Kinetic models

December 10, 2015

Summary

Background

In our project we really tried to make the dry lab and the wet lab go hand in hand. To achieve this we tried to solve some problems that we wanted to solve before we would alter strains in the wet lab with the help of genome-scale FBA models. After the (altered) strains were characterized, we used the parameters from physiology measurements for further kinetic modeling of our consortium. These kinetic models help us in making predictions for further applications.

Aim

Even before we would try to create the final consortium we wanted to fiend out how we could create a stable carbon compound producing strain. We also wanted to know what genes we could knock-out to create an auxotrophic organism. After we created and characterized these strains we also wanted to know if and under what conditions the ratio of the participants in our consortium would converge, and if so, to what value. Furthermore we wanted to know the influence of the presence of *E. coli* on the light and thus on the growth of *Synechocystis*.

Approach

We created two algorithms which work with genome-scale FBA models. One algorithm, we named the stable compound generator, searches for ways to make a strain genetically stable produce carbon compounds. It is based on the idea that if the production of a compound is growth coupled, it is harder of an organism to get rid of it. The second algorithm searches for ways to create an auxotrophic strain. It is based on the idea that an auxotroph is no longer able to grow without the compound it is dependent on in the medium. However if you add this compound to the medium it is suddenly able to grow. The kinetic models are sets of differential equations, which we then analyzed with pydstool, a python tool. The equations model the amount of biomass formed and are based upon monod equations.

Results

The algorithms can be found in the results section. The algorithm which searches for ways to create genetically stable producers found several compounds which can possibly be produced in a genetically stable way. We chose a *Synechocystis* strain which can produce acetate. The algorithm which searches for ways to create auxotrophs provided us with a list of compounds we could make *Synechocystis* dependent on. We chose argenine and proline. The kinetic models seem to suggest that in a flask and in a chemostat the ratio of the two organisms converges to the same ratio under different initial conditions. This ratio is however dependent on some parameters. The growth of *Synechocystis* is either light dependent, or dependent on a substance *E. coli* produces. When light dependent, the influence *E. coli* seems to have on the light and thus on the growth of *Synechocystis* seems not that big. When substrate dependent *E. coli* and *Synechocystis* seem to converge to the same ratio again.

Background

If you take a look at the modeling part of a lot of iGEM projects, you will find that they are often almost a separate part from the rest of the project. We wanted to use the modeling more as a tool. A part of the project

that would provide for answers used elsewhere in the project. Our team itself tried to become a consortium, with the modeling and the wet lab exchanging information. We wanted a clear connection between the modeling and the other modules of the project. At the start of the project, we used mostly models based on Flux Balance Analysis (FBA). These models provided information which is used in the lab. We created two algorithms which provided this information. One algorithm searches for ways to make a stable producing organism, the other one searches for ways to make an auxotrophid organism. In the lab they then characterized the organisms used in the consortium. From this data we could extract the parameters needed as input for kinetic models we created.

Algorithms

Aim

Genetically stable production

A big challenge in biotechnology and synthetic biology today is genetic instability. It has been called the elephant in the room which nobody could (or wanted) to see [Jones, 2014]. A lot of the applications in synthetic biology work because an organism contains a certain gene construct which contains a set of genes which, when expressed, produces enzymes responsible for the synthesis of certain products.
These products are then excreted from the cell to be used by other organisms (us humans for example). A lot of parts in the part registry of iGEM also work based on this 'plug-and-play' rationale. After we, iGEM teams, find such an interesting set of genes, we register it, we characterize it and we envision big scale applications. But in reality a lot of these applications wouldn't work. This has everything to do with genetic instability.
In most circumstances in the lab where organisms are cultivated, or in bioreactors of any form, species are under a constant selection pressure for growth rate. This means that mutants with a slightly higher growth rate will take over an entire population within a few generations. An organism which produces a compound and transports it out of the cell, always has to divide the resources it has between growth and the production of the compound. The producing cells continuously diverge a part of the mass out of the cell. This means it can no longer use this mass for growth. So a hypothetical mutant which loses this producing activity will have a (slightly) higher growth rate, and thus soon take over the culture, making it loose its production. So if an organism can easily accumulate mutations such that it will loose its production, it will be genetically very unstable. And sometimes all it takes for a producing strain to lose a gene activity is a single mutation in the promoter.
In the physiology section (link) you can see the results when we tried to characterize several producing strain. In the results, we can see that indeed after a while the growth rate of non stable producers goes up, while at the same time the production of compound goes down. These results seem to be in line with our ideas about genetic instability.
Certainly in large scale applications the costs and effort to re-inoculate a culture, because the production of the old one has stopped can add up to significant amounts. This makes these applications not only economically less feasible, it is also less sustainable, since a lot of recourses are needed to re-inoculate.
We see it as a major challenge in synthetic biology to improve this stability. So this was also the first question we had in the wet lab. How do we make a *Synechocystis* strain which stably produces a carbon compound? We created an algorithm to answer this question, called the Stable compound generator.

Creating an auxotroph

Auxotrophs are organisms which need a certain compound which they can take up from their environment to be able to grow. They can be very useful in synthetic biology, as an auxotrophy can regulate the growth of a certain organism. It is also useful synthetic consortium, as it can create a dependency of one species on another. In our project we wanted to make the chemoheterotroph dependent on the cyanobacterium by the carbon compound the cyanobacteria produces, because this will be part of the flux of CO_2 into product. However, we also wanted to create dependency of the cyanobacterium on the chemoheterotroph. Most important reasons are that this dependency will decrease the risk of the cyanobacterium surviving on its own in the environment in case of an outbreak. Some chemoheterotrophs are very efficient in the production of a certain compound. If the photoautotroph does not have to synthesize the compound itself, it may increase the growth rate of the photoautotroph. This increases the rate at which carbon is fixated and carbon compound is produced, which in its turn benefits the chemoheterotroph. It is also suggested that this interdependency may create a more robust system. So the second question that we had in the lab is how can we create an auxotroph? To answer this question we also created an algorithm, the Auxotrophy sniper.

Approach

For some questions we had (see Aim) we decided to look for an answer based on algorithms for genome scale flux balance analysis (FBA) models. We have used genome scale FBA that already existed of *Synechocystis* and PySCeS CBMpy, a tool for **C**onstraint **B**ased **M**odeling using **Py**thon **S**imulator for **Ce**llular **S**ystems [Olivier, 2014-2015]. In flux balance analysis (FBA) you assume that all reactions in a cell are in steady state and then you can then represent them in a set of linear equations. The objective function then optimizes something given a set of boundary conditions about the flux through each reaction. The objective function is often the formation of biomass. If a cell is producing biomass, means it grows.

Stable compound generator

In previous section we emphasized we wanted to create **stable** carbon compound producing cyanobacterial strain. But how do we do this? We created an algorithm which finds ways to make a stable producer. It is based on two ideas:

1. It takes a lot more evolutionary time to re-create a whole new gene than to accumulate a loss-of-function mutation.
2. If the gene responsible for the production of the compound is expressed only when the organism grows (growth coupled production), the organism cannot simply stop expressing the pathway.

Below follows a general outline of the algorithm.
The algorithm makes a list of carbon compounds associated with the production of biomass in the cell.
For each of these compounds, the algorithm does the following:

- Find sources reactions of the compound in the extracellular space and set the boundaries of these reactions to zero.
- Find all reactions associated with the compound, as the model only contains reactions with gene associations and no genes which directly influence flux. We will call these reactions primary reactions
- Find the genes associated to the reactions.
- For each of these genes, find all reactions which have the gene of interest in their gene association. We will call these reactions secondary reactions
- Make a list of combinations of the genes associated to the production of a compound.
- For each of the genes involved each combination, set flux boundaries of these reactions to zero. In that way we simulate the knock-out of the gene.
- Since accumulation of compound is not possible in FBA, while in reality most cells are leaking, we create sinks for compounds which would otherwise possibly accumulate. These sinks prevent the model from going to a non-growing steady state, because the compound would otherwise accumulate.
- Do a flux balance analysis on the model en check if biomass is formed.
- Check the value of the sink of the compound of interest.

If there is still formation of biomass (growth) and the sink of the compound of interest is used to export the compound out of the cell, the knock-out of the combination of genes is a good candidate for making a stable producer.

Auxotrophy Sniper

We wanted to create an auxotrophic *Synechocystis* strain as part of our consortium. But how can we make an auxotroph out of *Synechocystis*? To get the answer to this question we decided to create an algorithm that works on genome scale FBA models. The general idea is that an auxotroph can be created by knocking out a combination of genes involved in the production of the compound we want to make a dependency on. The organism should not be able to grow if there if these genes are knocked out and there is no source for the compound in the extracellular space. When this is the case we create a source in the extracellular space (add compound to the medium) and after that the organism should be able to grow again. Here follows an overview of how the Auxotrophy Sniper works:
The algorithm takes a list of compounds of which you want to make a synthesis deficiency in the organism of

choice. This list can contain for example vitamins, or amino acids. For each metabolite on the list it does the following:

- Set flux boundaries of source reaction in the extracellular space (if there is already one present) to zero.
- Find all primary reactions associated to the compound we want to make the organism dependent on.
- For each primary reaction, find genes which are associated to the primary reactions.
- For each gene find all reactions which have the gene in their gene association (secondary reactions).
- Make a list of possible combinations of genes which can be knocked out.
- Go over the combinations one by one, for each gene in a combination turn off primary and secondary reactions.
- Per combination, do a flux balance analysis and check for biomass formation (growth).
- If it does **not** form any biomass, a source reaction of the compound is added to the model
- Another flux balance analysis of the model is done and if there is biomass formation now, the combination of genes are good candidates. If this combination of genes is knocked out the organism will probably become an auxotroph.

Results

The software needed to run the algorithms can be downloaded here: [Olivier, 2014-2015]. The model of *Synechocystis* that we used can be found here. The algorithms can be downloaded here.

Stable compound generator

In this figure an output of the program can be seen in this picture (add picture). This is a list of the compounds that can be stably produced by *Synechocystis* according to the algorithm. In our consortium we chose to use a strain with a knocked out ACS-gene. This strain of *Synechocystis* produces acetate. In the physiology page you can find a characterization of this strain. In line with our expectations, this strain does not lose its production over time, although extended research may be needed to prove the longer stability.

Auxotrophy sniper

We wanted to create an amino acid dependency in *Synechocystis*. This picture of the output of the

Kinetic models

Aim

We also wanted to answer some questions about the final consortium which will help in envisions of a final application. What will happen to the growth rate of the organisms when With what initial conditions and parameters will the ratio between the biomass of the different species converge to the same ratio and what will the ratio be? What is an estimate of the yield? What is the influence of the light intensity on *Synechocystis* and how will this be influenced by a the presence chemoheterotroph which blocks and scatters light. To answer these questions we created some kinetic models consisting of differential equations.

Approach

We modeled the biomass per liter of *Synechocystis* and Esherichia coli in chemostat as well as batch cultures. We created a set of differential equations and analyzed them with pydstool. A python tool, which can solve differential equations numerically.

Batch

Unlimited cell growth is exponentially. The amount of biomass per time for an exponentially growing species can be given by the following different equation:

$$\frac{da}{dt} = \mu a \tag{1}$$

Herein is a the amount of biomass per liter and μ the growth rate normalized for biomass. One can easily verify that the solution of this differential equation is indeed exponential growth ($a = ce^{\mu t}$). Now from experimental data it has been shown that limited growth on a substrate is a bit different. The normalized growth rate is then dependent on the concentration of substrate. According to the monod equation, the μ is dependent on $[S]$ in the following way:

$$\mu = \mu_{max} \frac{[S]}{k_S + [S]} \tag{2}$$

Herein is μ_{max} the maximal growth rate (equal to the growth rate at unlimited growth), and $[S]$ the concentration of substrate. k_S is the concentration of $[S]$ at a rate $\frac{1}{2}\mu_{max}$. We also know this equation from enzyme kinetics as the Michaelis-Menten equation. In enzyme kinetics this equation is used to calculate the rates at which enzymes convert products. However the Michaelis-Menten equation is based on theoretical arguments, while Monod is based on experimental findings. According to Monod, the growth rate saturates as the concentration becomes higher (see figure 1).

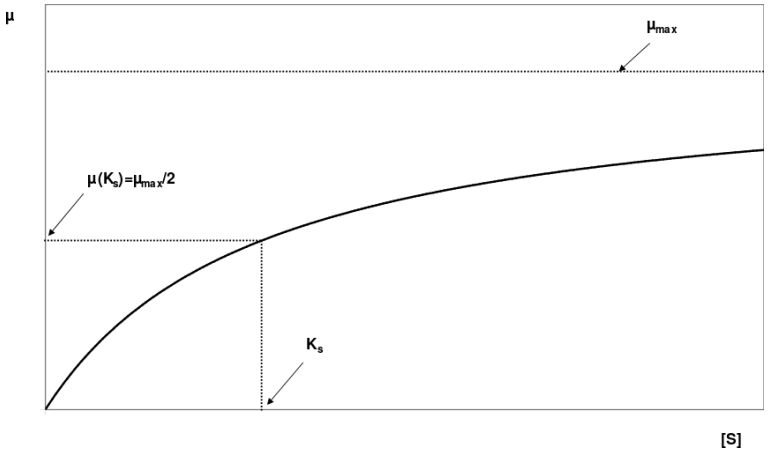

Figure 1: Limited growth on a substrate according to the Monod equation. μ is the normalized growth rate in units per hour μ_{max} is the maximal growth rate and $[S]$ is the substrate concentration. k_S is the concentration at the rate equal to $\frac{1}{2}\mu_{max}$.

In our consortium, *Escherichia coli* grows limited on acetate. So now we know how μ depends on the concentration of acetate. We now need to model the concentration of acetate. We assume the synthesis of acetate is growth coupled and depends linear on the growth of *Synechocystis*. We also know that the maximal uptake rate of acetate by *E. coli* is dependent $1/y$ herein is y the yield of *E. coli* on acetate in is in milligram Dry Weight *E. coli* per millimole acetate per liter. The uptake of substrate per unit time is also in a a saturable way dependent on the concentration of S, exactly the same way the growth speed is dependent on the the concentration of substrate. The synthesis of the substrate is growth coupled and is dependent on the amount of biomass of *Synechocystis* formed in time. The ch We now arrived at the following set of differential equations:

$$\frac{d\text{syn}}{dt} = \mu_{syn}\text{syn} \tag{3}$$

$$\frac{d\text{ec}}{dt} = \mu_{max,ec} \frac{[S]}{k_{s,ec} + [S]}\text{ec} \tag{4}$$

$$\frac{d[S]}{dt} = \frac{1}{y_{s,syn}}\mu_{syn}\text{syn} - \frac{1}{y_{s,ec}} \frac{[S]}{k_{s,ec} + [S]}\text{ec} \tag{5}$$

Herein is syn the amount of biomass of *Synechocystis* and ec the amount of biomass of *E. coli*. y_s is the substrate yield of *Synechocystis*. Since in this model the substrate is only formed when *Synechocystis* forms biomass, per amount of biomass formed, there is a constant amount of substrate formed. The yield is usually expressed in

gram dry weight mole substrate *used*. In this case we mean gram dry weight mole substrate *formed*. So to find the amount of substrate that is formed per amount of biomass that is formed we simply take $\frac{1}{y,syn}$.

It can be easily seen that such a relationship will not be stable if $\mu_{max,ec} << \mu_{max,cyn}$. In this case by stability we mean the convergence of the growth rate to the same value.

In this model, *Synechocystis* is not dependent on *E. coli*. There are two ways in which *Synechocystis* may be dependent on *E. coli* we have explored. Firstly, *E. coli* may produce a substrate *Synechocystis* grows on, as is the case with the auxotrophic *Synechocystis*. Secondly, *E. coli* may decrease the light intensity in the culture, in this way slowing down the growth of *Synechocystis*. The growth rate of *Synechocystis* can only be limited by one of these two processes and it will always be limited by the process that slows it the most. Either μ_{syn} is lower than $\mu_{max,syn}$ because there is a photon shortage, but then the amount of substrate available at that growth rate would be enough, or the amount of substrate is limiting, but then the amount of photons available would also be enough for that given growth rate. So actually the growth rate of *Synechocystis* would be

$$\min(\mu_{max,syn}\frac{[S_2]}{k_{s,syn}+[S_2]}, \mu_{max,syn} f(\text{syn}, \text{ec})) \tag{6}$$

Herein is $f(\text{syn}, \text{ec})$ a function which determines the factor of decrease in growth rate because of a photon shortage and it is a function of the amount of biomass per liter of *Synechocystis* as well as that of *E. coli*. If we now assume that the amount of substrate *E. coli* produces is going to be limiting we arrive at the following set of differential equations.

$$\frac{d\text{syn}}{dt} = \mu_{max,syn}\frac{[S_2]}{k_{s2,syn}+[S_2]}\text{syn} \tag{7}$$

$$\frac{d\text{ec}}{dt} = \mu_{max,ec}\frac{[S_1]}{k_{s1,ec}+[S_1]}\text{ec} \tag{8}$$

$$\frac{d[S_1]}{dt} = \frac{1}{y_{s,syn}}\mu_{max,syn}\frac{[S_2]}{k_{s2,syn}+[S_2]}\text{syn} - \frac{1}{y_{s1,ec}}\frac{[S]}{k_{s1,ec}+[S]}\text{ec} \tag{9}$$

$$\frac{d[S_2]}{dt} = Q_{p,ec}\text{ec} - \frac{1}{y_{s2,syn}}\frac{[S_2]}{k_{s2,syn}+[S_2]}\text{syn} \tag{10}$$

Herein is $Q_{p,ec}$ the amount of $[S_2]$ formed by *E. coli* per gram dry weight of *E. coli*. We assume *E. coli* doesn't produce in a growth coupled way, but has a constant production per amount of biomass.

0.0.1 Turbidostat

(link turbidostat) If we want to model the consortium in a turbidostat, we have to account for the fact that both *Synechocystis* and *E. coli* are increasing the OD as they grow. This means that the dilution rate is dependent on the biomass of *Synechocystis* as well as that of *E. coli*. For simplicity we make the assumption that there is a constant flow through the system, instead of only diluting when the threshold is reached. To understand this we first look at the case of a single strain, called b. In a chemostat the growth rate of the organism would become equal to the dilution rate. In a turbidostat however, an organism can grow at its maximal growth rate, but the amount of biomass must still become constant. This means the following:

$$\frac{db}{dt} = \mu \cdot b - D \cdot b = 0 \tag{11}$$

Where b is the amount of biomass of the strain b, mu is the growth rate and D the dilution rate.
In a chemostat it would mean that $D < \mu$ is a chosen dilution rate and that μ becomes equal to D due to substrate limitation. In a turbidostat however, D becomes equal to μ.
In the case where there are two strains, strain a and b, that share a turbidostat, the differential equations that then describe the system looks like the following:

$$\frac{da}{dt} = f(a,b,t) - Da \tag{12}$$

$$\frac{db}{dt} = g(a,b,t) - Db \tag{13}$$

Herein are f and g functions that describe the growth of the organisms a and b respectively. D is again the dilution rate. Now in a shared turbidostat it holds that $a + b = k$, where k is a constant amount of biomass.

Then the following holds:

$$a + b = k \tag{14}$$
$$\implies \frac{da}{dt} + \frac{db}{dt} = 0 \tag{15}$$
$$\implies f - Da + g - Db = 0 \tag{16}$$
$$\implies Da + Db = g + f \tag{17}$$
$$\implies D = \frac{g+f}{a+b} \tag{18}$$

For the turbidostat we then arrive at the following set of differential equations:

$$\frac{d\text{syn}}{dt} = \mu_{syn}\text{syn} - D\text{syn} \tag{19}$$
$$\frac{d\text{ec}}{dt} = \mu_{ec}\text{ec} - D\text{ec} \tag{20}$$
$$\frac{d[S_1]}{dt} = \frac{1}{y_{syn/S1}}\mu_{syn}\text{syn} - \frac{1}{y_{ec/S1}}\mu_{ec}\text{ec} - D[S_1] \tag{21}$$
$$\frac{d[S_2]}{dt} = Qp_{S2/ec}\text{ec} - \frac{1}{y_{syn/S2}}\mu_{syn}\text{syn} - D[S_2] \tag{22}$$

where

$$D = \frac{\mu_{syn} + \mu_{ec}}{\text{syn} + \text{ec}} \tag{23}$$
$$\mu_{syn} = \mu_{max,syn}\frac{[S_2]}{k_{s2,syn} + [S_2]} \tag{24}$$
$$\mu_{ec} = \mu_{max,ec}\frac{S_1}{k_{s1,ec} + [S_1]} \tag{25}$$

1 Light limited growth

When *Synechocystis* grows in a batch, after a while, the growth becomes limited to the light intensity. Like stated in section Franco-Lara et al. [2006],

$$\frac{d\text{syn}}{dt} = \mu_{max,syn}\frac{[S_2]}{k_{s2,syn} + [S_2]}\text{syn} \tag{26}$$
$$\frac{d\text{ec}}{dt} = \mu_{max,ec}\frac{[S_1]}{k_{s1,ec} + [S_1]}\text{ec} \tag{27}$$
$$\frac{d[S_1]}{dt} = \frac{1}{y_{s,syn}}\mu_{max,syn}\frac{[S_2]}{k_{s2,syn} + [S_2]}\text{syn} - \frac{1}{y_{s1,ec}}\frac{[S]}{k_{s1,ec} + [S]}\text{ec} \tag{28}$$
$$\frac{d[S_2]}{dt} = Q_{p,ec}\text{ec} - \frac{1}{y_{s2,syn}}\frac{[S_2]}{k_{s2,syn} + [S_2]}\text{syn} \tag{29}$$

Results

In figure

References

Patrik R Jones. Genetic instability in cyanobacteria–an elephant in the room? *Frontiers in bioengineering and biotechnology*, 2, 2014.

Brett G. Olivier. Python Simulator for Cellular Systems. `http://pysces.sourceforge.net/`, 2014-2015. [Online; accessed 16-September-2015].

Ezequiel Franco-Lara, Jan Havel, Frank Peterat, and Dirk Weuster-Botz. Model-supported optimization of phototrophic growth in a stirred-tank photobioreactor. *Biotechnology and bioengineering*, 95(6):1177–1187, 2006.