NUTRIENT-LIMITATION IN COMPLEX MEDIUM

(not a paper draft)

1 NUTRIENT LIMITATION

Definition 1: Nutrient Limitation

We consider a network to be nutrient limited by metabolite A if:

$$\delta|U_z|/\delta|U_A| > 0$$

for some range $U_A \in [0, C]$ (an intake regime). Where U_z is the biomass upper maximum feasible value and U_A is the maximum uptake value allowed for metabolite A.

Definition 2: Essential nutrient

We consider a metabolite E to be essential if:

$$\lim_{|U_E| \to 0} |U_z| = 0$$

and

$$\lim_{|U_E| \to +\infty} |U_z| > 0$$

Our aim is to discuss the main network configurations which might lead to condition 1.

1.1 MODEL

The aim is to explore the network reduction space so we can find configurations where the desired nutrient is limited. Using an abstract network model, we can study the different mechanisms which lead to nutrient limitation. The model is defined using total reactions, that is, $S \Rightarrow P$ does not necessarily means that it is a single reaction, it represents all reactions, or pathways, that produce P from S. The model has the follow general reactions:

Exchange reactions

Ex_A: $(-1)\ A \Leftarrow$

Ex_E: $(-1)\ E \Leftarrow$

Ex_S: $(-1)\ S \Leftarrow$

Biomass equation

z: $(-1)\ E + (-1)\ P_1 + (-1)\ P_2 \Rightarrow$

Internal reactions

: $(-1)\ A \Rightarrow\ P_1$

: $(-1)\ A \Rightarrow\ P_2$

: $(-1)\ S \Rightarrow\ P_1$

: $(-1)\ S \Rightarrow\ P_2$

The model has three nutrients E, A and S: E is an always essential, A is the metabolite of interest and S is another general metabolite, which compete with A for being limiting. Also, the model has internal general reactions connecting the nutrients with the biomass components P_1, P_2. That means, neither A nor S alone are essentials.

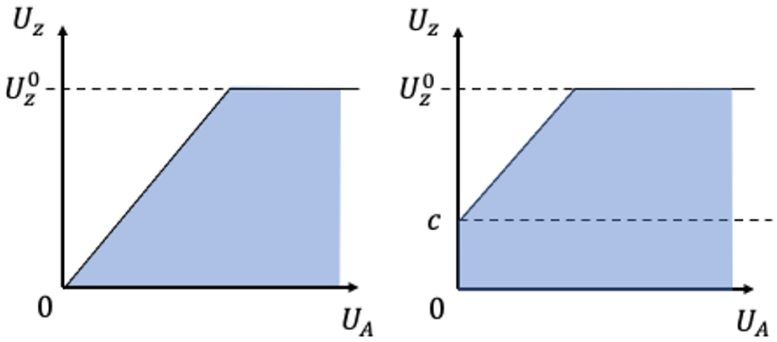

Figure 1: The two possible limiting cases. U_z is the upper limit for z given all constraints other than U_z^0. U_A is the upper limit for the intake of A. c is the value of z achievable without intaking any A. Note that in a general network, the line of dependency could have changes on slope (due to changes in yield), but the main feature of continuous regions of limitations (with non zero slope) remains general. See main text for more analysis.

Assuming that $z = 0$ is inside the polytope (a vertex), the convex nature of it allow us to conclude that there is only two possibilities for A to be limiting: i) A to be essential (left panel Figure 1), and ii) A to be what I call $z > c$ limiting (right panel).

In case (i), there is at least one biomass component that can't be produced without A. It would be the case if $S \Rightarrow\ P_1$ (or $S \Rightarrow\ P_2$) is knocked out, or if the intake of S is blocked. Case (ii) is just a relaxed version of the first. The network is able to grow without A, but till a maximum c. Above that threshold, A is required and becomes a limiting factor. In our model, setting $(S \Rightarrow\ P_1) < a$ produces such a situation, given that $z < U_z^0$ at $U_A = 0$. The same is true if the intake of S is restricted enough, A can become limiting. Note how in the last example, both nutrients S and A will be limiting.

As can be noted, even in this small model, there are several configurations which leads to nutrient limitation. Is this degeneracy which complicates the contextualization of metabolic networks aimed to reproduce such phenotype. One approach to learn about possible solutions is an exploration of the space of all limiting configurations. Regarding possible exploration strategies, the two cases mentioned before have important differences.

All instances of case (i) rely inside the 'KO space'. That is, all the configuration which has blocked reactions with respect to the original. Without further constraints this space would be exponentially big with respect to the number of reactions. But, we HOPE the network to be sensible enough so much of such space would be unfeasible. The more reactions are blocked, the more probable it is for the network to yield zero biomass. The good news is that at least, this space is well defined.

For the case (ii), the situation is more complicated. We have continuous constraints c, which is harder to explore. Plus, we might also have combinations of KO reactions and partial limitations which are compatible. This space will be both hard to define and big.

2 INTERESTING EXPERIMENTAL DATA

We are trying to find experimental data which allows us to evaluate the importance of the nutrient limitation constraint over a metabolic model. Ideally, we can manage to introduce into the model formulation the information contained in the phenotype of nutrient limitation, and with good luck, this extra information might explain an improvement in the model performance. Conversely, we can evaluate the risk of not taking into account such information. That is, we shouldn't use models which fail to reproduce the experimental pattern of nutrient limitation, which is not uncommon practice in literature. In general, for complex medium and/or mammalian cell cultures I haven't found good/clear/appropriate/interesting data in the literature. Here are the main ones.

2.1 CANCER AND GLYCINE

At [1] they evaluate the patterns of consumption/production of metabolites for a large set of cell lines derived from cancer tissues. A clear result is that for a subset of such lines glycine is strongly correlated with growth (see Figure 2, left panel).

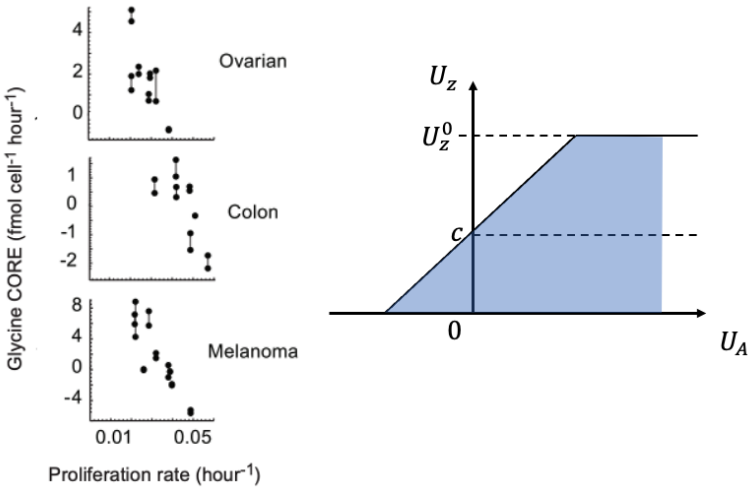

Figure 2: Left panel: Correlation between glycine exchange flux (y axis) and growth rate (x axis) for cell lines extracted from three types of cancers. Reproduced without authorization from [1]. Right Panel: If you flip the axis and the convention for intakes/production, you realize that the trend is compatible with a $z - c$ limitation (same as right panel on Figure 1), but this time the dependency is extended also toward negative exchanges (production). See main text for further analysis.

Quick analysis of this trend shows that glycine is not an essential metabolite, it is instead "$z > c$ limiting". Actually, it is even more complicated that our initial definition of "nutrient limitation" because the trend is traceable till the production regime (see Figure 2, right panel). That is, for $z < c$, glycine is still correlated with growth in the same way but it became a waste not a nutrient. The authors also point out that this subset of lines with good correlations between gly and the proliferation rate are the ones growing faster.

So, what can we do with this? The honest answer is "I don't know"! But let's talk about it. This is still the best paper describing some kind of relationship between complex medium cultures growth rate and an external metabolite exchange. Actually, the paper is very complete, it includes also transcriptomic screenings and an culture-supported hypothesis about why glycine is important for fast growing cells.

From our point of view, we might want to see if including "somehow" into the model formulation just the nutrient limitation phenotype, we can reproduce or get closer to, for instance, the same explanation for why glycine is important. Or reproduce the general transcriptomic pattern. Other approach could be to attempt to compute how likely from a naive point of view is the experimental configuration. We can find, for instance, that because of an entropic effect, the experimental phenotype is backed by a dominant number of configurations.

What are the problems? In previous attempts, we were trying to introduce/evaluate the phenotype constraint by exploring the ensemble of all nutrient-limited networks inside the KO feasible space. That is, we were targeting a Figure 1 left panel case. This is the easier one because it only depends on the topology of the network (no free parameters). But, even in this case, success was in jeopardy because the KO space is exponentially big and we are dearing to work with the biggest available network of all, Human1. On the other hand, the case at [1] is obviously a $z > c$ limiting case. Now we have a free parameter, the bottleneck constraint (the cause and value of c). We need to introduce extra information, for instance, an intake/production set of constraints (which aren't reported in the paper). Then, the results being sensible to its availability and accuracy. This extra degeneracy means that this space is even bigger than just the KO space. We now need to explore limiting (not blocking) the reactions of the network.

Ideas to think about: i) In the experimental data we have the slope of the relationship between glycine and growth, not only the trend. I suspect that this is a polytope edge. Yes, something is being optimized that is laid on this edge. This slope is the shadow price of the glycine. It will depend on the combination of the yields of the active elementary modes. Again, for a case (i) this only depends on the topology, but for a case (ii), this is harder to explore, because it depends on how you restrict the network. ii) The fact that the correlation goes till the production part of the glycine exchange tells us that the steady state assumption we make in the formulation of $Av = 0$ is accurate.

2.2 RATH AND THE CHEMOSTAT

Well, we have exchange fluxes from a set of cultures that the author [2] is reporting are glucose limited (see Figure 3). We can explore if marginalizing on the subnetworks which are glucose-limited we improve the reproduction of such data compare with a random subnetwork. We can explore just the KO space because it is easy. The bad news is that reproducing fluxes is harder than just finding general KO patterns (transcriptomic data) based on topology. Also, it is not a lot of data.

3 SUMMARY

Computations are probably unfeasible, questions aren't clear, data non available or noisy, rewards not very promising. Unfortunately, just like 6 months ago.

REFERENCES

[1] Mohit Jain, Roland Nilsson, Sonia Sharma, Nikhil Madhusudhan, Toshimori Kitami, Amanda L. Souza, Ran Kafri, Marc W. Kirschner, Clary B. Clish, and

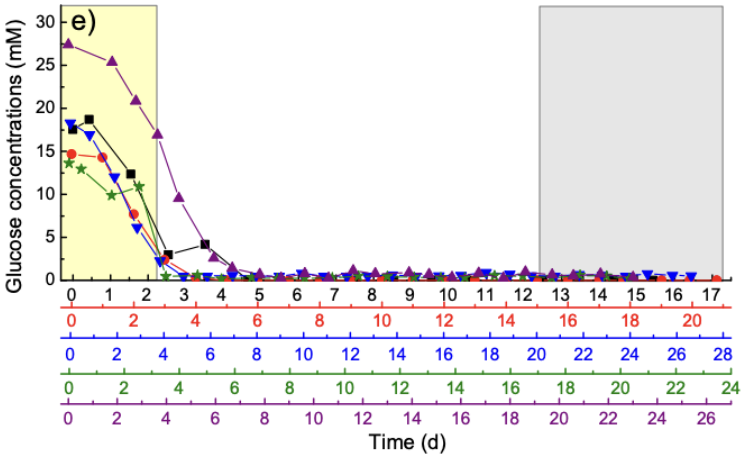

Figure 3: The time series of remaining glucose for 5 parallel continuous cultivations. Reproduced without request from [2]

Vamsi K. Mootha. Metabolite Profiling Identifies a Key Role for Glycine in Rapid Cancer Cell Proliferation. *Science*, 336(6084):1040–1044, May 2012.

[2] Alexander Rath. *Characterisation of Cell Growth, Metabolism and Recombinant Protein Production during Transient and Steady State Conditions for the Human Cell Line AGE1.HN-AAT*. PhD thesis, 2017.