**Simulations**

1. Predict the maximal growth rate

The ME model is expected to predict a maximal growth rate that the M model cannot predict. In this part, we used the M model and the ME model to predict the maximal growth rate using glucose as the main carbon source. Therefore, we gradually increased the upper bound of the glucose uptake rate to search the maximal growth rate while keeping unlimited the uptake rates of the other components in the medium, but 0 for oxygen uptake rate.

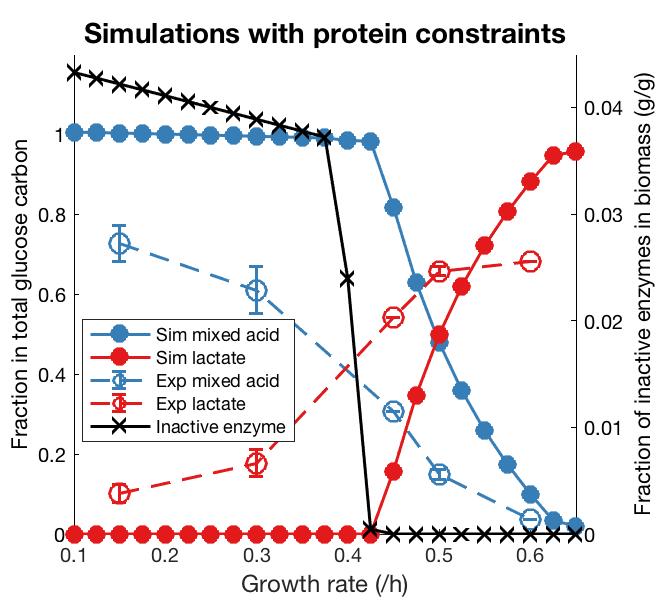
Due to the fact that L. lactis is auxotrophic for some amino acids, amino acids should be added to support growth *in vivo* and *in silico*. This, however, complicates the optimization process using either the M or the ME model since amino acids will be also used as the carbon source by the model if they are not constrained. Accordingly, we set the amino acid uptake rates as functions of growth rate, which were derived from experimental data ().

As amino acid uptake rates are functions of growth rate, the growth rate should be used as an input in each simulation. Therefore, binary search was performed for the simulations of both the M model and the ME model, although the ME model also contains more parameters using growth rate as the input. In the binary search process, growth rate will be assigned to see whether or not there is a feasible solution. Although the objective function does not affect the final maximal growth rate (24084808), we set the objective function as to minimize the glucose uptake rate.

1. Predict the metabolic shift

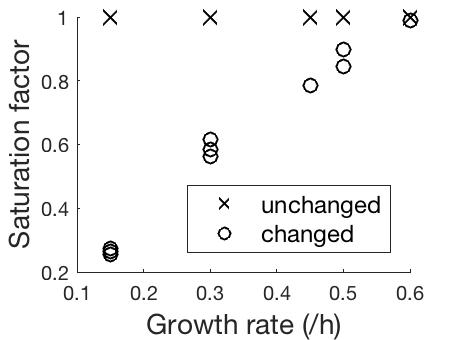
The metabolic shift from mixed acid to lactic acid fermentation of L. lactis cannot be captured by a normal GEM as the model only chooses high energy yield pathway. By integrating protein expression into the GEM, the model is expected to predict the metabolic shift. Therefore, simulations were performed using the protein-constrained GEM to try to explain the shift. We fixed the growth rate at a time in the range between 0.1 and 0.65, and also the upper bound of amino acid uptake rates while keeping unlimited the uptake rates of the other components in the medium. The objective function was to minimize the glucose uptake rate.

As we expected, the model with protein constraints can simulate the metabolic shift. Interestingly, the model produces an inactive enzyme in mixed acid fermentation while no inactive enzyme is produced in lactate fermentation. This indicates that the metabolic shift is caused by protein constraints.

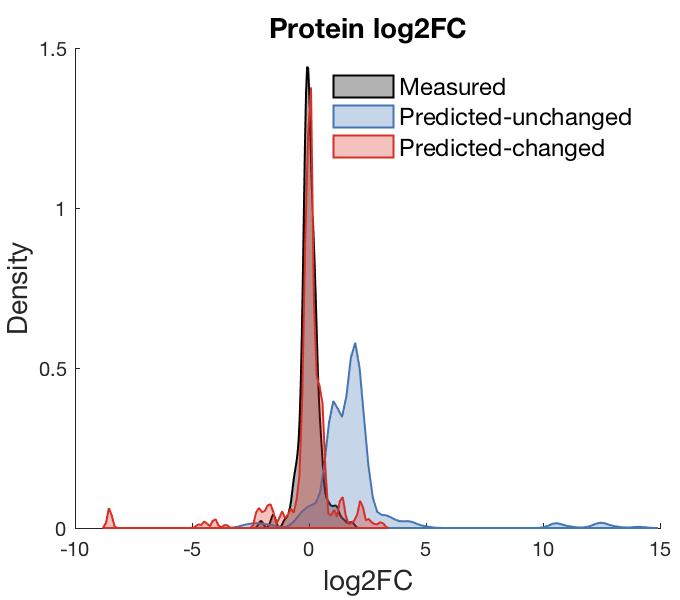
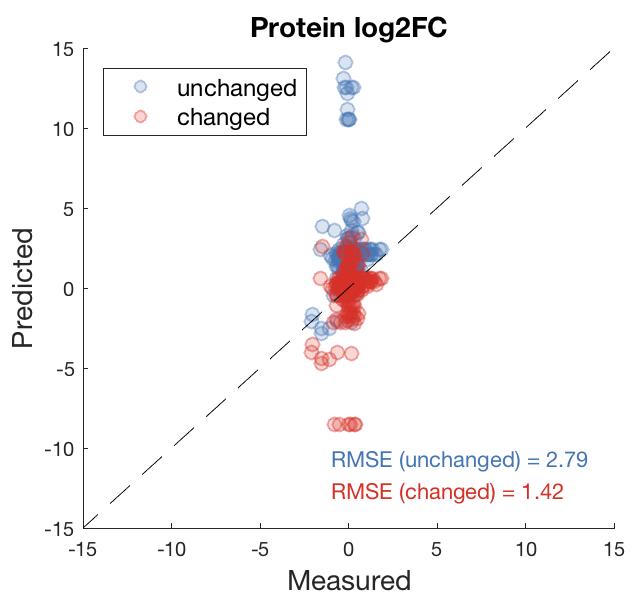


There comes the question: why the model produces the inactive enzyme at low growth rates. We found that the model has to produce the inactive enzyme, even though it does not carry any fluxes, at low growth rates to fill the gap between the sum of calculated active components and the given value (0.56 g/gCDW) of the total protein and RNA as an equality constraint. Even though we used inequalities to couple metabolic reaction rates and enzyme synthesis rate, the optimization process or the solver can only regard those as equalities, leading to the fact that all the active enzymes are operating at maximal kcat values. This seems to be biologically impossible. Therefore, we then proposed a method to eliminate the inactive enzyme by decreasing a global saturation factor. In this method, the production rate of the dummy protein is maximized while constraining some exchange reactions. Then the weight of inactive enzyme is calculated, and if it is greater than 0 the global saturation factor should be decreased. Again, maximize the dummy protein and check inactive enzyme until the weight is 0.

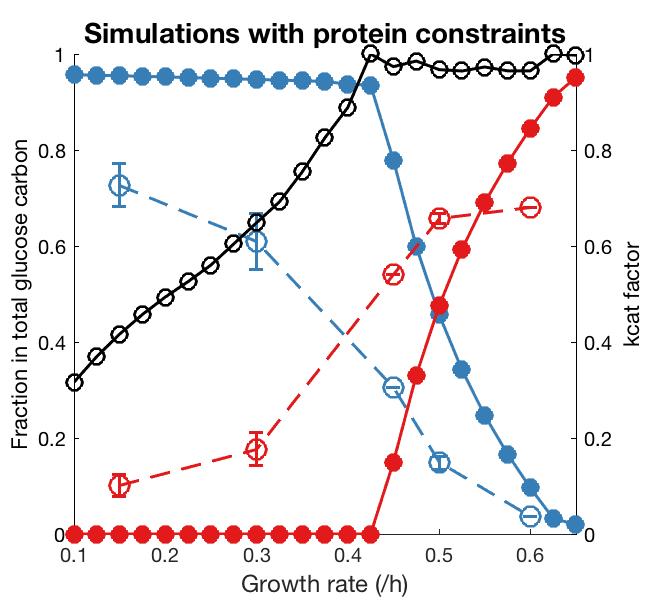
Next, experimental data () was used to evaluate this method by comparing simulations with and without changing the global saturation factor. We can see that the global saturation factors of any growth rates are below.



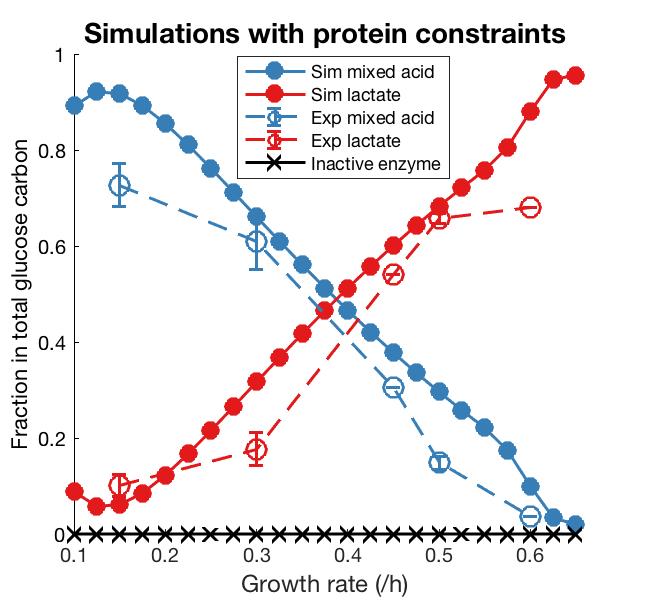
Then we checked if changing the global saturation factor can improve predictions. Firstly, we calculated total protein and RNA with and without changing the global saturation factor but found that it did not affect so much. Then, we calculated intracellular protein levels in metabolism and found that changing the global saturation factor enables more accurate predictions compared to unchanging. Since there is no absolute proteomics data, we used the relative data from the study (), which were normalized to the reference condition D = 0.15 h-1, i.e., log2 foldchange values. We found that neither changing nor unchanging the saturation factor can lead to a good agreement with experimental data in terms of correlation. This is because we just used a global saturation factor for all the enzymes while in reality each enzyme has its own one, so such a poor correlation between predicted and measured levels is expected if we compared for each enzyme. However, the calculated RMSE value of simulations with changed saturation factor is approximately half of the simulation with unchanged one. In addition, we can see from the density distribution that changing the saturation factor enables the predicted dataset to move much closer to the experimental data. And the predicted result is consistent with the conclusion that protein levels of most enzymes do not change with growth rate. Accordingly, it should be better to change the global saturation factor in each simulation.



Then we searched for global saturation factors for those simulations. It can be seen that the global saturation factor increases with growth rate before entering lactate fermentation. And during lactate fermentation, the global saturation factors are almost equal to 1.



However, this cannot explain why lactate is produced at low growth rates in reality. Then we found that saturation factors increase with growth rate (26498510, 27351952), and in *L. lactis* saturation factors from low to high growth rates (), but the model can only predict the increase from low to medium growth rates. Besides, we found from our simulations that the global saturation factor has a very good correlation with growth rate. Accordingly, we added the equation (saturation factor = 1.76 x mu) into our model. Then we can get a very good result.



It shows that at each growth rate the proteins are always constraining the model as no inactive enzyme is produced at each growth rate.

1. Estimate saturation factors for some enzymes with proteomics data

Considering the fact that changing the global saturation factor enables better predictions, we wanted to assign specific saturation factors for some enzymes to further improve model performance. Here, we used experimentally measured proteomics data combined with fluxes data calculated using the normal metabolic model to estimate specific saturation factors. This is based on the equation: V = saturation\_factor \* kcat \* [E]. We assumed that the kcat for each enzyme keeps constant across growth rates. Then when given changes in enzyme concentration and metabolic rate, we can calculate change in saturation factor between two steady state conditions. Accordingly, using fluxes and relative proteomics data, we estimated changes in saturation factors for many enzymes. We found that most of the enzymes show very good correlations between growth rate and saturation factors, and distribution…….. We also found that all the good correlations are positive, meaning that in most cases saturation factor increases with growth rate. As we found in the previous analysis that the global saturation factor is approximately equal to 1 at the growth rate of 0.6, we assumed that the absolute value of the saturation factor is 1 at 0.6 h-1 for each enzyme of interest. Then the saturation factor of the other growth rates can be calculated using the ratio. Then we established relationship between growth rate and saturation factor for each enzyme, and assumed that all the linear equations go through (0,0) point, i.e., the intercepts are all 0. Then we can use the information as input when doing simulations. Notably, we found that most of the negative correlations were the enzymes involved in mixed acid fermentation pathway. All of them showed highest saturation factors at growth rate of 0.3 h-1, thereby displaying an increase and then a decrease trend along growth rate. We fit all the curves with equations, which can be used in simulations. But when using all the specific saturation factors in simulations, they seemed to over-constrain the model. This could be explained by uncertainties in proteomics measurements and the collected kcat values in the model. So we should still use the global saturation factor. In fact, the calculated global saturation factor is very close to the median value of the estimated specific saturation factors.

There are three types of saturation factors in the model. The first is the saturation factor of glucose transporter, which was estimated using the method above. The second is about the transporters of unlimited compounds in the medium, whose saturation factor is assumed to be 1 at any condition. The last is the rest of the enzymes, and we used the method mentioned above to search for the global saturation factor for them.

1. Determination of PTS saturation factor

It was determined according to the analysis above and calculated to be around 1.5 mu, which increases almost linearly with growth rate. This is in line with PTS activity in E. coli (ME model of E. coli).

1. Fdgdfgd
2. Sersfsdfs