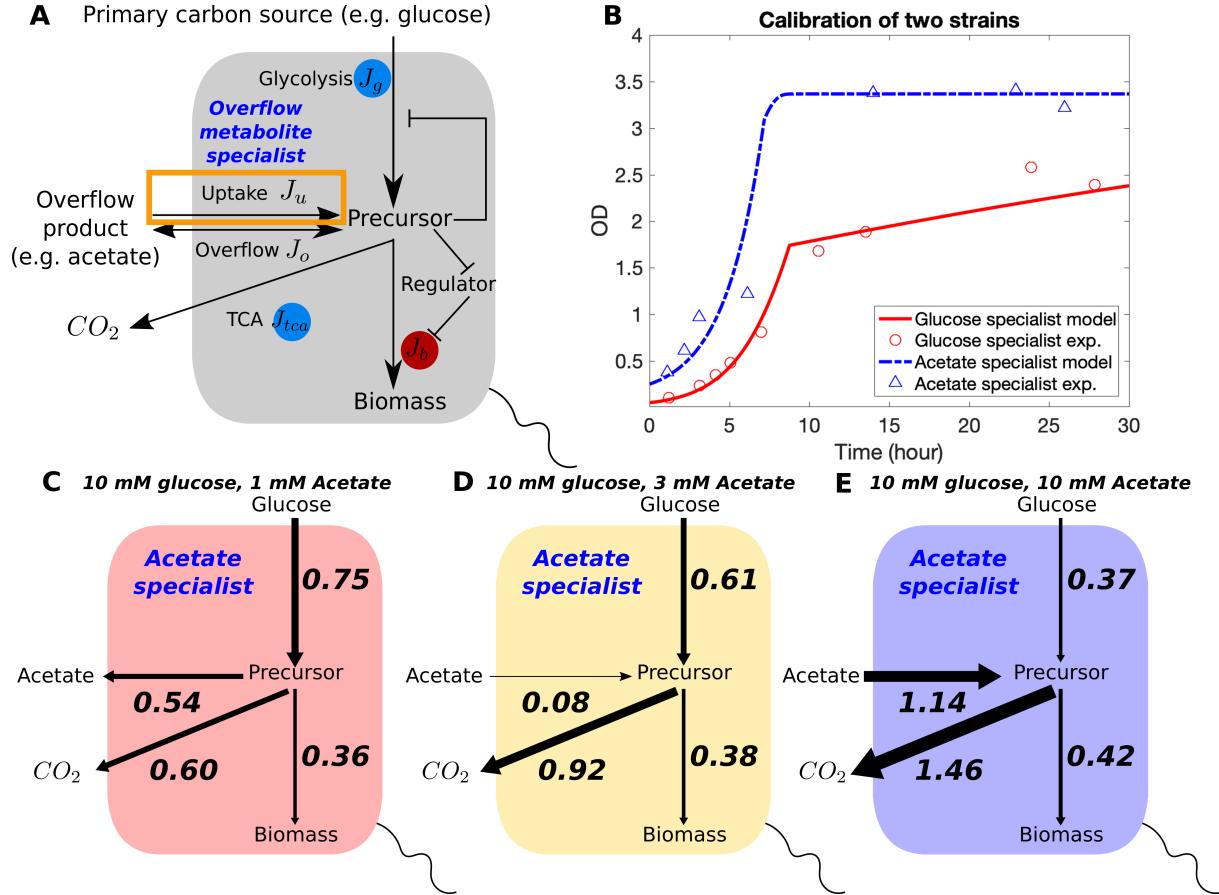


Figure 3: Modeling of the byproduct specialist and its calibration to experimentally measured data from polymorphism of *E. coli*. (A) The byproduct specialist has a unique high-affinity acetate uptake pathway (orange box). (B) Kinetic parameters of uptake rates for both strains were calibrated to their OD growth curves [6]. The balance of metabolic fluxes of the byproduct specialist when the extracellular acetate concentration is (C) 1mM, (D) 3mM, and (E) 10mM.

experimental examination of both strains' physiology [6]: (1) the glucose uptake rate of CV103 is much stronger than that of CV101, (2) CV103 produces acetate that is mainly consumed by CV101, and (3) the high-affinity acetate uptake pathway Acetyl-CoA Synthetase (ACS) is barely expressed in CV103, while it is constitutively overexpressed in CV101.

To capture the physiology difference, we modified the model setup for the overflow metabolite specialist in fig. 3A by adding a high-affinity overflow metabolite uptake pathway (shown in an orange square box) with its flux specified as J_u . The growth medium, used in evolutionary experiments [6], where microbial polymorphisms and coexistence were observed, is Davis and Mingoli minimal medium, quite different from the M9 minimal medium used in the previous ^{13}C -MFA study [6]. To make fine adjustments of maximum uptake rates of glucose and acetate in case of Davis and Mingoli minimal medium, these rates were adjusted to best fit OD growth curves of both *E. coli* strains growing on the medium with glucose supplied as the sole carbon source. The best-fitted OD growth curves generated by the model (solid lines in fig. 2B) match the experimentally observed OD growth curves (dots and triangles in fig. 2B).

For three different extracellular concentrations of acetate, the distribution of metabolic fluxes for the acetate specialist (fig. 3C-E) differs from that for the glucose specialist (fig. 2E-G).

Overall, as the extracellular acetate concentration increases from 1 mM to 10 mM the metabolic strategies of the two strains show a similar pattern, shifting from glucose consumption generating acetate as a metabolic byproduct to co-utilization of glucose and acetate. Besides similar changes in the metabolic pattern, there are 3 major differences between the two strains. First, the byproduct fraction is lower for the acetate specialist than the glucose specialist. For example, for 1 mM acetate, the byproduct fraction for the acetate specialist is 72%, compared to 106% for the glucose specialist. Second, the transition from acetate production to consumption in the acetate specialist happens when the acetate concentration is around 3 mM, which is lower than 8 mM for the glucose specialist. Third, as the acetate concentration increases, the growth rate of the acetate specialist increases from 0.36 to 0.42 hour⁻¹, showing an opposite trend to acetate inhibiting the growth of the glucose specialist (see growth rates in figs. 2E-G).

2.4 A Consumer-Resource Model of the two-strain microbial community

The MacArthur's consumer-resource model (CRM) and its variants [20, 21] are widely used to describe the process of resource consumption by biological species. The central tenet of a CRM is simple: depletion of resources gives rise to the competition between consumer species. CRM requires relatively small number of parameters and has a good interpretability. However, the traditional CRM and its extensions largely ignore important factors such as proteome allocation and thermodynamic constraints giving rise to product inhibition of individual reactions. These simplifications may result in completely wrong assumptions regarding the fraction of metabolic byproduct excreted under different environmental conditions. Here, we incorporate the single-cell thermodynamic model described above into a modified CRM and predict the ecological dynamics of a two-strain microbial community. We first use our single-cell model to generate all fluxes and growth rates under different extracellular concentrations of glucose and acetate for both strains: the glucose specialist and the acetate specialist (shown as dots in figs. 4A-D). Then we propose a mathematical modification of nutrient uptake and growth terms (curved surfaces in figs. 4A-D) in the CRM that fit the above data remarkably well.

To account for the inhibition effect brought in by the acetate, a mass-action multiplier with the acetate concentration was added to the glucose uptake rate [SM: can we drop (CRM) from these formulas? It makes them hard to read. Also, I would also drop the square brackets for concentrations since this notation is rarely used in CRM literature] :

$$J_g = \frac{k_g^{(\text{CRM})}[G]}{K_g^{(\text{CRM})} + [G]} \frac{K_{ig}^{(\text{CRM})}}{K_{ig}^{(\text{CRM})} + [A]}, \quad (1)$$

where $[G]$ is the concentration of glucose, $[A]$ is the concentration of acetate, $k_g^{(\text{CRM})}$ is the maximum uptake rate of glucose, $K_g^{(\text{CRM})}$ is the Monod constant for the glucose uptake, and $K_{ig}^{(\text{CRM})}$ is the inhibition constant of the acetate on the glucose uptake. In our fitting procedure, we use $K_g^{(\text{CRM})} = 5 \mu\text{M}$ previously reported in the literature [22]. The acetate accumulation flux is decomposed into two terms. The negative term accounts for the overflow metabolism, and it is proportional to the glucose uptake rate. The positive term accounting for the acetate uptake comes from either the reverse uptake of the acetate through J_o or the via the high-affinity uptake pathway J_u . Thus, we fitted the acetate uptake flux into the cell as:

$$J_a = -f \frac{k_g^{(\text{CRM})}[G]}{K_g^{(\text{CRM})} + [G]} \frac{K_{ig}^{(\text{CRM})}}{K_{ig}^{(\text{CRM})} + [A]} + \frac{k_a^{(\text{CRM})}[A]}{K_a^{(\text{CRM})} + [A]} \frac{K_{ia}^{(\text{CRM})}}{K_{ia}^{(\text{CRM})} + [A]}, \quad (2)$$

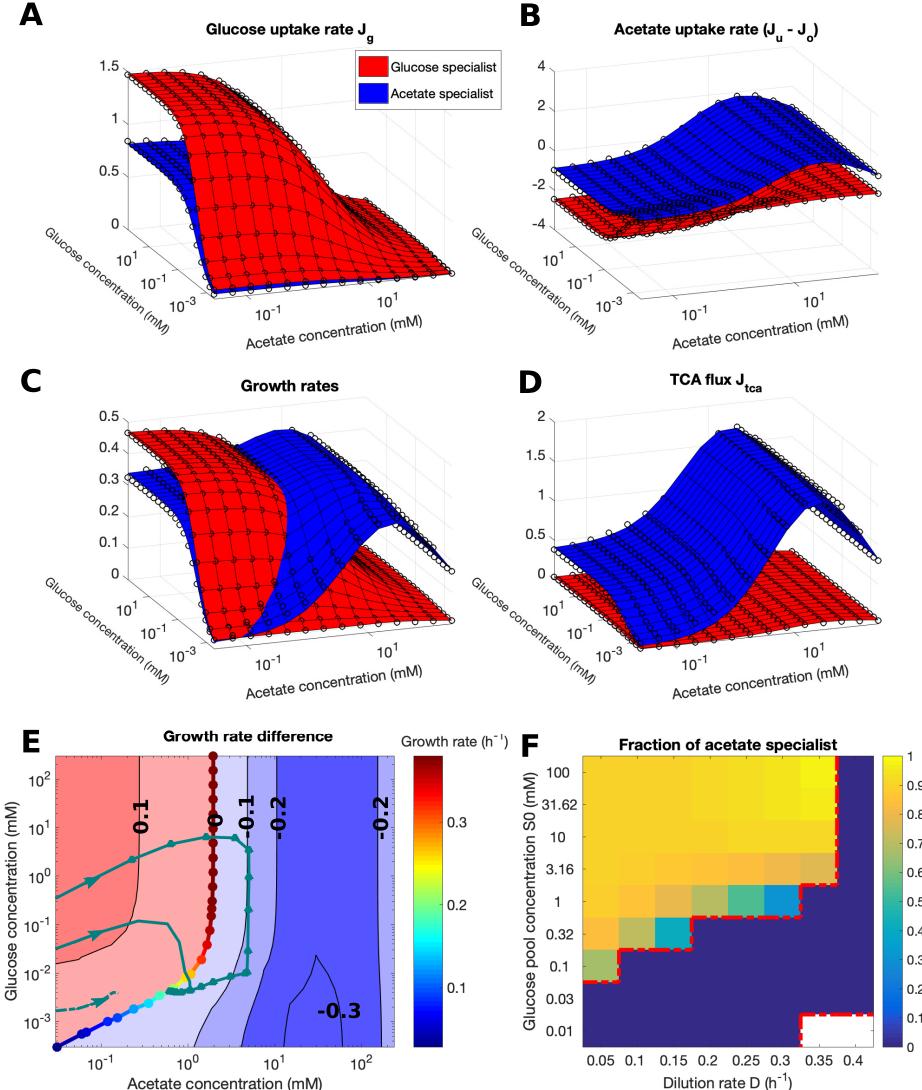


Figure 4: Conditions for coexistence of the glucose specialist and the acetate specialist in a chemostat environment. (A) Glucose uptake rates, (B) acetate uptake rates, (C) growth rates, and (D) fluxes through the TCA cycle plotted as a function of extracellular glucose and acetate concentrations for the glucose specialist (red) and the acetate specialist (blue). All dots are the results generated by our single-cell model, while curved surfaces are the best-fits from the modified Consumer-Resource Model proposed in our study. (E) The difference between growth rates of two strains. The region in which the growth rate of the glucose specialist is larger than that of the acetate specialist is colored red, while the region where the opposite inequality holds is colored blue. Growth rates on the equal-growth-rate isocline (EGRI) where two strains have the same growth rates are color-coded and the values are given in the color bar. Three trajectories colored green (the solid green line with triangles, the solid green line without any symbols, and the dashed line) correspond to initial glucose concentrations of 10mM, 1mM and 0.1 mM respectively. Note the absence of coexistence steady state in case of the dashed line. (F) The fraction of the acetate specialist in the steady-state simulated with different glucose pool concentrations and dilution rates. Two red dashed lines separate the phase space into three regimes arranged from top to bottom: coexistence of two strains, exclusion of the acetate specialist, and extinction of both strains.

where f is the ratio between the glucose uptake flux and the acetate overflow flux excreted from the cell (note that f could exceed 1 due to the difference in the number of carbon atoms in glucose and acetate molecules), $k_a^{(\text{CRM})}$ is the maximum uptake rate of acetate, $K_a^{(\text{CRM})}$ is the Monod constant for the acetate uptake, and $K_{ia}^{(\text{CRM})}$ is the inhibition constant of the acetate in the acetate uptake flux. The best fitted parameters for the modified CRM above are summarized in the table 1. [SM: we need standard deviations in this table. In the SI we may include the standard parameter sensitivity analysis including correlations between fitted values of all parameters. How accurate are our fits of large constants 139 and 81.8?]

Table 1: Fitted parameters for the modified CRM. [SM: We need to say something about parameters so large that we don't believe them to be true. We also should include the error bars.]

| Parameters | Unit | Glucose specialist | Acetate specialist |
|-------------------------|--------------------|--------------------|--------------------|
| $k_g^{(\text{CRM})}$ | hour ⁻¹ | 1.48 | 0.82 |
| $k_a^{(\text{CRM})}$ | hour ⁻¹ | 1.30 | 2.76 |
| f | none | 1.66 | 1.10 |
| $K_a^{(\text{CRM})}$ | mM | 81.8 | 5.09 |
| $K_{ig}^{(\text{CRM})}$ | mM | 7.45 | 6.53 |
| $K_{ia}^{(\text{CRM})}$ | mM | 14.4 | 139 |

Under the energy-limited conditions, three independent fluxes J_g , $J_a (= J_u - J_o)$, J_{tca} , and the growth rate λ satisfy the two constraints: (1) the balance of three carbon fluxes and the carbon flux into the biomass defined by the growth rate λ (conservation of carbon), [SM: perhaps we shouls explicitly write down the equation for carbon balance since it uses well-known integer stoichiometry (the number of carbon atoms in individual molecules).] and (2) the energy-limited growth rate λ is a linear combination of three fluxes (conservation of energy) [SM: how about the energy lost as heat? We need to rephrase the statement about the conservation of energy since it is confusing]. we selected J_g and J_a as two independent variables, and allowed λ and J_{tca} to be determined by J_g and J_a via the above two conservation laws. The mathematical derivation of these formulas is shown in supplemental information 2.4. Curved surfaces, shown in fig. 4C-D, were derived using $\lambda = 0.60J_g + 0.17J_a$ and $J_{tca} = 1.40J_g + 0.83J_a$.

The fitting of parameters in the modified CRM naturally gives us a way to understand the cell physiology and the ecological relationship between the two strains. There are several lessons one can learn. First, the above formulas show that as in the MacArthur's CRM, the growth rate of two strains is determined by a linear combination of two carbon sources (e.g. glucose and acetate) taken up independently. Second, we found that in order to properly model product inhibition due to thermodynamics, fluxes of the primary carbon source (glucose) and of the overflow metabolite (acetate) have an extra suppression term. Third, the best fit parameters of the modified CRM give a good interpretation of differences in metabolic strategies used by two strains: (1) the maximal uptake rate $k_g^{(\text{CRM})}$ of the glucose is higher in the glucose specialist than in the acetate specialist, (2) the maximal uptake rate $k_a^{(\text{CRM})}$ of the acetate is much lower in the glucose specialist than in the acetate specialist, (3) The Monod constant $K_a^{(\text{CRM})}$ for the acetate is much higher in the glucose specialist than that of the acetate specialist, and (4) the byproduct fraction f of the glucose specialist is higher than that of the acetate specialist. [SM: say something along the lines that our fits with Monod constants with values around 100mM have a large margin of error (we need to add \pm values to the table). For most experimental conditions the uptake of these substrates would be far from saturation. In the linear regime of

the Mondo curve experimentally measurable parameter is given by the ratio k/K .]

2.5 Stable coexistence of cross-feeding strains

To model steady state properties of the ecosystem populated by two cross-feeding strains, we extended our single-cell model to incorporate the population dynamics of both strains (glucose and acetate specialist) in a chemostat environment. Fig. 4E shows the growth rate difference between the glucose specialist and the acetate specialist under different concentrations of glucose and acetate in the chemostat. Fig. 4E is derived by subtracting the growth rates of two strains shown as curved surfaces in fig. 4C. In order for two strains to be able to coexist in a chemostat, their growth rates have to be equal to each other and to the dilution rate D of the chemostat. The one dimensional curve in fig. 4C where two strains have the same growth rates is called the equal-growth-rate isocline (EGRI). On this EGRI, marked with rainbow-colored symbols in fig. 4E, the real growth rates of two strains are equal to each other. The value of the growth rate is shown in the color bar. Two coexisting strains deplete resources until they end up at the EGRI. The exact steady-state concentrations of two resources are determined by the intersection between the EGRI [SM: can we mark the EGRI in panel C with the same markers and colorbar as in panel E?] in fig. 4C and the horizontal plane at which the growth rate is equal to D . Thus, as is always the case in CRMs, the steady-state concentrations of resources don't depend on their supply fluxes, but are fully determined by the dilution rate D of the chemostat. To illustrate this point for $D = 0.2 \text{ hour}^{-1}$, two solid green lines are trajectories of resource concentrations for the case where the glucose is supplied as the sole carbon source with different concentrations: 1 mM (solid green line) and 10 mM (solid green line with triangles). Two lines eventually end at the same point with identical steady state concentrations because this position is the only place that satisfies the environmental condition required for coexistence between strains when the dilution rate $D = 0.2 \text{ hour}^{-1}$.

However, the mere fact that the EGRI and intersect the horizontal plane at which the growth rate is equal to D , does not guarantee coexistence of two strains. An interesting and important feature of the above two simulations (with glucose pool concentration 1 mM and 10 mM) is that both trajectories in the space of resource concentrations cross through EGRI from the left side to the right side as glucose depletes and acetate accumulates (fig. 4E). This crossing over of the EGRI marks a shift of the chemostat's environment from the growth rate of the acetate specialist from being smaller than that of the glucose specialist (the left side of the EGRI) to being larger than it (the right side of the EGRI). This crossing over is a reflection of the negative frequency dependent selection between two cross-feeding microbial strains, where the growth rate of one strain is larger when the frequency of this strain is lower. Therefore, by lowering the glucose pool concentration to 0.1 mM (shown as the lowest green line [SM: dashes on this line are not visible] in fig. 4E), the accumulation of acetate cannot cross the EGRI thereby destroying the coexistence between two strains. To explore the influence of the glucose pool concentration and the dilution rate of the chemostat on the coexistence of two strains, the fraction of the acetate specialist as a function of two environmental factors is shown in fig. 4F. Overall, there are three regimes: (1) extinction of both strains (the white region in the lower right corner of fig. 4F), (2) survival of glucose specialist only (the dark blue region in the middle of fig. 4F), and (3) coexistence (the multicolored region around the upper left corner of fig. 4F). When the glucose pool concentration is less than the glucose concentration of the intersection between the EGRI and the horizontal plane at growth rate D , neither of the two strains can survive. Otherwise, at least the glucose specialist can survive. The survival of the acetate specialist depends on whether the dynamic trajectory of environmental concentrations crosses the EGRI.

3 Discussion

In this study, we introduced a model of a two-species microbial community growing on glucose and cross-feeding acetate. Our coarse-grained description of intracellular dynamics takes into account thermodynamically-controlled overflow metabolism, proteome allocation, and gene regulation. To demonstrate the predictive power of this general modeling framework, we calibrated our model to fit the experimental data, such as ^{13}C -MFA and growth curves of two *E. coli* strains. We then computationally predicted the patterns of coexistence and competition between these two strains in a chemostat. To our knowledge, the proposed single-cell model is among the first to explore the influence of thermodynamics on physiology of cells and stability of cross-feeding interactions.

Overall, coarse-grained models of intracellular dynamics may provide a good middle ground between traditional consumer resource models in which intracellular metabolic reactions are completely ignored [10, 17] and Flux Balance Analysis (FBA) models, which include too detailed description of the intracellular metabolism (100s to 1000s of reactions) at the price of omitting important factors such as thermodynamics of reversible reactions, gene regulation, and limitations imposed by the proteome allocation [23, 24, 11, 25]. Coarse-grained models (1) are easy to calibrate based on limited experimental data: many kinetic parameters are ignored, and the remaining parameters can be inferred from single-cell experiments, (2) are generalizable to a wide range of environmental conditions: they take into account how gene regulation, proteome allocation, and thermodynamics of metabolic pathways are shifting in response to changes in the environment (3) are interpretable: the model variables such as nutrient fluxes, growth rates, and half maximal inhibitory concentrations correspond to important ecological properties of the community. Mismatches between the predictions of the model and the experimentally observed ecological patterns may suggest which extra properties need to be added to the coarse-grained description of intracellular dynamics.

Our model predictions of two-strain coexistence for different environmental concentrations of glucose and acetate approximately agrees with the existing experimental data [6] [SM: What exactly agrees with these experiments? Can you be more specific?]. This agreement is remarkable given that kinetic parameters for two strains in our model were calibrated based only on their OD growth curves. This shows the power of our approach, in which we first calibrate the intracellular model of each strain for a wide range of environmental conditions and only later assemble these strains into a microbial ecosystem.

Our model also suggest several general modifications one could make to standard consumer resource models for a more realistic description of cross-feeding via overflow metabolites. One modification is the inhibitory effect of the cross-feeding metabolite (acetate in our case) on the growth rate and nutrient uptake rates of all species. Such product inhibition, discussed e.g. in Ref. [15], is an inevitable consequence of the thermodynamics of the overflow metabolism. We found that a standard multiplier with the Hill coefficient equal to 1, $A/(K_A + A)$, adequately describes this inhibitory effect in our model (see Fig. fig4).

The other modification is that the rate of uptake/excretion of the cross-feeding metabolite in general depends on its concentration in the environment. In fact, for high concentrations the sign of the excretion rate can be reversed, corresponding to the net uptake of the overflow metabolite by the species. This is exactly the strategy adopted by the acetate specialist in our model, which consumes acetate instead of excreting it to the environment. Many traditional CRMs [10, 16, 17, 18] assume that a fixed fraction of carbon consumed as primary metabolites is always excreted to the environment in the form of metabolic byproducts. Our model shows that instead of being a stoichiometric constant, this fraction decreases and perhaps even changes sign as metabolic byproducts are being accumulated in the environment.

Our model can also make several predictions on when one would expect coexistence of two cross-feeding strains. First, two different strains can only coexist when they differ significantly in their metabolic strategies. Specifically, survival of the acetate specialist depends on irreversible high-affinity uptake pathway of acetate. Without such a high-affinity acetate uptake pathway, this strain cannot claim acetate as its niche. [SM: Where do we show this?] Second, thermodynamic properties of the overflow metabolism cause acetate to inhibit the growth of the glucose specialist, while increasing the growth rate of the acetate specialist. Such negative-frequency dependent selection acts on both strains helping to stabilize their coexistence. Third, species in our model are energy-limited (instead of carbon-limited) for most of environmental conditions. This provides an intuitive explanation of why they excrete the cross-feeding metabolite (acetate). Indeed, growing cells need to balance the carbon precursor flux and the energy flux. When the carbon precursor flux is larger than the energy flux, this excess of carbon needs to be excreted to the environment as metabolic byproducts. This also explains why the byproduct fraction is usually higher for anaerobic conditions than for aerobic conditions. Indeed, the energy flux generated by cells in the absence of oxygen tends to be smaller than that generated when oxygen is abundant. So, in anaerobic environments characterized by smaller values of J_e , maintaining the balance between these two fluxes requires a smaller J_c leading to a higher fraction of carbon excreted as metabolic byproducts.

4 Methods

4.1 The coarse-grained model of *E. coli*

A comprehensive dynamical model of *E. coli* metabolism includes thousands of kinetic parameters describing ~ 744 chemical reactions [26], ~ 607 enzymes that catalyze reactions [26], ~ 314 regulators [27], and more than 2600 metabolites [28]. The task of fitting all of these parameters to the existing experimental data, such as e.g. dynamic ^{13}C -MFA is currently impossible. Therefore, some coarse-graining of metabolic pathways, proteome, precursors, and regulators needs to be performed. The summary of the coarse-graining procedure used in our study is outlined below:

- **(1) Coarse-graining of core metabolic pathways:** Multiple chemical reactions within the same linear chain or module are coarse-grained into a single step. Our model uses two energy- and precursor-related pathways: the glycolysis (flux J_g) and the TCA cycle (flux J_{tca}). Because the focus of our model is on cross-feeding of the acetate and the thermodynamic nature of the overflow metabolism, we extend our coarse-grained description to include two more pathways: the acetate overflow metabolism (such as the Pta-AckA pathway with flux J_o) and the acetate high-affinity uptake pathway (such as the ACS pathway with flux J_u). The exponential growth rate g is assumed to be limited by either the carbon assimilation flux J_c or the energy flux J_e . Liebig's law of the minimum [29] describes the resulting growth rate: $g = \min(g_c, g_e)$ where g_c is the growth rate when it is limited by the carbon flux J_c , and g_e - when it is limited by the energy flux $J_e = n_g J_g + n_{tca} J_{tca} + n_o (J_o - J_u)$. Here n_g , n_{tca} , and n_o are the numbers of ATP molecules released, by the glycolysis, the TCA cycle and the Pta-AckA pathways correspondingly.
- **(2) Coarse-grained proteome allocation:** We use the simplest partition of the proteome into the ribosomal fraction, R , with the remaining fraction E occupied by global metabolic enzymes. The ribosomal fraction R is used as a multiplier of the carbon precursor assimilation flux J_c . The global metabolic enzyme fraction E proportionally influences the glycolysis flux J_g and the TCA cycle flux J_{tca} . We assume that the flux through the ACS

pathway is not affected by E because past experiments reported that the enzyme for the ACS is constitutively expressed [6].

- (3) **Coarse-grained precursor dynamics:** Typically, the number of biomass precursors is larger than 10 [30, 31]. Since pyruvate is an important metabolic intermediate for almost all biomass precursors [32], one single universal precursor P is assumed.
- (4) **Global regulation:** Since many metabolic pathways are regulated by global metabolic regulators such as Cra and Crp [13], a single global regulator X is assumed to inhibit the production of ribosomes. The production of the global regulator X is assumed to be inhibited as the accumulated concentration of the biomass precursor P . Thus, the production term of X is given by $k_x \frac{K_{mp}}{K_{mp} + [P]}$, where k_x is the maximal production rate, K_{mp} is the inhibition constant from the precursor, and $[P]$ is the concentration of the precursor. Also, the degradation of the regulator is assumed to happen at a constant rate $d_x X$.

Hence, catabolic pathway fluxes in each of two strains are given by $J_{g,i} = k_{g,i}[E_i] \frac{[G]}{K_{mg,i} + [G]} \frac{K_{ip,i}}{K_{ip,i} + [P_i]}$,

$$J_{o,i} = \frac{v_{a,i}[P_i](1 - \frac{1}{K_{ca,i}} \frac{[A]}{[P_i]})}{K_{mp,i} + [P_i](1 + k_{ra,i} \frac{1}{K_{ca,i}} \frac{[A]}{[P_i]})}, \quad J_{tca,i} = k_{tca,i}[E_i] \frac{[P_i]}{K_{tca,i} + [P_i]}, \text{ and } J_{u,i} = \frac{k_{u,i}[A]}{K_{mu,i} + [A]},$$

where i is the strain label, $[G]$ is the glucose concentration, $[A]$ is the acetate concentration, and the concentration of metabolic enzymes for the strain i is $[E_i] = \frac{\beta - m_r[R_i]}{m_e}$ (conservation of the proteome). The anabolic pathway determines the carbon-limited growth rate $g_{c,i} = \frac{k_{c,i}[R_i][P_i]}{\beta(K_{mp,i} + [P_i])}$, energy-limited growth rate $g_{e,i} = n_g J_{g,i} + n_{tca} J_{tca,i} + n_o (J_{o,i} - J_{a,i})$, growth rate $g_i = \min(g_{c,i}, g_{e,i})$ and biomass flux $J_{b,i} = g_i \beta$. Descriptions and values of all unspecified kinetic parameters can be found in the supplemental information.

Experimental studies of the thermodynamically-controlled overflow pathway remain very limited. Thus, all kinetic parameters related to this pathway (such as its maximum rate, Monod constant, and the equilibrium constant) are not well-known. To calibrate these kinetic parameters, we used the experimental data about how the growth rate and metabolic fluxes varies as a function of the environmental concentration of the acetate from the dynamic ^{13}C -MFA data [7].

4.2 Cross-feeding community dynamics in a chemostat

Several evolutionary experiments starting with a single strain of *E. coli* growing in a glucose-limited environment resulted in a reproducible emergence of a two-strain stable microbial community consisting of a wild type strain, specializing in a primary resource (glucose), while the evolved strain co-utilizing glucose along with the overflow metabolic product (acetate) [5, 6]. The growth medium used in these evolutionary experiments is Davis minimal medium, quite different from the M9 minimal medium used in the previous dynamic ^{13}C -MFA experiments. Therefore, the OD growth curve of the glucose specialist strain CV103 from Ref. is used to fit maximal rates for glycolysis and the TCA cycle. Furthermore, it has been shown that, for the glucose specialist CV103, the high-affinity acetate uptake pathway Acetyl-CoA Synthetase (ACS) is barely expressed. So the acetate uptake rate contributed by the ACS pathway $J_{a,i}$ is set as 0 by taking $k_{a,i} = 0$.

Similar to the calibration for the glucose specialist, similar calibration for the acetate specialist CV101 is needed to account for different growth media. The OD growth curve of CV101 is used

to fit maximal rates for glycolysis, ACS pathway, and TCA cycle. Based on the established coarse-grained *E. coli* model, the population dynamics for the chemostat is only needed to study the coexistence problem. The dynamics for the microbial community between the glucose specialist CV103 (strain 1) and the acetate specialist CV101 (strain 2) is:

$$\frac{d[P_1]}{dt} = 2J_{g,1} - J_{o,1} - J_{b,1} - J_{tca,1} - g_1[P_1], \quad (3)$$

$$\frac{d[R_1]}{dt} = g_1\left(\frac{K_{ix,1}}{K_{ix,1} + [X_1]} - [R_1]\right), \quad (4)$$

$$\frac{d[X_1]}{dt} = k_{x,1}\left(\frac{K_{mp,1}}{K_{mp,1} + [P_1]}\right) - d_{x,1}[X_1], \quad (5)$$

$$\frac{dN_1}{dt} = (g_1 - D)N, \quad (6)$$

$$\frac{d[P_2]}{dt} = 2J_{g,2} - J_{o,2} - J_{b,2} + J_{u,2} - J_{tca,2} - g_2[P_2], \quad (7)$$

$$\frac{d[R_2]}{dt} = g_2\left(\frac{K_{ix,2}}{K_{ix,2} + [X_2]} - [R_2]\right), \quad (8)$$

$$\frac{d[X_2]}{dt} = k_{x,2}\left(\frac{K_{mp,2}}{K_{mp,2} + [P_2]}\right) - d_{x,2}[X_2], \quad (9)$$

$$\frac{dN_2}{dt} = (g_2 - D)N, \quad (10)$$

$$\frac{d[G]}{dt} = D([G_0] - [G]) - J_{g,1}N_1V_r - J_{g,2}N_2V_r, \quad (11)$$

$$\frac{d[A]}{dt} = D([A_0] - [A]) + J_{o,1}N_1V_r + (J_{o,2} - J_{u,2})N_2V_r, \quad (12)$$

where N_1 is number of the glucose specialist, N_2 is number of the acetate specialist, D is the dilution rate of the chemostat, $[G_0]$ is the glucose pool concentration, $[A_0]$ is the acetate pool concentration, and V_r is the volume size for a single *E. coli* cell.

Conflicts of interest

The authors declare that there are no competing interests.

References

- [1] Wolfe, A. J. The acetate switch. *Microbiol. Mol. Biol. Rev.* **69**, 12–50 (2005).
- [2] De Mey, M., De Maeseneire, S., Soetaert, W. & Vandamme, E. Minimizing acetate formation in *e. coli* fermentations. *Journal of industrial microbiology & biotechnology* **34**, 689–700 (2007).
- [3] De Dekker, R. The crabtree effect: a regulatory system in yeast. *Microbiology* **44**, 149–156 (1966).
- [4] Basan, M. *et al.* Overflow metabolism in escherichia coli results from efficient proteome allocation. *Nature* **528**, 99 (2015).

- [5] Helling, R. B., Vargas, C. N. & Adams, J. Evolution of escherichia coli during growth in a constant environment. *Genetics* **116**, 349–358 (1987).
- [6] Rosenzweig, R. F., Sharp, R., Treves, D. S. & Adams, J. Microbial evolution in a simple unstructured environment: genetic differentiation in escherichia coli. *Genetics* **137**, 903–917 (1994).
- [7] Enjalbert, B., Millard, P., Dinclaux, M., Portais, J.-C. & Létisse, F. Acetate fluxes in escherichia coli are determined by the thermodynamic control of the pta-acka pathway. *Scientific reports* **7**, 42135 (2017).
- [8] Richter, H. *et al.* Ethanol production in syngas-fermenting clostridium ljungdahlii is controlled by thermodynamics rather than by enzyme expression. *Energy & Environmental Science* **9**, 2392–2399 (2016).
- [9] Pfeiffer, T. & Morley, A. An evolutionary perspective on the crabtree effect. *Frontiers in molecular biosciences* **1**, 17 (2014).
- [10] Gudelj, I., Kinnersley, M., Rashkov, P., Schmidt, K. & Rosenzweig, F. Stability of cross-feeding polymorphisms in microbial communities. *PLoS computational biology* **12**, e1005269 (2016).
- [11] San Roman, M. & Wagner, A. An enormous potential for niche construction through bacterial cross-feeding in a homogeneous environment. *PLoS computational biology* **14**, e1006340 (2018).
- [12] Rozen, D. E. & Lenski, R. E. Long-term experimental evolution in escherichia coli. viii. dynamics of a balanced polymorphism. *The American Naturalist* **155**, 24–35 (2000).
- [13] Nanchen, A., Schicker, A., Revelles, O. & Sauer, U. Cyclic amp-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in escherichia coli. *Journal of bacteriology* **190**, 2323–2330 (2008).
- [14] Hoh, C.-Y. & Cord-Ruwisch, R. A practical kinetic model that considers endproduct inhibition in anaerobic digestion processes by including the equilibrium constant. *Biotechnology and bioengineering* **51**, 597–604 (1996).
- [15] Großkopf, T. & Soyer, O. S. Microbial diversity arising from thermodynamic constraints. *The ISME journal* **10**, 2725–2733 (2016).
- [16] Goyal, A. & Maslov, S. Diversity, stability, and reproducibility in stochastically assembled microbial ecosystems. *Physical review letters* **120**, 158102 (2018).
- [17] Marsland III, R. *et al.* Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities. *PLoS computational biology* **15**, e1006793 (2019).
- [18] Wang, T., Goyal, A., Dubinkina, V. & Maslov, S. Evidence for a multi-level trophic organization of the human gut microbiome. *PLOS Computational Biology* **15**, 1–20 (2019). URL <https://doi.org/10.1371/journal.pcbi.1007524>.
- [19] Goyal, A., Wang, T., Dubinkina, V. & Maslov, S. Ecology-guided prediction of cross-feeding interactions in the human gut microbiome. *Nature communications* **12**, 1–10 (2021).

- [20] Mac Arthur, R. Species packing, and what competition minimizes. *Proceedings of the National Academy of Sciences* **64**, 1369–1371 (1969).
- [21] MacArthur, R. Species packing and competitive equilibrium for many species. *Theoretical population biology* **1**, 1–11 (1970).
- [22] Jahan, N., Maeda, K., Matsuoka, Y., Sugimoto, Y. & Kurata, H. Development of an accurate kinetic model for the central carbon metabolism of escherichia coli. *Microbial cell factories* **15**, 112 (2016).
- [23] Klitgord, N. & Segrè, D. Environments that induce synthetic microbial ecosystems. *PLoS computational biology* **6**, e1001002 (2010).
- [24] Harcombe, W. R. *et al.* Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell reports* **7**, 1104–1115 (2014).
- [25] Chen, Y. & Nielsen, J. Energy metabolism controls phenotypes by protein efficiency and allocation. *Proceedings of the National Academy of Sciences* **116**, 17592–17597 (2019).
- [26] Ouzounis, C. A. & Karp, P. D. Global properties of the metabolic map of escherichia coli. *Genome research* **10**, 568–576 (2000).
- [27] Pérez-Rueda, E. & Collado-Vides, J. The repertoire of dna-binding transcriptional regulators in escherichia coli k-12. *Nucleic acids research* **28**, 1838–1847 (2000).
- [28] Guo, A. C. *et al.* Ecmbdb: the e. coli metabolome database. *Nucleic acids research* **41**, D625–D630 (2012).
- [29] Tilman, D. *Resource competition and community structure* (Princeton university press, 1982).
- [30] Savinell, J. M. & Palsson, B. O. Network analysis of intermediary metabolism using linear optimization. i. development of mathematical formalism. *Journal of theoretical biology* **154**, 421–454 (1992).
- [31] Feist, A. M. & Palsson, B. O. The biomass objective function. *Current opinion in microbiology* **13**, 344–349 (2010).
- [32] Noor, E., Eden, E., Milo, R. & Alon, U. Central carbon metabolism as a minimal biochemical walk between precursors for biomass and energy. *Molecular cell* **39**, 809–820 (2010).
- [33] Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
- [34] Hui, S. *et al.* Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Syst. Biol.* **11**, 784 (2015).
- [35] Kubitschek, H. E., Baldwin, W. W., Schroeter, S. J. & Graetzer, R. Independence of buoyant cell density and growth rate in *Escherichia coli*. *J. Bacteriol.* **158**, 296–299 (1984).
- [36] Basan, M. *et al.* Inflating bacterial cells by increased protein synthesis. *Mol. Syst. Biol.* **11**, 836 (2015).

- [37] Bremer, H. & Dennis, P. P. Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal Plus* **3** (2008).
- [38] Aidelberg, G. *et al.* Hierarchy of non-glucose sugars in *Escherichia coli*. *BMC Syst. Biol.* **8**, 133 (2014).
- [39] Gourse, R. L., Gaal, T., Bartlett, M. S., Appleman, J. A. & Ross, W. rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu. Rev. Microbiol.* **50**, 645–677 (1996).
- [40] Lemke, J. J. *et al.* Direct regulation of *Escherichia coli* ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc. Natl. Acad. Sci. USA* **108**, 5712–5717 (2011).
- [41] Piir, K., Paier, A., Liiv, A., Tenson, T. & Maiväli, Ü. Ribosome degradation in growing bacteria. *EMBO Rep.* **12**, 458–462 (2011).
- [42] Deutscher, M. P. Degradation of stable RNA in bacteria. *J. Biol. Chem.* **278**, 45041–45044 (2003).
- [43] Traxler, M. F. *et al.* The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol. Microbiol.* **68**, 1128–1148 (2008).
- [44] Klumpp, S., Scott, M., Pedersen, S. & Hwa, T. Molecular crowding limits translation and cell growth. *Proc. Natl. Acad. Sci. USA* **110**, 16754–16759 (2013).
- [45] Maitra, A. & Dill, K. A. Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proc. Natl. Acad. Sci. USA* **112**, 406–411 (2015).
- [46] Marr, A. G. Growth rate of *Escherichia coli*. *Microbiol. Rev.* **55**, 316–333 (1991).

Supplemental materials

Tong Wang, Chen Liao, Sergei Maslov

September 18, 2021

1 Simple cell growth model

1.1 Derivation of a Minimal Cell Growth Model for *E. coli*

Our kinetic bacterial growth model defines the interplay between the concentrations of precursor $[P]$, ribosomes $[R]$, the global regulator $[X]$, and the specific growth rate of the cell Λ . The dynamics of $[R]$, $[P]$, and $[X]$ is modeled by considering their rates of production minus their rates of consumption and degradation. Namely,

$$\frac{d[P]}{dt} = J_p^{in} - J_p^{out}, \quad (1)$$

$$\frac{d[R]}{dt} = J_r^{in} - J_r^{out}, \quad (2)$$

$$\frac{d[X]}{dt} = J_x^{in} - J_x^{out}, \quad (3)$$

where J_x^{in} is the production rate of molecule x and J_x^{out} is its combined consumption rate and dilution due to the growth of cell volume. The specific growth rate is determined with respect to cell volume Ω , which grows exponentially as

$$\Lambda = \frac{1}{\Omega} \frac{d\Omega}{dt}. \quad (4)$$

The mathematical representations of these reaction rates are detailed as follows.

1.1.1 A coarse-grained model of *E. coli* proteome

Recent experimental studies [33, 34] have shown that the proteome of exponentially growing *E. coli* cells can be minimally partitioned into two sectors which each have a distinct role in determining its physiology. Ribosomal and affiliated proteins belong to the R sector, while catabolic and anabolic enzymes for nutrient uptake and conversion belong to the E sector. Since the *E. coli* mass density is relatively constant in the presence of various perturbations [35, 36] and the majority of biomass is protein [37], the protein density $[O]$ remains unchanged under different growth conditions. Thus the allocation of protein density $[O]$ to R sector and E sector can be written as:

$$[O] = m_r[R] + m_e[E] = \beta, \quad (5)$$

where m_r and m_e is the average number of amino acids in a ribosome and a metabolic protein respectively.

1.1.2 The precursor production, consumption and dilution rates

Under constant or saturating nutrient concentration, the nutrient uptake rate is determined by the availability of metabolic proteins and subject to feedback inhibition by its end products, precursor. That is, when the metabolic intermediates between nutrient and precursor are ignored, the production rate of precursor is

$$J_p^{in} = f_e([E]) \frac{[P]}{K_{ip} + [P]}, \quad (6)$$

where the function $f_e([E])$ relates the influx of precursor to the concentration of coarse-grained metabolic proteins. For most rate-limiting enzymes that absorb nutrients from environment, the function is linear [38], so we have

$$f_e([E]) = k_e[E], \quad (7)$$

where k_e is a linear coefficient that represents the nutrient quality. The other term $\frac{[P]}{K_{ip} + [P]}$ is used to describe the feedback inhibition from the precursor where K_{ip} is the inhibition constant.

In this framework, precursors are utilized solely for protein synthesis and not for other metabolic reactions. Since each active ribosome is able to initiate translation, the consumption rate of precursor (J_p^{con}) in protein synthesis is modelled using a Michaelis-Menten equation

$$J_p^{con} = \frac{k_r[R][P]}{K_{mp} + [P]}. \quad (8)$$

We assume precursors are only diluted via volume expansion, so the dilution rate of precursor is

$$J_p^{dil} = \Lambda[P]. \quad (9)$$

The total precursor outflux rate is the sum of the reaction rates from both utilization via protein synthesis and dilution via volume expansion, that is,

$$J_p^{out} = J_p^{con} + J_p^{dil}. \quad (10)$$

1.1.3 The ribosome synthesis and dilution rates

It is generally believed that the synthesis of ribosomal RNAs is the limiting step in ribosome production due to excessive ribosomal proteins than ribosomal RNAs [39]. On the other hand, ribosome production is also limited by the synthesis of ribosomal proteins in terms of reaction dynamics because protein synthesis takes a much longer time as compared to RNA synthesis. Considering the dynamic nature of our minimal model, we assume the latter, meaning that the ribosome production rate J_r^{in} is determined by the fraction of total translational capacity allocated for producing R sector proteins. Additionally, because ppGpp inhibits transcription initiation of ribosomal proteins both *in vitro* and *in vivo* [40], a function $\alpha_r([X])$ that describes the degree of inhibition is included. The ribosome production rate is then

$$J_r^{in} = \frac{J_p^{con}}{m_r} \alpha_r([X]). \quad (11)$$

The regulation of ribosome by the global regulatory X is written as

$$\alpha_r = \frac{K_{ix}}{[X] + K_{ix}}, \quad (12)$$

where K_{ix} is the regulator concentration at which half of the ribosome synthesis is halted.

Ribosomes are very stable in fast growth but are degraded in starvation, entry into stationary phase, and slow growth [41, 42]. For simplicity, we assume ribosomes are not degraded. The only reduction of ribosome concentration comes from the dilution effect, thus we have the out flow of the ribosome as

$$J_r^{out} = \Lambda[R]. \quad (13)$$

1.1.4 The synthesis and degradation rates of global regulator

When *E. coli* is under amino acid starvation, ppGpp regulator is activated [43]. When the concentration of ppGpp regulator is increased, the synthesis of ribosome protein is inhibited [40] and *E. coli* stops growth. As a result, we assume the synthesis rate of global regulator X is inversely regulated by the precursor concentration

$$J_x^{in} = k_x \frac{K_{mp}}{K_{mp} + [P]}, \quad (14)$$

where k_x is the maximal synthesis rate and K_{mp} is the dissociation constant of precursor.

The turnover rate of global regulator is generally much faster than the reduction rate due to dilution. Therefore, we assume first-order active degradation and ignore the dilution effect

$$J_x^{out} = d_x[X], \quad (15)$$

where d_x is the first order degradation rate constant.

1.1.5 The specific growth rate

Using the previous assumption that the intracellular protein density is constant, the specific growth rate Λ is determined by the rate of protein production

$$\Lambda = \frac{1}{\Omega} \frac{d\Omega}{dt} = \frac{1}{\Omega} \frac{d}{dt} \left(\frac{O}{\beta} \right) = \frac{1}{\beta} \left(\frac{1}{\Omega} \frac{dO}{dt} \right) = \frac{J_p^{con}}{\beta}. \quad (16)$$

1.1.6 Summary

The entire model equations are summarized below

$$\frac{d[P]}{dt} = \underbrace{k_e[E]}_{\substack{\text{maximum precursor} \\ \text{synthesis rate}}} \underbrace{\frac{K_{ip}}{K_{ip} + [P]}}_{\substack{\text{feedback} \\ \text{inhibition}}} - \underbrace{\frac{k_r[R][P]}{K_{mp} + [P]}}_{\substack{\text{precursor} \\ \text{consumption rate}}} - \underbrace{\Lambda[P]}_{\substack{\text{precursor} \\ \text{dilution rate}}} \quad (17)$$

$$\frac{d[R]}{dt} = \underbrace{\frac{k_r[R][P]}{m_r(K_{mp} + [P])}}_{\substack{\text{ribosome} \\ \text{synthesis rate}}} \underbrace{\frac{K_{ix}}{K_{ix} + [X]}}_{\substack{\text{ribosome} \\ \text{dilution rate}}} - \underbrace{\Lambda[R]}_{\substack{\text{ribosome} \\ \text{dilution rate}}} \quad (18)$$

$$\frac{d[X]}{dt} = \underbrace{k_x \frac{K_{mp}}{K_{mp} + [P]}}_{\substack{\text{regulator} \\ \text{synthesis rate}}} - \underbrace{d_x[X]}_{\substack{\text{regulator} \\ \text{degradation rate}}} \quad (19)$$

$$\Lambda = \underbrace{\frac{k_r[R][P]}{\beta(K_{mp} + [P])}}_{\substack{\text{specific growth rate}}} \quad (20)$$

$$[E] = \underbrace{\frac{\beta - m_r[R]}{m_e}}_{\text{metabolic protein concentration}} \quad (21)$$

1.2 Nondimensionalization

The following transformations are made:

$p = [P]/\beta$, $r = m_r[R]/\beta$, $x = [X]/\beta$, $\tau = k_r t/m_r$, $\lambda = m_r \Lambda/k_r$, $k'_e = m_r k_e/(k_r m_e)$, $k'_x = (k_x m_r)/(k_r \beta)$, $d'_x = d_x m_r/k_r$, $K'_{mp} = K_{mp}/\beta$, $K'_{ix} = K_{ix}/\beta$, $K'_{ip} = K_{ip}/\beta$.

$$\frac{dp}{d\tau} = k'_e(1-r) \frac{K'_{ip}}{K'_{ip} + p} - \lambda(1+p) \quad (22)$$

$$\frac{dr}{d\tau} = \lambda \left(\frac{K'_{ix}}{K'_{ix} + x} - r \right) \quad (23)$$

$$\frac{dx}{d\tau} = k'_x \frac{K'_{mp}}{K'_{mp} + p} - d'_x x \quad (24)$$

$$\lambda = \frac{rp}{K'_{mp} + p} \quad (25)$$

1.3 Parameters

| Symbol | Description | Value | Unit | Formula | Source |
|----------|--|--------------------|---------|--|--------|
| m_r | # of precursor in a ribosome | 11738 | | 7336×1.6 | [44] |
| m_e | # of precursor in a metabolic protein | 325 | | | [45] |
| k_r | maximum rate of peptide elongation | 7.56×10^4 | 1/hr | $21 \frac{1}{s} \times 3600 \frac{s}{hr}$ | [46] |
| k_x | maximum rate of ppGpp synthesis | 3.6×10^3 | 1/hr | $1 \frac{1}{s} \times 3600 \frac{s}{hr}$ | [46] |
| d_x | first-order rate constant of ppGpp degradation | 1.26×10^2 | 1/hr | $0.035 \frac{1}{s} \times 3600 \frac{s}{hr}$ | [46] |
| K_{mp} | dissociation constant for amino acid | 2.00×10^1 | μM | | [46] |
| K_{ix} | dissociation constant for ppGpp | 6.00×10^1 | μM | | [46] |
| β | total peptide concentration | 3.00×10^6 | μM | | [46] |
| K_{ip} | precursor feedback inhibition constant | 1.00×10^4 | μM | | [45] |

Table 2: Model parameters used in the model.

1.4 Steady state solution

Let $p \ll K'_{ip}$ so that $K'_{ip}/(K'_{ip} + p) \approx 1$. Combining Eq. 22 and Eq. 25 gives

$$\lambda = \frac{k'_e}{1 + p + k'_e \frac{K'_{mp} + p}{p}}. \quad (26)$$

Similarly, combining Eq. 23 and Eq. 24 gives

$$\lambda = \frac{K'_{ix}p}{K'_{ix}(K'_{mp} + p) + \frac{k'_x K'_{mp}}{d'_x}}. \quad (27)$$

Combining the two equations above gives $p^2 + p - \frac{k'_e k'_x K'_{mp}}{K'_{ix} d'_x} = 0$, whose solution is given by

$$p = \frac{-1 + \sqrt{1 + \frac{4k'_e k'_x K'_{mp}}{K'_{ix} d'_x}}}{2}. \quad (28)$$

Accordingly, the growth rate λ is rewritten as

$$\lambda = \frac{k'_e}{k'_e + \frac{1}{2} \left(1 + \sqrt{1 + \frac{4k'_e k'_x K'_{mp}}{K'_{ix} d'_x}} \right) \left(1 + \frac{K'_{ix} d'_x}{k'_x} \right)}. \quad (29)$$

2 Cell growth model with glycolysis, fermentation, TCA cycle, and energy limitation

2.1 Derivation of a cell growth model for *E.coli* with glycolysis, fermentation, TCA cycle and energy limitation

In this section, the previous minimal single-cell model is extended to include fermentation and TCA cycle. To do this, two carbon sources are considered: a primary carbon source and a fermentation product which is a byproduct of the primary carbon source. And pathways involving glycolysis, fermentation, and TCA cycle are also properly modeled. In the previous minimal model, the growth rate of the single cell is considered to be limited by precursor consumption rate J_p^{con} . However, in reality, the growth of cells can also be limited by the energy/ATP flux. Thus the contribution of the energy/ATP flux to the growth rate is also considered here.

2.1.1 New environmental variables included

Two environmental variables are included: extracellular concentrations for the primary carbon source and the fermentation product. For simplicity, the intracellular concentrations of those two carbon sources are ignored and only the extracellular concentrations are considered: $[G_{ex}]$ represents the concentration of the primary carbon source outside the cell and $[A_{ex}]$ represents the concentration of the fermentation/overflow product outside the cell.

2.1.2 Pathways involving glycolysis, fermentation and TCA cycles

The glycolysis pathway is modeled as the conversion of the primary carbon source $[S_{ex}]$ to the precursor $[P]$. And it is written as

$$J_g = \frac{k_e [E] [S_{ex}]}{K_{ms} + [S_{ex}]} \frac{K_{ip}}{K_{ip} + [P]}, \quad (30)$$

where the Michaelis-Menten kinetics and the feedback inhibition is included.

The fermentation pathway is modeled as the conversion of the precursor $[P]$ to the fermentation product $[A_{ex}]$. The fermentation pathway in many different microbial species is controlled by thermodynamics rather than by enzyme expression [7, 8, 9]. For instance, *E. coli* cells have a famous acetate switch where *E. coli* will ferment glucose into acetate and secrete acetate into the extracellular environment when the glucose is sufficient in the environment. The overflow metabolism of generating acetate is thermodynamically controlled by Pta-Acka pathway [7]. When the glucose is depleted, *E. coli* can

re-uptake the acetate previously excreted into the environment. This uptake of acetate can be done by reversing the direction of the Pta-Acka pathway from overflow to assimilation, but the affinity for acetate assimilation is typically low with a value around 7 to 10 mM [7]. There is another acetate assimilation pathway Acs with a higher affinity close to 200 μM [6].

Because the overflow pathway is thermodynamically controlled, it is written as

$$J_o = \frac{v_a[P](1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]})}{K_{mp} + [P](1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]})}, \quad (31)$$

where the formalism follows the previous derivation with thermodynamic constraints [14, 15]. In the formula, the most important parameter reflecting the thermodynamic control is K_{ca} which is the equilibrium constant. When $\frac{[A_{ex}]}{[P]} < K_{ca}$, J_o is positive which means that the cell is generating the fermentation product. While if $\frac{[A_{ex}]}{[P]} > K_{ca}$, J_o is negative which means that the cell is using the fermentation product. The explanation of all other parameters is in the Supplemental Table 2.

The high-affinity assimilation of the fermentation product and the flux of the uptake is written as

$$J_u = \frac{k_u[E][A_{ex}]}{K_{mu} + [A_{ex}]}, \quad (32)$$

where the flux is regulated by the enzyme $[E]$.

The TCA cycle is modeled as the oxidation of precursor to CO_2 . Because TCA cycle is also regulated by the enzyme $[E]$, the flux of the TCA cycle is expressed as

$$J_{tca} = k_{tca}[E] \frac{[P]}{K_{tca} + [P]}. \quad (33)$$

2.1.3 Growth rate considering carbon/precursor limitation and energy limitation

Previously, in the minimal model of a single cell, the growth rate is considered as carbon-limited (or precursor-limited) where the growth rate is $\frac{k_r[R][P]}{\beta(K_{mp} + [P])}$. However, in reality, the biomass synthesis requires good coordination between the precursor and ATP. To include both factors, we consider a carbon-limited growth rate Λ_c and an energy/ATP-limited growth rate Λ_e .

When energy is sufficient, the growth rate is only limited by the concentration of precursor, so the carbon-limited growth rate

$$\Lambda_c = \frac{k_r[R][P]}{\beta(K_{mp} + [P])} \quad (34)$$

which is the same as the growth rate used in the previous section.

When the precursor is sufficient, the growth rate is only limited by the ATP flux. The ATP flux comes from various pathways generating ATP: glycolysis, overflow pathway and the TCA cycle. By summing up all ATP fluxes, we get the total ATP flux produced $J_{ATP} = n_g J_g + n_{tca} J_{tca} + n_o (J_o - J_u)$ where n_g , n_{tca} and n_o represents the ATP produced per flux for glycolysis, TCA cycle and overflow pathways respectively. Then the energy-limited growth rate Λ_e is assumed to be proportional to the total ATP flux

$$\Lambda_e = \sigma J_{ATP} = \sigma \{n_g J_g + n_{tca} J_{tca} + n_o (J_o - J_u)\}, \quad (35)$$

where σ is a parameter representing growth rate per ATP production flux.

For a general case, the growth rate should be the minimal of the Λ_c and Λ_e , so growth rate $\Lambda = \min\{\Lambda_c, \Lambda_e\}$. Then the precursor consumption rate is

$$J_p^{con} = \beta \Lambda = \beta \min\{\Lambda_c, \Lambda_e\}. \quad (36)$$

The ribosome production rate is determined by the growth rate and the regulator $[X]$ together:

$$J_r = \frac{\beta \Lambda}{m_r} \frac{K_{ix}}{K_{ix} + [X]}. \quad (37)$$

2.1.4 Summary

The entire model equations are summarized below:

$$\begin{aligned}
\frac{d[P]}{dt} &= 2J_g - J_o + J_u - J_b - J_{tca} - \Lambda[P] \\
&= \underbrace{\frac{2k_e[E][S_{ex}]}{K_{ms} + [S_{ex}]}}_{\text{maximum precursor synthesis rate}} \underbrace{\frac{K_{ip}}{K_{ip} + [P]}}_{\text{feedback inhibition}} - \underbrace{\frac{v_a[P] \left(1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]}\right)}{K_{ma} + [P] \left(1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]}\right)}}_{\substack{\text{reversible fermentation rate} \\ \text{with thermodynamic constraint}}} + \underbrace{\frac{k_u[E][A_{ex}]}{K_{mu} + [A_{ex}]}}_{\substack{\text{fermentation product uptake rate} \\ \text{precursor consumption rate}}} - \underbrace{\frac{\beta\Lambda}{\Lambda[P]}}_{\text{precursor dilution rate}} \\
&\quad - \underbrace{\frac{k_{tca}[E][P]}{K_{tca} + [P]}}_{\text{rate of TCA cycle}} - \underbrace{\frac{\Lambda[P]}{\Lambda[P]}}_{\text{precursor dilution rate}} \\
\frac{d[R]}{dt} &= J_r - \Lambda[R] \\
&= \underbrace{\frac{\beta\Lambda}{m_r K_{ix} + [X]}}_{\substack{\text{ribosome synthesis rate} \\ \text{ribosome dilution rate}}} - \underbrace{\frac{\Lambda[R]}{\Lambda[R]}}_{\text{ribosome dilution rate}} \\
\frac{d[X]}{dt} &= k_x \underbrace{\frac{K_{mp}}{K_{mp} + [P]}}_{\substack{\text{regulator synthesis rate} \\ \text{regulator degradation rate}}} - \underbrace{\frac{d_x[X]}{d_x[X]}}_{\text{regulator degradation rate}} \\
\Lambda_c &= \underbrace{\frac{k_r[R][P]}{\beta(K_{mp} + [P])}}_{\text{carbon limited growth rate}} \\
\Lambda_e &= \sigma J_{ATP} = \sigma \{n_g J_g + n_{tca} J_{tca} + n_o (J_o - J_u)\} \\
&= \sigma \underbrace{\left(n_g \frac{k_e[E][S_{ex}]}{K_{ms} + [S_{ex}]} \frac{K_{ip}}{K_{ip} + [P]} + n_{tca} \frac{k_{tca}[E][P]}{K_{tca} + [P]} + n_o \left(\frac{v_a[P] \left(1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]}\right)}{K_{ma} + [P] \left(1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]}\right)} - \frac{k_u[E][A_{ex}]}{K_{mu} + [A_{ex}]} \right) \right)}_{\text{energy limited growth rate}} \\
\Lambda &= \min\{\Lambda_b, \Lambda_e\} \\
[E] &= \underbrace{\frac{\beta - m_r[R]}{m_e}}_{\substack{\text{metabolic protein concentration} \\ \text{concentration}}}
\end{aligned}$$

Note that there is a factor of 2 in front of glycolysis flux J_g to the production of precursor because there is a stoichiometry of 1:2 from the primary carbon source to the precursor. For example, the stoichiometry of glycolysis pathway in the *E. coli* will convert 1 molecule of glucose into 2 molecules of pyruvate.

2.2 Nondimensionalization

The transformations are as follows: $p = [P]/\beta$, $r = m_r[R]/\beta$, $x = [X]/\beta$, $s_{ex} = [S_{ex}]/\beta$, $a_{ex} = [A_{ex}]/\beta$, $\tau = k_r t / m_r$, $\lambda = m_r \Lambda / k_r$, $\lambda_c = m_r \Lambda_c / k_r$, $\lambda_e = m_r \Lambda_e / k_r$, $J'_g = m_r J_g / (k_r \beta)$, $J'_o = m_r J_o / (k_r \beta)$, $J'_u = m_r J_u / (k_r \beta)$, $J'_b = m_r J_b / (k_r \beta)$, $J'_{tca} = m_r J_{tca} / (k_r \beta)$, $J'_r = m_r J_r / (k_r \beta)$, $k'_e = m_r k_e / (k_r m_e)$, $K'_{ms} = K_{ms} / \beta$, $K'_{ip} = K_{ip} / \beta$, $v'_a = m_r v_a / (k_r \beta)$, $K'_{ms} = K_{ms} / \beta$, $K'_{ma} = K_{ma} / \beta$, $k'_u = m_r k_u / (k_r m_e)$, $K'_{mu} = K_{mu} / \beta$, $k'_{tca} = m_r k_{tca} / (k_r m_e)$, $K'_{tca} = K_{tca} / \beta$, $K'_{ix} = K_{ix} / \beta$, $k'_x = m_r k_x / (k_r \beta)$, $K'_{mp} = K_{mp} / \beta$, $d'_x = m_r d_x / k_r$, $n'_g = \sigma \beta n_g$, $n'_{tca} = \sigma \beta n_{tca}$, $n'_o = \sigma \beta n_o$.

$$\frac{dp}{d\tau} = 2J'_g - J'_o + J'_u - J'_b - J'_{tca} - \lambda p \quad (38)$$

$$= \frac{2k'_e(1-r)s_{ex}}{K'_{ms} + s_{ex}} \frac{K'_{ip}}{K'_{ip} + p} - \frac{v'_a p \left(1 - \frac{1}{K_{ca}} \frac{a_{ex}}{p}\right)}{K'_{ma} + p \left(1 + \frac{k_{ra}}{K_{ca}} \frac{a_{ex}}{p}\right)} + \frac{k'_u(1-r)a_{ex}}{K'_{mu} + a_{ex}} \quad (39)$$

$$- \lambda - \frac{k'_{tca}(1-r)p}{K'_{tca} + p} - \lambda p \quad (40)$$

$$\frac{dr}{d\tau} = J'_r - \lambda r \quad (41)$$

$$= \lambda \frac{K'_{ix}}{K'_{ix} + x} - \lambda r \quad (42)$$

$$\frac{dx}{d\tau} = k'_x \frac{K'_{mp}}{K'_{mp} + p} - d'_x x \quad (43)$$

$$\lambda_b = \frac{rp}{K'_{mp} + p} \quad (44)$$

$$\lambda_e = n'_g J'_g + n'_{tca} J'_{tca} + n'_o (J'_o - J'_u) \quad (45)$$

$$= n'_g \frac{k'_e(1-r)s_{ex}}{K'_{ms} + s_{ex}} \frac{K'_{ip}}{K'_{ip} + p} + n'_{tca} \frac{k'_{tca}(1-r)p}{K'_{tca} + p} \quad (46)$$

$$+ n'_o \left(\frac{v'_a p \left(1 - \frac{1}{K_{ca}} \frac{A_{ex}}{p}\right)}{K'_{ma} + p \left(1 + \frac{k_{ra}}{K_{ca}} \frac{A_{ex}}{p}\right)} - \frac{k'_u(1-r)a_{ex}}{K'_{mu} + a_{ex}} \right) \quad (47)$$

$$\lambda = \min\{\lambda_b, \lambda_e\} \quad (48)$$

2.3 Parameters

| Symbol | Description | Value | Unit | Source |
|-----------|--|--------------------|-----------|--------------------|
| k_e | maximum rate of glucose uptake | 2.07×10^3 | h^{-1} | estimated from [7] |
| K_{ms} | Michaelis constant for glucose uptake | 6.67 | μM | [22] |
| K_{ip} | amino acids/aa. (precursor) feedback inhibition constant | 1.00×10^4 | μM | [45] |
| v_a | maximum rate of acetate (fermentation product) production | 1.62×10^8 | $\mu M/h$ | [22] |
| K_{ma} | Michaelis constant for acetate (fermentation product) production | 1.60×10^2 | μM | [22] |
| K_{ca} | equilibrium constant | 174.2 | | [22] |
| k_{ra} | ratio between maximal forward and reverse rate | 1.00 | | assumed |
| k_u | maximum rate of acetate (fermentation production) reuptake | 1.13×10^3 | h^{-1} | estimated from [7] |
| K_{mu} | Michaelis constant for acetate (fermentation product) reuptake | 200 | μM | [22] |
| k_{tca} | maximum rate of the TCA cycle | 1.13×10^3 | h^{-1} | estimated from [7] |
| K_{tca} | Michaelis constant for the TCA cycle | 200 | μM | estimated from [7] |

| | | | | |
|----------|--|-----------------------------|--------------|------------|
| m_r | # of aa. (precursor) in a ribosome | 11738 (7336×1.6) | | [45] |
| m_e | # of aa. (precursor) in a metabolic protein | 325 | | [45] |
| k_r | maximum rate of peptide elongation | 7.56×10^4 | h^{-1} | [46] |
| k_x | maximum rate of ppGpp (regulator) synthesis | 3.6×10^3 | h^{-1} | [46] |
| d_x | first-order rate constant of ppGpp (regulator) degradation | 1.26×10^2 | h^{-1} | [46] |
| K_{mp} | dissociation constant for aa. (precursor) | 2.00×10^1 | μM | [46] |
| K_{ix} | dissociation constant for ppGpp (regulator) | 6.00×10^1 | μM | [45] |
| β | total aa. (precursor) concentration | 3.00×10^6 | μM | [46] |
| $P:O$ | ATP produced per NADH or FADH ₂ | 2 | | |
| n_p | ATP yield of glycolysis | 10 | | $2+4(P:O)$ |
| n_a | ATP yield of fermentation | 1 | | $1+0(P:O)$ |
| n_r | ATP yield of TCA | 9 | | $1+4(P:O)$ |
| σ | growth rate per ATP production flux | 8.87×10^{-9} | μM^{-1} | [22] |

Table 3: Model parameters used in the model.

2.4 Steady state under the energy-limited condition

Whether the growth of cells is energy-limited or carbon-limited is generally hard to conclude. Here we focus on the energy-limited growth condition because the single-cell model is energy-limited for most environmental conditions. The energy-limited case is very common because microbes often use a particular carbon source in the medium, while they produce another carbon source. Under this circumstance, not all carbon is fixed because the energy produced by microbes is the limiting factor. For example, *E. coli* under the aerobic environment with sufficient glucose will consume glucose and secrete acetate.

The case of energy-limited growth rate means that $\lambda = \lambda_e$. We can combine this relation together with the Eq. 45 to obtain

$$\lambda = n'_g J'_g + n'_{tca} J'_{tca} + n'_o (J'_o - J'_u). \quad (49)$$

Set the right hand side of the differential equation Eq. 38 be zero, we have

$$2J'_g - J'_o + J'_u - J'_b - J'_{tca} - \lambda p = 0. \quad (50)$$

Because the concentration of the precursor is usually small, so the term λp is typically much smaller than all other flux terms. Then we get of a condition reflecting the balance of fluxes:

$$2J'_g - (J'_o - J'_u) - J'_b - J'_{tca} = 0. \quad (51)$$

Because $J'_b = \lambda$, so we have

$$\lambda = 2J'_g - (J'_o - J'_u) - J'_{tca}. \quad (52)$$

Combining Eq. 49 and Eq. 52 will give the following equation:

$$\lambda = \frac{(n'_g + 2n'_{tca})}{(1 + n'_{tca})} J'_g + \frac{(n'_{tca} - n'_o)}{(1 + n'_{tca})} (J'_u - J'_o). \quad (53)$$

The above derived relation is similar to the form used by the MacArthur's Consumer-Resource Model [21] where the growth rate of a consumer is the summation of the contribution to growth rate from all resources.

If we use the parameters used for *E. coli* consuming glucose and acetate, we will have $\lambda = 2.8J'_g + 0.8(J'_u - J'_o)$. The coefficients in front of the glucose flux J'_g and the acetate uptake flux $(J'_u - J'_o)$ are 2.8 and 0.8. They are interesting because they reflect their yields to the growth rate.

3 Multiple strains community model with population dynamics

3.1 Derivation of multiple strains community model for *E.coli*

To bridge the gap between the single-cell model and the community model of multiple strains, the population dynamics for each strain and nutrient dynamics are needed to be specified. Here the case of the chemostat is considered, multiple strains of microbes are living in the chemostat where only the primary carbon source is provided into the system. The form of population-nutrient dynamics used here follows MacArthur's consumer-resource model [21].

3.1.1 Population dynamics

The growth of each strain follows the single-cell model defined in the second section. The number cell of the microbial strain i in the chemostat N_i is introduced. The dynamics of the strain i is simply expressed as

$$\frac{dN_i}{dt} = \underbrace{\Lambda_i N_i}_{\text{growth rate}} - \underbrace{\delta N_i}_{\text{dilution rate}},$$

where Λ_i represents the growth rate of the microbial strain i and δ represents the dilution rate of the chemostat.

3.1.2 Nutrient dynamics

In the chemostat, only the primary carbon source is provided. The consumption and the generation of carbon sources are needed in the dynamics of the primary carbon source $[S_{ex}]$ and the secondary fermentation product $[A_{ex}]$. The consumption of the primary carbon source $[S_{ex}]$ is written as:

$$J_{S_{ex}}^{con} = J_{g,i} N_i V_r,$$

where V_r is the volume ratio between the volume of a single cell and the volume of the whole chemostat. The net production of the secondary fermentation/overflow product $[A_{ex}]$ is written as:

$$J_{A_{ex}}^{prod} = (J_{o,i} - J_{u,i}) N_i V_r.$$

3.1.3 Summary

Two strains of bacteria are introduced into the system, labeled by i ($=1,2,\dots$) The model equations with fermentation and overflow are as follows:

$$\begin{aligned} \frac{dN_i}{dt} &= \underbrace{\Lambda_i N_i}_{\text{growth rate}} - \underbrace{\delta N_i}_{\text{dilution rate}} \\ \frac{d[S_{ex}]}{dt} &= \underbrace{\delta([S_0] - [S_{ex}])}_{\text{dilution of primary carbon source}} - \sum_i \underbrace{\frac{2k_{e,i}[E]_i[S_{ex}]}{K_{ms,i} + [S_{ex}]}}_{\text{maximum precursor synthesis rate}} \underbrace{\frac{K_{ip,i}}{K_{ip,i} + [P]_i}}_{\text{feedback inhibition}} N_i V_r \\ \frac{d[A_{ex}]}{dt} &= - \underbrace{\delta([A_{ex}])}_{\text{diffusion of fermentation product}} + \sum_i \left(\underbrace{\frac{v_{a,i}[P]_i(1 - \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i})}{K_{ma,i} + [P]_i(1 + k_{ra,i} \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i})}}_{\text{reversible fermentation rate with thermodynamic constraints}} - \underbrace{\frac{k_{u,i}[E]_i[A_{ex}]}{K_{mu,i} + [A_{ex}]}}_{\text{fermentation product uptake rate}} \right) N_i V_r \end{aligned}$$

$$\begin{aligned}
\frac{d[P]_i}{dt} &= 2J_{g,i} - J_{o,i} + J_{u,i} - J_{b,i} - J_{tca,i} - \Lambda_i [P]_i \\
&= \underbrace{\frac{2k_{e,i}[E]_i[S_{ex}]}{K_{ms,i} + [S_{ex}]}}_{\text{maximum precursor synthesis rate}} \underbrace{\frac{K_{ip,i}}{K_{ip,i} + [P]_i}}_{\text{feedback inhibition}} - \underbrace{\frac{v_{a,i}[P]_i \left(1 - \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i}\right)}{K_{ma,i} + [P]_i \left(1 + k_{ra,i} \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i}\right)}}_{\substack{\text{reversible fermentation rate} \\ \text{with thermodynamic constraint}}} + \underbrace{\frac{k_{u,i}[E]_i[A_{ex}]}{K_{mu,i} + [A_{ex}]}}_{\substack{\text{fermentation product uptake rate}}} \\
&\quad - \underbrace{\frac{\beta \Lambda_i}{K_{tca,i} + [P]_i}}_{\substack{\text{precursor consumption rate}}} - \underbrace{\frac{\Lambda_i [P]_i}{\text{rate of TCA cycle}}}_{\text{precursor dilution rate}} \\
\frac{d[R]_i}{dt} &= J_{r,i} - \Lambda_i [R]_i \\
&= \underbrace{\frac{\beta \Lambda_i}{m_r K_{ix,i} + [X]_i}}_{\substack{\text{ribosome synthesis rate}}} - \underbrace{\frac{\Lambda_i [R]_i}{\text{ribosome dilution rate}}}_{\text{ribosome dilution rate}} \\
\frac{d[X]_i}{dt} &= \underbrace{k_{x,i} \frac{K_{mp,i}}{K_{mp,i} + [P]_i}}_{\substack{\text{regulator synthesis rate}}} - \underbrace{d_{x,i} [X]_i}_{\substack{\text{regulator degradation rate}}} \\
\Lambda_{c,i} &= \underbrace{\frac{k_{r,i}[R]_i[P]_i}{\beta(K_{mp,i} + [P]_i)}}_{\substack{\text{carbon limited growth rate}}} \\
\Lambda_{e,i} &= \sigma J_{ATP,i} = \underbrace{\sigma \{n_g J_{g,i} + n_{tca} J_{tca,i} + n_o (J_{o,i} - J_{u,i})\}}_{\text{energy limited growth rate}} \\
&= \sigma n_g \frac{k_{e,i}[E]_i[S_{ex}]}{K_{ms,i} + [S_{ex}]} \frac{K_{ip,i}}{K_{ip,i} + [P]_i} + \sigma n_{tca} \frac{k_{tca,i}[E]_i[P]_i}{K_{tca,i} + [P]_i} \\
&\quad + \sigma n_o \left(\underbrace{\frac{v_{a,i}[P]_i \left(1 - \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i}\right)}{K_{ma,i} + [P]_i \left(1 + k_{ra,i} \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i}\right)}}_{\substack{\text{reversible fermentation rate} \\ \text{with thermodynamic constraint}}} - \underbrace{\frac{k_{u,i}[E]_i[A_{ex}]}{K_{mu,i} + [A_{ex}]}}_{\substack{\text{fermentation product uptake rate}}} \right) \\
\Lambda_i &= \min\{\Lambda_{b,i}, \Lambda_{e,i}\} \\
[E]_i &= \underbrace{\frac{\beta - m_r [R]_i}{m_e}}_{\substack{\text{metabolic protein concentration}}}
\end{aligned}$$

4 Application to the stable acetate cross-feeding polymorphism of the *E. coli*

4.1 Background

It has been reported many times that the long term evolution of the single-clone *E. coli* in an environment with glucose fed as the sole carbon source will become polymorphic [5, 6, 12]. The coexistence of several different clones with glucose as the only carbon source could be explained by differential patterns of the secretion and uptake of two alternative metabolites, acetate and glycerol. For example, it has been observed during the evolution, a stable coexistence between two strains CV101 and CV103 can be established. CV103 is a glucose specialist, and it consumes glucose and secretes acetate into the environment. While

the other strain CV101 is an acetate specialist where it can consume the acetate secreted. Thus the coexistence between two strains can be established because of the cross-feeding via acetate.

The physiology of both strains CV101 and CV103 was studied previously in [6]. Several important observations are summarized here:

1. The glucose uptake rate of the glucose specialist CV103 is much stronger than the ability of CV101.
2. The high-affinity acetate uptake pathway Acetyl-CoA Synthetase (ACS) is barely expressed in CV103. However, the ACS enzyme is constitutively overexpressed in CV101.

It may look like that the coexistence of two strains through acetate cross-feeding is a coincidence. However, a long time balanced polymorphism is reported [12]. Over the entire 14,000-generation, the coexistence of two strains (designated L and S) was maintained, although the frequency of the S type fluctuated between approximately 10% and 85%. Experiments with conditioned media demonstrate that L and S secrete one or more metabolites that promote the growth of S but not of L.

4.2 The thermodynamic model for cross-feeding two strains

The model framework posed in the third section can be applied to model the stable acetate cross-feeding polymorphism of the *E. coli*. Therefore the coexistence of a glucose specialist and an acetate specialist is analyzed using the previous framework. To reduce the parameters used in the model, the majority of parameters related to the *E. coli* strains are kept the same for both strains. Because the glucose uptake rate and the acetate uptake rate caused by the ACS pathway is significantly different for two strains, only the maximal uptake rate of glucose $k_{e,i}$ and acetate $k_{u,i}$ is assumed to be different. All other parameters are kept the same for both strains.

Two strains of *E. coli* are introduced into the system with $i = 1, 2$ representing the glucose specialist and the acetate specialist respectively. There are several assumptions taken according to the physiology of the two strains introduced above:

1. $k_{e,1} > k_{e,2}$: the glucose uptake rate of the first one is stronger than the second one.
2. $k_{u,1} = 0$: the acetate uptake rate of the first strain is set as zero because the Acs pathway of the glucose specialist is barely expressed.
3. $J_{u,2} = \frac{k_{u,2}[A_{ex}]}{K_{mu} + [A_{ex}]}$: the acetate uptake rate of the acetate specialist is independent of the metabolic enzyme $[E]_2$ because the ACS enzyme is constitutively expressed.

$$\begin{aligned}
\frac{dN_i}{dt} &= \underbrace{\Lambda_i N_i}_{\text{growth rate}} - \underbrace{\delta N_i}_{\text{dilution rate}} \\
\frac{d[S_{ex}]}{dt} &= \underbrace{\delta([S_0] - [S_{ex}])}_{\substack{\text{dilution of} \\ \text{primary carbon source}}} - \sum_i \underbrace{\frac{2k_e[E]_i[S_{ex}]}{K_{ms} + [S_{ex}]}}_{\substack{\text{maximum precursor} \\ \text{synthesis rate}}} \underbrace{\frac{K_{ip}}{K_{ip} + [P]_i}}_{\substack{\text{feedback} \\ \text{inhibition}}} N_i V_r \\
\frac{d[A_{ex}]}{dt} &= - \underbrace{\delta([A_{ex}])}_{\substack{\text{diffusion of} \\ \text{fermentation product}}} + \sum_i \left(\underbrace{\frac{v_{a,i}[P]_i(1 - \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i})}{K_{ma,i} + [P]_i(1 + k_{ra,i} \frac{1}{K_{ca,i}} \frac{[A_{ex}]}{[P]_i})}}_{\substack{\text{reversible fermentation rate} \\ \text{with thermodynamic constraints}}} \right) N_i V_r - \underbrace{\frac{k_{u,2}[A_{ex}]}{K_{mu} + [A_{ex}]}}_{\substack{\text{fermentation product} \\ \text{uptake rate}}} N_2 V_r \\
\frac{d[P]_1}{dt} &= 2J_{g,1} - J_{o,1} + -J_{b,1} - J_{tca,1} - \Lambda_1 [P]_1
\end{aligned}$$

$$\begin{aligned}
&= \underbrace{\frac{2k_{e,1}[E]_1[S_{ex}]}{K_{ms} + [S_{ex}]}}_{\text{maximum precursor synthesis rate}} \underbrace{\frac{K_{ip}}{K_{ip} + [P]_1}}_{\text{feedback inhibition}} - \underbrace{\frac{v_a[P]_1 \left(1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_1}\right)}{K_{ma} + [P]_1 \left(1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_1}\right)}}}_{\substack{\text{reversible fermentation rate} \\ \text{with thermodynamic constraint}}} - \underbrace{\frac{\beta \Lambda_1}{\beta \Lambda_1}}_{\substack{\text{precursor consumption rate}}} \\
&\quad - \underbrace{\frac{k_{tca}[E]_1[P]_1}{K_{tca} + [P]_1}}_{\substack{\text{rate of TCA cycle}}} - \underbrace{\frac{\Lambda_1[P]_1}{\Lambda_1[P]_1}}_{\text{precursor dilution rate}} \\
\frac{d[P]_2}{dt} &= 2J_{g,2} - J_{o,2} + J_{u,2} - J_{b,2} - J_{tca,2} - \Lambda_2[P]_2 \\
&= \underbrace{\frac{2k_{e,2}[E]_2[S_{ex}]}{K_{ms} + [S_{ex}]}}_{\text{maximum precursor synthesis rate}} \underbrace{\frac{K_{ip}}{K_{ip} + [P]_2}}_{\text{feedback inhibition}} - \underbrace{\frac{v_a[P]_2 \left(1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_2}\right)}{K_{ma} + [P]_2 \left(1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_2}\right)}}}_{\substack{\text{reversible fermentation rate} \\ \text{with thermodynamic constraint}}} + \underbrace{\frac{k_{u,2}[A_{ex}]}{K_{mu} + [A_{ex}]}}_{\text{fermentation product uptake rate}} \\
&\quad - \underbrace{\frac{\beta \Lambda_2}{\beta \Lambda_2}}_{\text{precursor consumption rate}} - \underbrace{\frac{k_{tca}[E]_2[P]_2}{K_{tca} + [P]_2}}_{\substack{\text{rate of TCA cycle}}} - \underbrace{\frac{\Lambda_2[P]_2}{\Lambda_2[P]_2}}_{\text{precursor dilution rate}} \\
\frac{d[R]_i}{dt} &= J_{r,i} - \Lambda_i[R]_i \\
&= \underbrace{\frac{\beta \Lambda_i}{m_r} \frac{K_{ix}}{K_{ix} + [X]_i}}_{\substack{\text{ribosome synthesis rate}}} - \underbrace{\frac{\Lambda_i[R]_i}{\Lambda_i[R]_i}}_{\text{ribosome dilution rate}} \\
\frac{d[X]_i}{dt} &= \underbrace{k_x \frac{K_{mp}}{K_{mp} + [P]_i}}_{\substack{\text{regulator synthesis rate}}} - \underbrace{\frac{d_x[X]_i}{d_x[X]_i}}_{\text{regulator degradation rate}} \\
\Lambda_{c,i} &= \underbrace{\frac{k_r[R]_i[P]_i}{\beta(K_{mp} + [P]_i)}}_{\text{carbon limited growth rate}} \\
\Lambda_{e,1} &= \sigma J_{ATP,1} = \sigma \{n_g J_{g,1} + n_{tca} J_{tca,1} + n_o J_{o,1}\} \\
&= \sigma \underbrace{\left(n_g \frac{k_{e,1}[E]_1[S_{ex}]}{K_{ms} + [S_{ex}]} \frac{K_{ip}}{K_{ip} + [P]_1} + n_{tca} \frac{k_{tca}[E]_1[P]_1}{K_{tca} + [P]_1} + n_o \left(\frac{v_a[P]_1 \left(1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_1}\right)}{K_{ma} + [P]_1 \left(1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_1}\right)} \right) \right)}_{\text{energy limited growth rate of the glucose specialist}} \\
\Lambda_{e,2} &= \sigma J_{ATP,2} = \underbrace{\sigma \{n_g J_{g,2} + n_{tca} J_{tca,2} + n_o (J_{o,2} - J_{u,2})\}}_{\text{energy limited growth rate of the acetate specialist}} \\
&= \sigma n_g \frac{k_{e,2}[E]_2[S_{ex}]}{K_{ms} + [S_{ex}]} \frac{K_{ip}}{K_{ip} + [P]_2} + \sigma n_{tca} \frac{k_{tca}[E]_2[P]_2}{K_{tca} + [P]_2} \\
&\quad + \sigma n_o \left(\frac{v_a[P]_2 \left(1 - \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_2}\right)}{K_{ma} + [P]_2 \left(1 + k_{ra} \frac{1}{K_{ca}} \frac{[A_{ex}]}{[P]_2}\right)} - \frac{k_{u,2}[A_{ex}]}{K_{mu} + [A_{ex}]} \right)
\end{aligned}$$

$$\begin{aligned}\Lambda_i &= \min\{\Lambda_{b,i}, \Lambda_{e,i}\} \\ [E]_i &= \underbrace{\frac{\beta - m_r[R]_i}{m_e}}_{\text{metabolic protein concentration}}\end{aligned}$$

4.3 Model calibration

Although many parameters used in the model can be found in other articles, there are still some parameters unknown: the maximal rate of glucose uptake k_e , the maximal rate of acetate uptake of the Acs pathway k_u , the maximal rate of the TCA cycle k_{tca} , the Michaelis constant of the TCA cycle K_{tca} .

To estimate those parameters, the single-cell growth model is calibrated to the experimentally observed data. The experimentally measured data we used is from the paper [7] that proved that the acetate flux in *E. coli* is determined by the thermodynamic control of the Pta-AckA pathway. In the paper, glucose is provided sufficiently and the acetate concentration is varied. With the variation of the acetate concentration, the change of the net acetate accumulation flux, the glucose uptake flow, and the growth rate of the *E. coli* was measured. This data provides an ideal case for calibrating our model by varying unknown parameters.

The calibration process is performed by using a nonlinear optimization technique (implemented as `lsqnonlin` in MATLAB R2016b, Mathworks Inc.). Unknown parameters are varied to find to minimize the distance between the experimentally measured data and the data generated by the single-cell growth model.

4.4 Coexistence under different nutrient conditions

To study the coexistence of two strains under different nutrient conditions, the maximal glucose and acetate uptake rates $k_{e,i}$ and $k_{u,i}$ of both strains are needed to be specified. Here $k'_{e,1} = 3$, $k'_{u,1} = 0$, $k'_{e,2} = 1$ and $k'_{u,2} = 3$. The patterns of coexistence is generated for different glucose pool concentrations $[S_0]$ and different dilution rates δ of the chemostat.

4.5 Coexistence of evolved strains

To study the coexistence of evolved strains, the glucose pool concentrations $[S_0]$ and the dilution rate of the chemostat δ is fixed: $[S_0] = 10$ mM and $\delta = 0.2$ h⁻¹. To simplify the calculation, parameters of the first strain is fixed: $k'_{e,1} = 3$ and $k'_{u,1} = 0$. While $k_{e,2}$ and $k_{u,2}$ of the second strain is changed, and it is used to calculate the coexistence pattern with the first strain under the chemostat condition previously mentioned.