Summary

<p>Modeling played a pivotal role in our project. It helped us: (i) incorporate expert's input in designing novel metabolic engineering strategies; (ii) consolidate our understanding of the photosynthetic host and its adaptive strategies to diurnal light fluctuations; and (iii) guide our experimental design and the mimicking of industrial conditions. We extended a genome-scale metabolic reconstruction of <i>Synechocystis</i> such that it could simulate the behavior of different synthetic constructs. This model was analyzed using state-of-the-art constraint-based techniques with the goal of designing a synthetic organism in which fumarate production directly from CO₂ is aligned with microbial fitness. Mimicking of industrial light regimes in the lab was achieved by deducing sinusoidal functions coupled with algorithms that generate stochasticity resembling the one cells encounter in production scenarios. Our modelling efforts combined culminated in the first <i>in silico</i> photosynthetic cell factories that produce fumarate day and night. This design has been successfully validated in other modules of our project </p>

<h1>Aligning biomass and product formation</h1>

<p>We have been alerted by experts in biotechnology at the onset of this project that phenotypic stability can be a real hurdle in the scaling up of biotechnology processes. A clever way to stabilize a production trait is to align the anthropogenic needs (i.e. high level production) with those of the production host (i.e. fitness). Although previously attempted in chemoheterotrophs [1], and also in cyanobacteria [2], this is easier said than done. Particularly for the latter, as typically these strategies involve coupling product formation to the ability of a cell to regenerate energy and redox cofactors. The caveat for photoautotrophs is that their systems to capture the energy of photons into chemical bonds has evolved to contain a large set of safety valves for electrons - alternative electron flow pathways. Removing these <i>in silico</i> leads to stable growth coupling on the computer screen - but unfortunately, not in the lab - as these cells become very crippled and unable to deal with very small light variations [3].</p>

<p>A closer inspection of metabolic network maps reveals that, Although often neglected, anabolic pathways are in reality mined with reaction that produce side-products, in addition to the intermediates that will subsequently be (used to synthesize) biomass components. These side-products are usually recycled back via dedicated metabolic routes to guarantee an efficient use of cellular resources. In the case of photosynthetic organisms, given the large energetic requirements associated with carbon fixation, this recycling is particularly important, as it really does not pay off to allow the carbon contained in side-products of anabolism to be lost. This has been used as an argument to explain the extended anabolic versatility of photoautotrophs and their reduced number of auxotrophies (i.e. nutrient requirements for growth) [4]. But hitherto has never been exploited to develop growth-coupled production strategies.</p>

Introducing "FRUITS"

Building up on the ideas brought forward by the [2015 Amsterdam iGEM team](http://2015.igem.org/Team:Amsterdam), an algorithm has been developed to 'Find Reactions Usable In Tapping Side-products' – FRUITS. This is freely-available at https://gitlab.com/mmp-uva/fruits.git. It identifies side-products of anabolism that are suitable to be coupled to growth of the cells by deletion of their re-utilization reaction(s). A constrained (e.g. here mimicking industrial conditions) genome-scale model (GSM) of the desired production organism (e.g. here *Synechocystis*), is the only input required. Implemented as a pipeline of Python scripts, FRUITS produces (i) a list of all the modeled metabolites that can be produced in a growth-coupled fashion; (ii) the gene deletions that are required for it; and (iii) the computed maximal biomass formation rate along with the predicted minimum flux towards the identified target compounds.

FRUITS starts by evaluating the GSM for reactions responsible for the synthesis of macromolecules (e.g. proteins, DNA or RNA) and their precursors (e.g. amino acids or nucleotides). All reactions identified are then dissected into their products and substrates. A distinction for the products is drawn between biomass precursor metabolites (i.e. strictly used within the anabolic pathway) and co-produced molecules (i.e. that are not necessarily directly used as a substrate in an anabolic pathway). The list of co-produced compounds is then shortened by removing molecules that do not contain carbon (e.g. ions) or that are cofactors that act as energy carriers or redox equivalents (e.g. ATP, NADPH, NADH or CoA). The remaining metabolites constitute the list of preliminary candidates that will be further considered.

Modeling wise, a compound can only be accumulated if it can be exchanged within the boundaries of the defined system. In modeling terminology, this means that a metabolite must be a *boundary species* for which a so-called *sink reaction* is present. The metabolites in the preliminary list of candidates do not necessarily have a sink reaction, which would not allow them to ever be predicted to accumulate. This is resolved by creating a model version for each compound of the list in which, if necessary, a single new "artificial" sink reaction is added for the respective target compound. These multiple models are then further analyzed individually. Obviously, in reality it is very important to know the exact localization of where a product will accumulate. In this project particular attention has been devoted to [transport](/Export.html) as [experts alerted us](/HP/Gold_Integrated.html) to the great impact that it has on the economic feasibility of a biotechnological process as well.

The topology of the metabolic network of each of these models is then evaluated to assess the possibility to make each respective candidate compound in a strict growth-coupled way. Basically we want to find whether the re-utilization of a target compound can be completely disrupted through gene deletions; and if so, which one(s). This is done using an in-house Python implementation extended for gene deletions [5] of Optknock [6]. We then limit

the maximum number of deletions permitted according to our experimental capacity. While Optknock identifies the deletions necessary to maximize a specific objective flux, it does not guarantee that the product is uniquely coupled to an anabolic pathway. The strict stoichiometric coupling of the rates of biomass and product formation are therefore tested via Flux Variability Analysis (FVA) on the constrained model with the proposed gene deletions, while using the maximization of biomass formation as the objective function. If the exchange reaction for the target compound has a minimum flux greater than zero, then the viability of making it growth-coupled is confirmed (<i>in silico</i> of course!). The output of this FRUITS will report a list containing the identified target compounds along with the associated gene knockouts necessary and the maximal biomass and minimum target product formation rates.
During this iGEM project, our supervisors have validated this approach, showcasing it experimentally on the growth-coupled production of acetate. A manuscript describing the algorithm and its experimental validation was submitted at the end of the Summer and is currently under review [7]. In our project, we have in parallel applied FRUITS to <i>Synechocystis</i> focusing on a different target product - fumarate. </p>

<h1>'Collecting the FRUITS of <i>Synechocystis</i> without damaging the branches"</h1>

<p>FRUITS was applied to the GSM of <i>Synechocystis</i> <i>i</i>JN678 with default constraints that support a light-limited growth rate of 0.052 h⁻¹ [8]. The output of FRUITS lists nine target metabolites (Table 5.1). Based on the ratio between the production rate of each compound, and the respective maximal growth rate, the predicted product yield on biomass (Y_{p/x}) can then be calculated.</p>

<figure id='tab51'>
   <img src='images/figures/Model1_tab5.1.png' class='module-figure-image'>
       <figcaption class='module-figure-text'>Table 5.1 <i>Metabolites that can be produced in a growth coupled way based on in silico simulations during constant light.</i></figcaption>
</figure>

Metabolite	Growth rate (h^{-1})	Yield (mmol gDW^{-1})
5-Methylthioadenosine	0.052	0.007
Acetate	0.052	0.195

Mercaptopyruvate	0.034	5.702
5'-Deoxyadenosine	0.052	0.044
3,4-dihydroxy-2-butanone 4-phosphate	0.051	0.732
Adenine	0.052	0.032
Adenosine	0.052	0.032
S-Adenosyl-L-homocysteine	0.052	0.025
Fumarate	0.051	0.848

<p>An important pattern in the output of FRUITS (table N) is the inverse proportionality between the predicted maximal growth rates and the Y_{p/x} normalized over the number of carbon atoms per molecule (C-mol). To put it bluntly, the more carbon is deviated towards product formation, the slower the culture is predicted to grow. The same pattern was observed experimentally for lactate production in <i>Synechocystis</i> [9], when the flux towards lactate was allosterically increased without changing the expression level of the responsible enzyme (L-lactate dehydrogenase). This suggests that the fitness trade off at the core of the phenotypic instability that haunts biotechnology according to transport as experts and literature [10, 11], is hardwired already in the structure of the metabolic network.
All the target compounds identified by FRUITS have never been tested experimentally to be indeed possible to couple to growth. To the best of our knowledge, the only exception is acetate, which as mentioned above, transport as our supervisors have used to validate the method [7]. We convinced them to let us try to test this for fumarate within this iGEM project mainly because out of all these compounds, fumarate is the one that currently is mostly derived from oil. Luckily, they agreed! </p>

<h1>Model guided photoautotrophic growth-coupled production of fumarate. </h1>

<p>We started by performing Flux Balance Analysis (FBA) of the *Synechocystis* model *i*JN678, constrained to simulate photoautotrophic growth, and using maximization of the flux through the biomass equation as the objective function. Examination of the optimal flux distributions rapidly gives a sense of how fumarate is being produced and consumed at the same rate (hence, no accumulation!) The picture that emerges is very different from what one is accustomed when looking for instance at respiring *E. coli* or other chemoheterotrophs (i.e. fumarate is produced and consumed within the TCA cycle). In stark contrast, the flux towards fumarate in photoautotrophically cultured *Synechocystis* is coming exclusively as a side-product of anabolic reactions within both purine and urea metabolism. The fumarate that is released in the latter two is then re-assimilated through the TCA cycle (which is actually not really working as a cycle then) via the activity of a class II fumarase that converts it to malate. Fumarase is encoded by the gene *fumC* (slr0018) and, as first revealed by FRUITS, independent model simulations using a variant of *i*JN678 in which *fumC* is silenced, indeed predict production of fumarate with a $Y_{p/x}$ of 0.848 mmol gDW^{-1}.

The prediction that a *fumC* deletion results in growth coupled fumarate production was subsequently tested experimentally. How this was done and the results obtained are reported here.</p>

<h1>Conclusion</h1>
<p> The modeling framework applied here has allowed us to select a suitable product (fumarate) and devise a strategy to align its production with biomass formation. The experimental validation of the main predictions made here is an important step towards achieving one of our overall goals - " to create stable and robust photosynthetic cell factories". To make it explicit, this modeling exercise predicts that:</p>

<ul class='produce-list'>
A *fumC* deletion strain of *Synechocystis* forcibly makes fumarate </l>
 The rates of production and growth are stoichiometrically coupled (i.e. faster growth implies faster production, and vice-versa)
This results in phenotypic stability

<p>Finally, it is important to note that all these predictions have been done for cells cultured photoautotrophically in the presence of light. We did perform FRUITS also using *i*JN678 but now under darkness using glycogen consumption as the carbon and energy source (i.e. simulating chemotrophic growth). We were curious to see whether there would be something interesting related to fumarate there, or potentially, to identify other interesting targets. Unfortunately, this did not yield many new compounds besides the ones we had already picked up before, except for 3-Phospho-D-glycerate. Although the predicted $Y_{p/x}$ is very tempting (184.4 mM.gDW^{-1} - not a typo!), this comes at the cost of a drastic drop in growth rate.</p>

<figure id="tab52">
  <img src="images/figures/Model1_tab5.2.png" class="module-figure-image">
  <figcaption class="module-figure-text">Table 5.2. <i>Metabolites that can be produced in a growth coupled way based on in silico simulations during darkness relying on glycogen as carbon and energy source.</i></figcaption>
</figure>

Metabolite	Growth rate (h^{-1})	Yield (mmol gDW^{-1})
Adenosine	0.052	0.032
Mercaptopyruvate	0.034	5.702
5'-Deoxyadenosine	0.052	0.044
5-Methylthioadenosine	0.052	0.007
Adenine	0.052	0.032
S-Adenosyl-L-homocysteine	0.052	0.025
3-Phospho-D-glycerate	0.007	184.4

<p>Although, fumarate is not on the list of compounds on Table 5.2, we did not give up on trying to stabilize its production also during darkness.</p>

<H2 id=ppp>Modeling the native metabolism under nighttime conditions</H2>

<p>As discussed on our production page, being able to produce during the night as well as the day is a prerequisite for any industrially viable, cyanobacterial production strain. Towards this end, we continued to use FBA to analyze a variant of the <i>i</i>JN678 model, where the reaction corresponding to the one catalyzed by the <i>fumC</i> product is silenced (i.e. our in silico <i>Δ</i><i>fumC</i> strain). In order to get an accurate representation of nighttime metabolism, we constrained our model to simulate heterotrophic growth, where the primary substrate is glycogen. Our model predicts that even under this different trophic mode, the same Y_{p/x} of 0.848 mmol gDW⁻¹ is obtained for fumarate production. However, there is one 'small' caveat that our model does not explicitly address.</p>

<p>As a photosynthetic organism, <i>Synechocystis</i>' primary metabolism has evolved towards primarily exploit the energy of light. This does not mean that it is completely metabolically inactive at night, as it still needs to produce some energy to cover certain maintenance costs. However, compared to the day, <i>Synechocystis</i> is more or less 'asleep' during the night, metabolically speaking</p>

[12]. So although the fumarate yield is not predicted to change, due to drastically decreased overall flux through the system, the fumarate production rate is expected to be much less during the night. In order to test if our model corroborates this expectation, we turned to dynamic FBA constrained by experimental data acquired during circadian regimes. Simulating a sinusoidal light regime and using measured glycogen consumption rates [13], we were able to get a more accurate representation of fumarate production for our *in silico* <i>Δ</i><i>fumC</i> strain over a 24 hour period (**fig. 5.1**).</p>
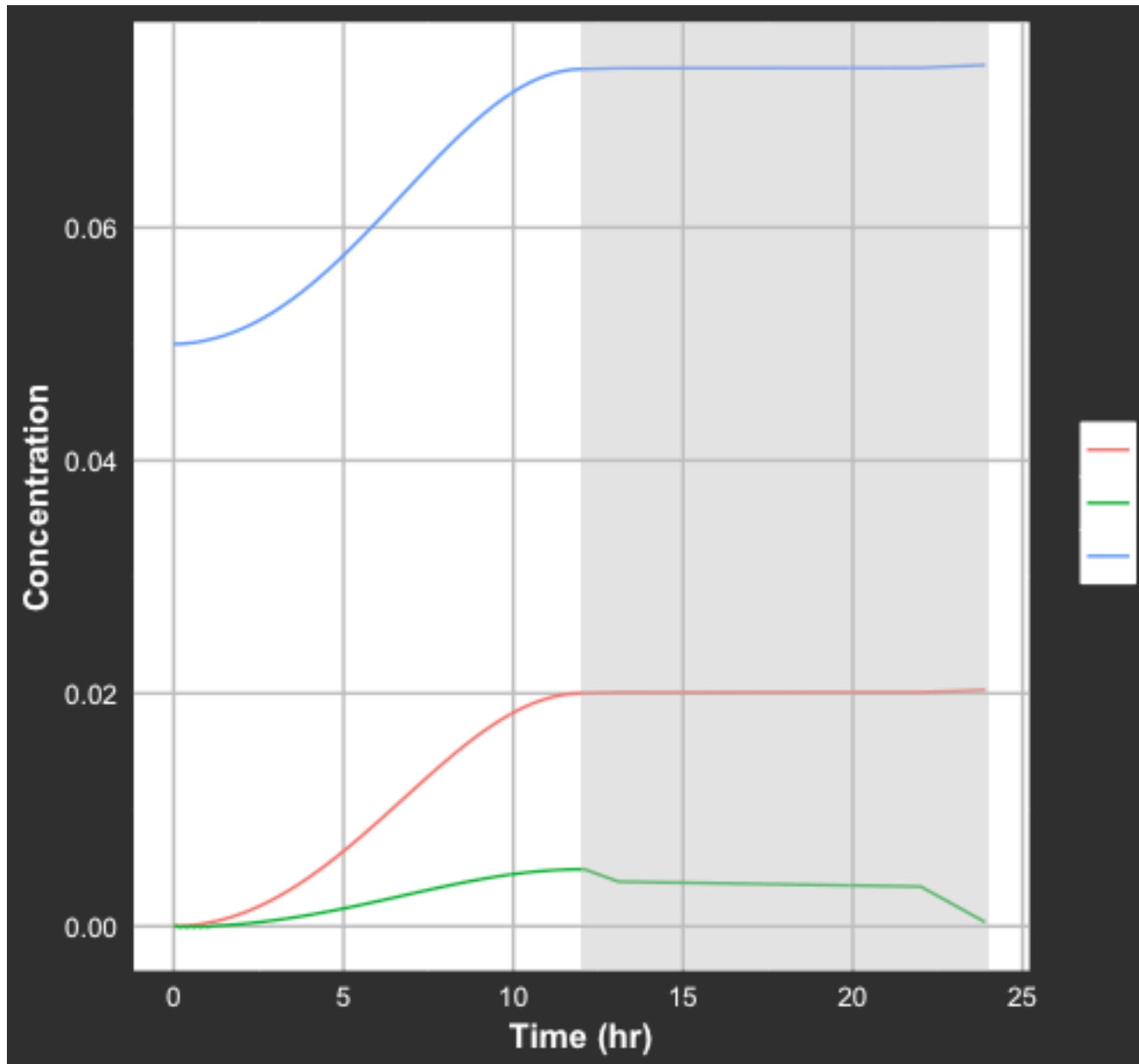
<figure id=fig51><img src=images/figures/Model3_5.1.png class=module-figure-image><figcaption class=module-figure-text>Figure 5.1. <i>**Dynamic Simulation of fumarate Production from <i>Δ</i><i>fumC</i>.** The grey shaded area signifies the night where the white background signifies the day. Fumarate and glycogen are both given in units of mmol L⁻¹, whereas biomass is in g L⁻¹.</i></figure>

As expected, when our model is fitted with experimental data, it predicts that during the night both the growth and fumarate production rates drop drastically. From this simulation, it appears that deleting the *fumC* gene is not enough for stable nighttime production. So in order to get a higher nighttime production rate, the first thing we had to ask ourselves was, why?

Examining our model under simulated night, it became apparent that the majority of the flux from glycogen was passing through the Pentose Phosphate pathway (PPP). As explained [here](/Produce.html), diverting flux away from the PPP and towards the TCA cycle could increase the nighttime fumarate production. However, since the PPP consists of many reactions involved in both day and nighttime metabolism, we needed to find a gene (or genes) which when removed would not only silence the PPP during the night, but also not affect the growth rate, and thus fumarate production, during the day. Using both literature [14] and our model, we were able to determine that the *zwf* gene was just the gene we were looking for. In fact, when silencing the reactions associated with the *zwf* gene on the *Δ**fumC* background (*Δ**zwf**Δ**fumC*) in our model during the night, the $Y_{p/x}$ on fumarate jumps up from 0.848 to 3.17 mmol gDW^{-1}. During the day, knocking out *zwf* is predicted to have no effect. This is exactly what we were looking for and so we tested it [experimentally](/Produce.html). (Spoiler alert: IT WORKS!)

Modeling beyond the native metabolism to further increase stable fumarate production

The double deletion strain *Δ**fumC**Δ**zwf* produces fumarate around the clock, which is quite an achievement, but still we wanted to try to improve it further. In this strain, during the night, the deletion of the *zwf* blocks the PPP forcing the carbon flux towards the TCA cycle. However, the *fumC* deletion suppresses the cyclic nature of the TCA. This means that the maximum stoichiometry that we could hope to achieve is to produce one fumarate per glucose catabolized. In order to achieve any further improvement, we now knew that we would have to shift somehow this stoichiometry. Only way to do that is to rewire the network. We looked into nature for inspiration.

Many organisms possess a glyoxylate shunt. This pathway, consisting of isocitrate lyase (ICL) and malate synthase (ML), will split isocitrate into glyoxylate and succinate in a first step, and then convert glyoxylate together with acetyl-CoA into malate. It creates a sort of shortcut into the TCA, in which the activity of succinate dehydrogenase could be very relevant for this project as it converts succinate to reduced co-factors (FADPH2) and fumarate. We hypothesized that a synthetic glyoxylate shunt in *Synechocystis*, would not only reconnect the broken TCA cycle of the *Δ**fumC* strain, but also increase the flux towards fumarate production, by feeding into reactions which produce electron carriers. This would potentially align fitness during the night and fumarate production. This is exactly the type of positive selection pressure, which could stabilize the expression of this heterologous pathway in a production strain.

Using our new *i*JN678 model variant, we first investigated the individual effect for each glyoxylate shunt enzyme on growth rate and fumarate production(**fig. 5.2**). This was done by forcing the flux through one of the reactions while the other was kept off (i.e. constrained to a flux of 0). We simulated both day (photoautotrophic) and night

(chemoheterotrophic) metabolism, as well as the *zwf* deletion during the night. Note that the *zwf* deletion was not simulated in conjunction with daytime metabolism as the reactions catalyzed by *zwf* are not used during the day, and hence its deletion would have no effect. Moreover, due to thermodynamic constraints, the flux through MS (J^{MS}) was only simulated in the positive direction whereas the flux through ICL (J^{ICL}) was simulated in both the positive and negative directions. </p>

<figure id=fig52>
   <img src=images/figures/Model4_5.2.png class=module-figure-image>
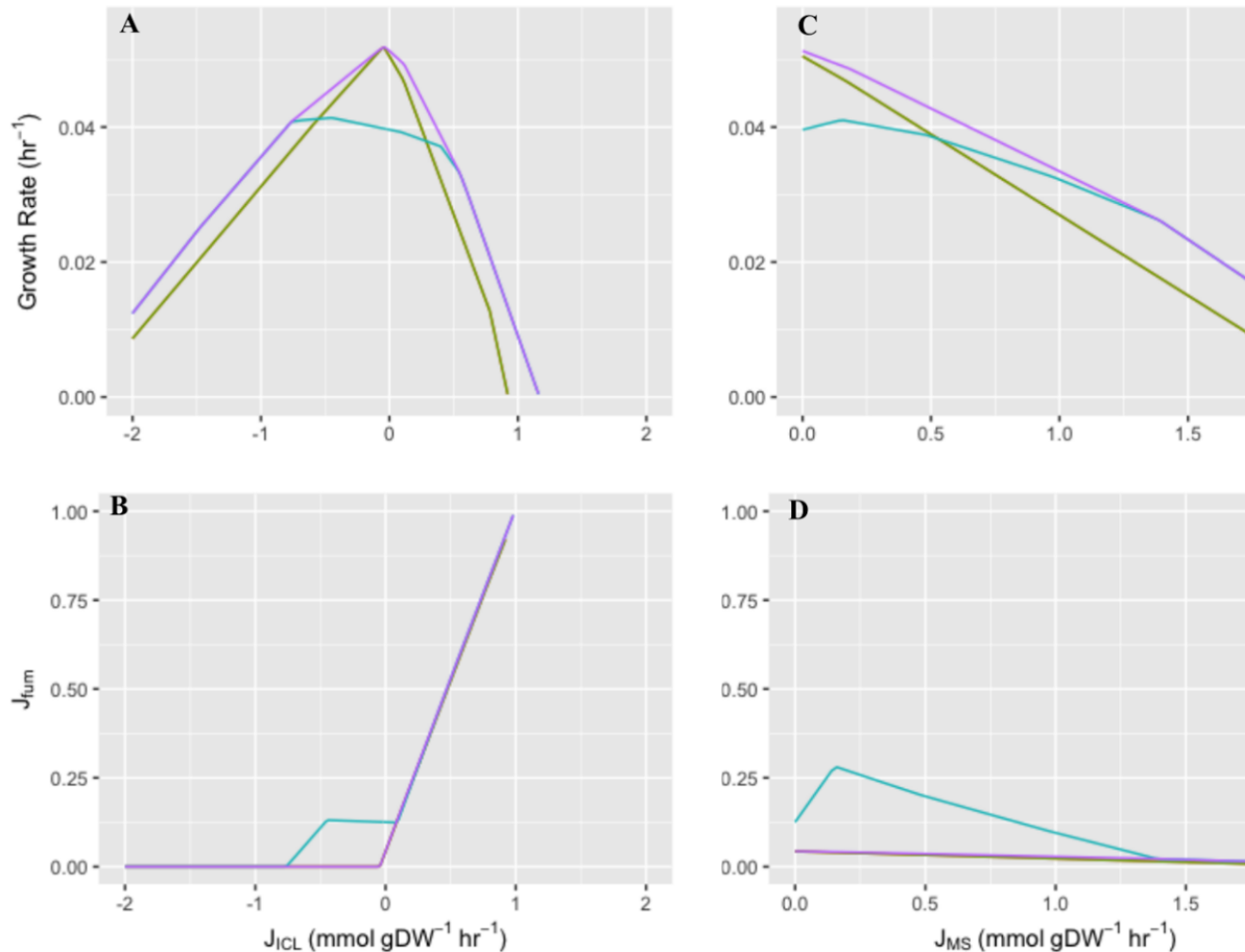
       <figcaption class=module-figure-text>Figure 5.2. <i>**The effect of either ICL or MS on growth and fumarate production.** The x axis is the value of the forced flux through the reaction as denoted by the x axis label. For plots A and C, the y axis is growth rate (hr⁻¹) and for plots B and D, the y axis is fumarate production (mmol gDW⁻¹ hr⁻¹). Three different scenarios are also shown in each graph: (green) Day, (blue) Night <i>Δ</i><i>zwf</i>, and (purple) Night.</i>
</figure>

<p>In these simulations what we are looking for at first is for peaks in growth rate. These can act as attractor states for Darwinian selection. Then for our application specific interests, these peaks will only interest us, if they also require an increase fumarate production. In other words, we want to align the microbial interests, to ours. </p>

The first thing we noticed was that the growth rate for Δ*fumC* peaks when either J^{ICL} or J^{MS} are 0 during the Day and Night scenarios (fig. **5.2A and fig. 5.2C**). This signifies that flux through either of these reactions under those scenarios is predicted to reduce growth rate, and therefore, lead to the phenotypic instability we aim to avoid. Simply put, expressing (part of) the glyoxylate shunt in Δ*fumC* is predicted to be unstable. This is a prediction that we actually decided to a href='/Produce.html#intro3' class='in-text-link' target='_blank'>try to test even though we know that it will not lead to increased stable fumarate production (if our predictions are right, of course).

In contrast the Δ*fumC*Δ*zwf* strain under a night scenario displays a peak in growth rate at non-zero fluxes for the two enzymes. Moreover, this growth rate peak coincides also with a peak in the predicted fumarate production. This is exactly what we were looking for - a glimpse of hope that if we would rewire the metabolic network we could potentially achieve higher fitness and production (fig. **5.2D**).

However, it is important to keep in mind that these simulations were done by forcing the flux through either ICL or MS one at a time, while the other reaction was off. A closer inspection of the flux distributions indicated that the network had sufficient plasticity to still find a steady-state solution in the absence of one or the other shunt enzymes. This happens because, as it is common practice in this type of modeling, the absolute flux bounds through each internal node of the system is unconstrained. Some of the solutions found seem non-physiological, since they involve drastic increases of fluxes that are typically very small in the cell.
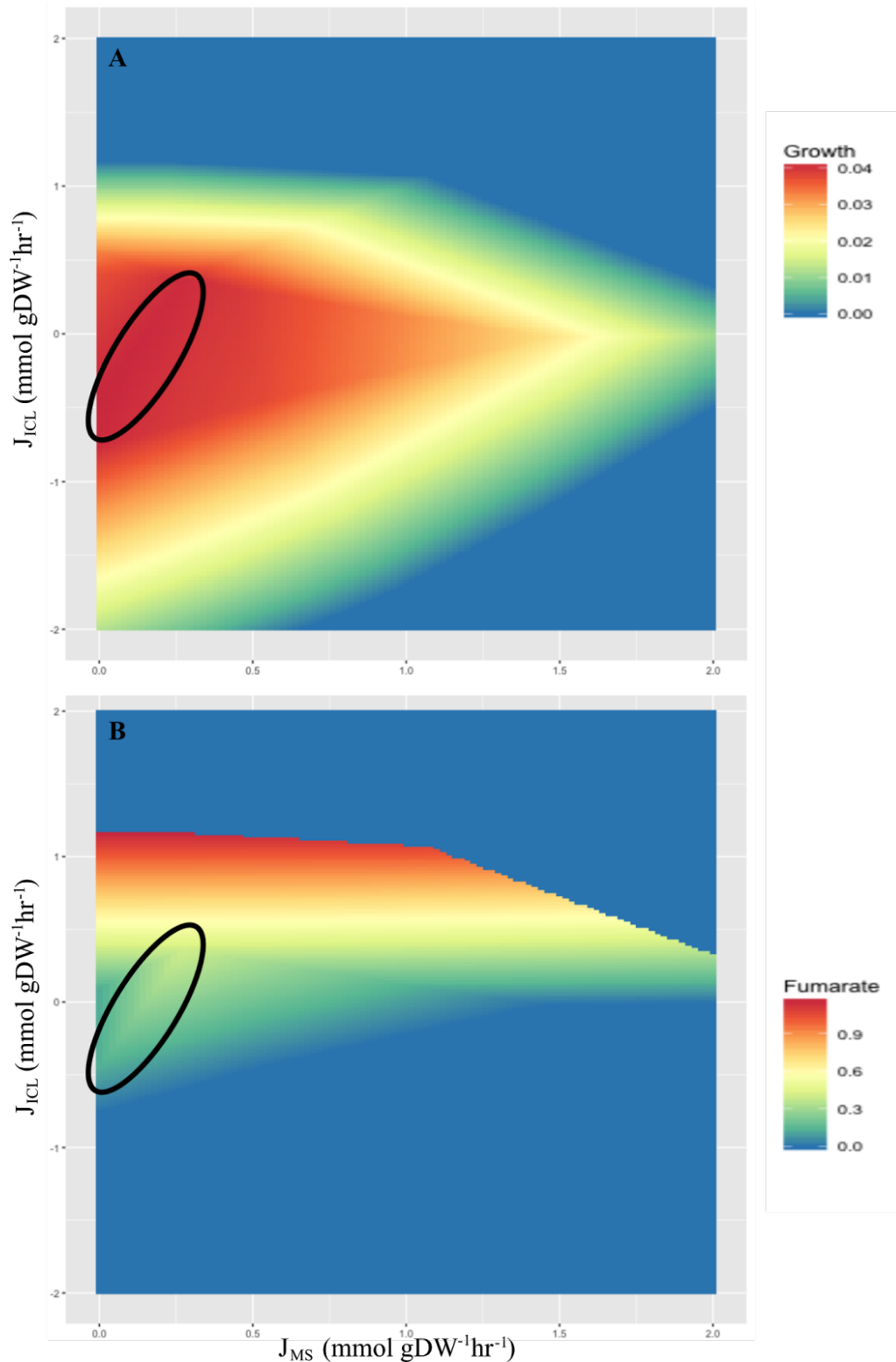
Figure 5.3. *The effect of both ICL and MS on growth and fumarate production. The x axis is flux through MS and the y axis is flux through R_ICL. The color indicates either the growth rate (A) in units hr^{-1} or fumarate production (B) in units $mmol\ gDW^{-1}\ hr^{-1}$. The circled area in both plots corresponds to the area with the highest growth rate.*

Therefore, we conducted simulations where flux is forced through both reactions at the same time (fig. 5.3). These simulations were all conducted under the Night *Δ*zwf* scenario. In plot 3A, the peak growth rate occurs where there is a small flux through MS and either a small or 0 flux through ICL, which is encircled in both plots. However, the direction of the flux through ICL within this optimal region has an effect on fumarate production. If ICL is in the reverse direction, fumarate production drops when compared to when ICL is off. On the other hand, if ICL is in the forward direction, there is a slight increase in production with increasing MS as well. One difference between 2A and 3A is that a small forward flux through ICL no longer leads to a reduced growth rate when MS also has a small flux. This implies that the detrimental effects of a forward flux through ICL when MS is off can be mitigated by a flux through MS. The takeaway from this figure is that expression of ICL may only be beneficial, in terms of fumarate production if MS is also expressed and ICL catalyzes in the forward reaction. A thermodynamic analysis of ICL under physiological conditions predicts that it will have a Gibbs Free energy (*Δ*G) of -8.4 KJ/mol. This suggests that the reversibility of ICL will most likely not be the case in such a cell *in vivo*.

We have decided to put this new insight to the test by running dFBA simulations over a 24h circadian period (fig. 5.2). Comparisons were established for model variations constrained to mimic the following genotypes: *Δ*fumC*, *Δ*fumC**Δ*zwf* and *Δ*fumC**Δ*zwf* expressing the complete shunt (*Δ*fumC**Δ*zwf*::MS::ICL). Simulations were performed as previously described for figure 5.1 and both MS and ICL were modelled as irreversible and left unconstrained, i.e. at any given point they are not required to carry a flux. The analysis of these new batches of dFBA simulations clearly indicate that, during the day there is no difference between the models. But this is only so because during day neither MS nor ICL are carrying a flux (not active). During the night, again we see that the *Δ*fumC**Δ*zwf* is beneficial compared to the *Δ*fumC* alone. But more interesting, we see that the glyoxylate shunt on top of the *Δ*fumC**Δ*zwf* further boost fumarate production.

<figure id=fig54<img src=images/figures/Model6_5.4.png class=module-figure-image><figcaption class=module-figure-text>Figure 5.4. <i>Some caption</i></figcaption></figure>
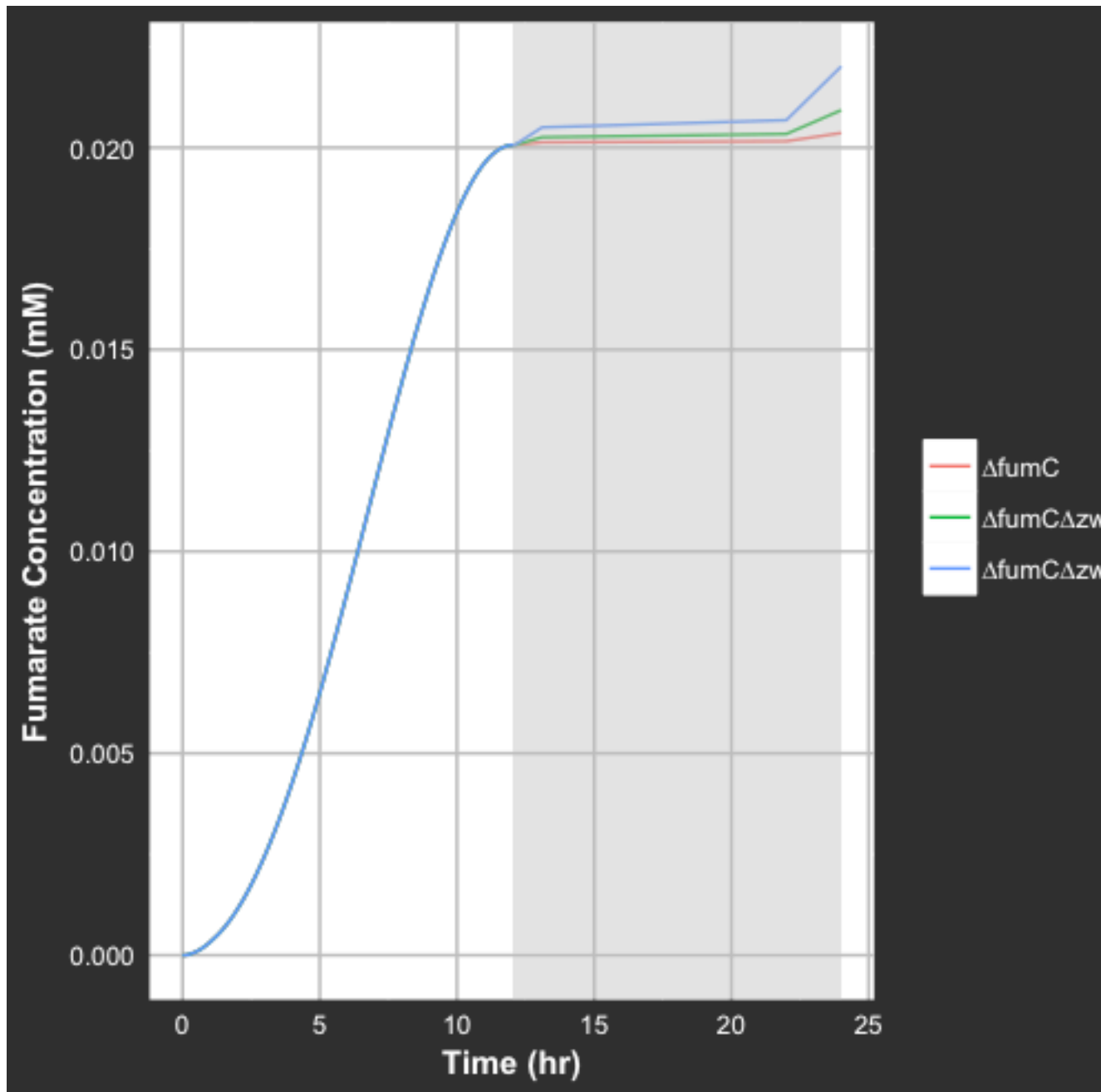
<p>When taken together, all the modeling observations here reported indicate that the conditional expression of the glyoxylate shunt over a double mutant of <i>fumC</i> and <i>zwf</i> would be beneficial. It would lead to an increase in fumarate production during the night beyond what we had achieved so far when focusing on the native metabolic network alone. And more importantly, this increase would be aligned with fitness, hence, stable. Although very promising, this <I>in silico</i> result comes with string attached… the timing has to be perfect. And this is timing of activity, not expression *per se*. So how can we time this perfectly? We dived into the exploration of the native regulatory network of an organism that has evolved to exploit circadian cycles - our very own <i>Synechocystis</i>! If there is a perfect promoter out there to control the shunt genes, chance are that <i>Synechocystis</i>

will have it. And so the search using libraries began… but about this you can read more href='/Produce.html#intro3' class='in-text-link' target='_blank'>here.   </p>

<h1>Modeling industrial conditions</h1>
<p>At the industrial scale, <i><i>Synechocystis</i></i> and other cyanobacteria are often grown in large outdoor ponds or in greenhouses [15], where natural solar radiation is the primary source of light. This means that the cultures are subject to an oscillating light-dark cycle with a period of 24 hours. Additionally, these cultures must also deal with natural fluctuations in light intensity which can be due to a variety of factors including self-shadowing of the cells and mixing.</p>

<p>Culturing photoautotrophic cells in a controlled fashion in a laboratory is typically done in lab-scale photobioreactors. At the Molecular Microbial Physiology group at the University of Amsterdam we had access to Multi-Cultivators (PSI, Czech Republic) to carry out the physiological characterization of the strains we developed in this project. There they have developed an in-house software package to control the many pieces of hardware that surround these photobioreactors (e.g. LED panel, online OD measurement, temperature, etc.). PyCultivator, as it is called, is freely available at https://gitlab.com/mmp-uva/pycultivator, and while it provides one platform for modelling and interacting with these devices, it still requires further implementations to tailor it to the purpose of the experiments - here mimicking industrial conditions in the lab.</p>

<p>In order to model the light-dark cycle, we chose a sinusoidal function as represented below:
$$
y=\alpha \sin(2 \pi \frac{t}{24} + \Delta t) + \Delta y
$$
where <i>y</i> is the light intensity ($\mu E\ s^{-2}$) at time <i>t</i> (hr). By setting the amplitude (α) and <i>Δ</i>y (y-offset) appropriately, we can simulate cycles with unequal periods of light and dark. An example of a 8hr^{-1}6hr light-dark cycle with a peak light intensity of 120 $\mu E\ s^{-2}$ is shown here. Additionally, to avoid potential errors, this function was clamped such that the minimum value allowed was 0 $\mu E\ s^{-2}$. </p>

<p>To model the effects of self-shadowing and mixing on the perceived light for single cells, we calculate the 2D light distribution of a simulated culture within a cylindrical vessel illuminated by an external light source. The simulated culture is described by three parameters: (i) the extinction coefficient of the culture, ε (L gDW^{-1} cm$_{-1}$), (ii) the biomass concentration, C (gDW L^{-1}), and (iii) the radius of the vessel, R (cm). The values for these three parameters will be used in the fluctuation experiments are listed in Table 5.3.</p>

Name (symbol)	Value
Extinction Coefficient (ε)	2.55 L gDW^{-1}cm^{-1}
Biomass Concentration	0.148 gDW L^{-1}
Vessel Radius (R)	5 cm

<figure id=tab53><img src=images/figures/Model7_tab5.3.png class=module-figure-image><figcaption class=module-figure-text>Table 5.3. <i>Parameter values for the fluctuation algorithm</i></figcaption></figure>

<p>For a specific position within the simulated vessel, given in polar coordinates, the perceived light can be calculated using the following equation [16]:

$$
I(r, \theta) = I_0 10^{-\epsilon C (\sqrt{R^2-(r \cos(\theta))^2} - r \sin(\theta))}
$$

where I_0 is the incident light intensity (μE m$_{-2}$ s$_{-1}$), and r and θ are the polar coordinates of the simulated position. Additionally, we also use the following two equations to calculate the average light intensity and the minimum light intensity, respectively, within the simulated vessel:

$$
I_{ave} = \frac{1}{\pi R^2}\int_{0}^{2\pi}\int_{0}^{R} r I_0 10^{-\epsilon C (\sqrt{R^2-(r \cos(\theta))^2} - r \sin(\theta))} dr d\theta \\
I_{min} = I_0 10^{-\epsilon C 2 R}
$$

Using these three parameters and equations, we developed an algorithm which returns a list of light intensities that correspond to realistic light fluctuations. An outline of this procedure is shown in Algorithm 1.</p>

Algorithm 1: Fluctuation Algorithm

Input:
I_{max}: the incident light intensity
N_f: number of intensities (fluctuations)
N_i: number of iterations

Output:
I_{list}: a list of intensities (fluctuations)

1 $I_{ave} \leftarrow Eq\,2; I_{min} \leftarrow Eq\,3$
2 $I_{list} \leftarrow [I_{ave}] * N_f$
3 **for** $i < N_i$ **do**
4     $r \leftarrow RandomFloat(0, R)$
5     $\theta \leftarrow RandomFloat(0, 2\pi)$
6     $I_1 \leftarrow Eq\,1(r, \theta)$
7     $x \leftarrow RandomInt(1, N_f - 1)$
8     $\Delta I \leftarrow I_{list}[x] - I_1$
9     $y \leftarrow RandomInt(1, N_f - 1)$
10    $I_2 \leftarrow I_{list}[y] - \Delta I$
11   **if** $I_{min} \leq I_2 \leq I_{max}$ **then**
12       $I_{list}[x] \leftarrow I_1$
13       $I_{list}[y] \leftarrow I_2$
14   **end**
15    $i++$
16 **end**
17 **return** I_{list}

<figure id=tab55<img src=images/figures/Model8_5.5.png class=module-figure-image><figcaption class=module-figure-text>Figure 5.5. <i>Some caption</i></figure>

<p>This algorithm was designed such that the average light intensity of all the intensities within I_{list} never changes during the procedure. This

constant average is ensured by having uniform separation between intensities and then selecting two intensities ($I_{list}[x]$, $I_{list}[y]$) in each iteration, and changing each one with the same magnitude but in opposite directions.

Since the average light intensity never changes, this means that the initial intensity with which I_{list} is seeded determines what the average will be. Therefore, as shown in lines 1 & 2, we use the calculated I_{ave} to seed I_{list}, assuring that the fluctuations correspond to the values of the parameters in Table 5.3. Note that line 2 uses python syntax whereby multiplying a list containing a single element by an integer, n, creates a list containing n copies of the original element.

Furthermore, there is no check that I_1 is within the appropriate bounds, unlike how I_2 is checked in line 11, because the value of I_1 is calculated using the formula in equation 1 given a random position within the simulated vessel. Therefore I_1 will always be within the appropriate bounds. One scenario to be aware of is when either x or y is equal to either the first or last index of I_{list}, but the other is not. When this occurs, the intensity associated with y needs to be changed by a value of either $-2\Delta I$ or $-0.5\Delta I$, depending on whether it is the boundary index or not. However, in practice, it is simpler to not let x or y equal a boundary index as this makes it easier to stitch multiple I_{list}s together throughout an entire experiment.

These two models for lighting conditions were then incorporated into the in-house software to control the Multi-Cultivators with which we used to characterize our strains. Furthermore, these two models can be used in conjunction to generate realistic industrial light conditions, exhibiting both oscillatory and fluctuating behavior. With these new tools, we are now able to characterize our strains in an industrial setting from the comfort of our own lab!

Methods

FBA was performed using the CBMPy (http://cbmpy.sourceforge. net/) package and IBM ILOG CPLEX Optimizer by IBM. To simulate the different trophic modes of *Synechocystis* between day and night lighting regimes, we changed the objective function of our model to either maximize the reaction R_BiomassAuto in the day, or R_BiomassHetero in the dark. The day regime is defined as setting the lower bound of light exchange reaction, R_EX_photon_LPAREN_e_RPAREN , to a value of -30 and the lower bound of the glycogen sink reaction, R_Sink_glycogen, set to 0. The night regime has the lower bound of the light reaction set to 0 and lower bound of the glycogen reaction set to -0.6816. This value for the glycogen reaction was chosen so as to have the same growth rate under both lighting regimes when no other reactions are perturbed to allow for ease of comparison. Note that negative values imply consumption, whereas positive is production. These definitions for the light and dark regimes hold true for all simulations, unless stated otherwise. All reactions mentioned here were already present in the original iJN678.

Dynamic FBA (dFBA) simulations were performed using an in-house script which tracks the biomass and relevant chemical concentrations via integration between FBA calls. The two integrals [2] used were:

$$X = X_{0} e^{\mu \cdot \Delta t}$$
$$C_i = C_{i0} + \frac{q_{ci}}{\mu} X_0 (1 - e^{\mu \cdot \Delta t})$$

where X is the current biomass concentration (gDW L^{-1}), X0 is the biomass concentration of the previous step, μ is the current growth rate (hr^{-1}), Δt is the time between FBA calls (hr^{-1}), Ci is the current concentration of chemical i (mmoli L^{-1}), Ci0 is the previous concentration of chemical i, and qci is the current production rate of chemical i (mmoli gDW^{-1} hr^{-1}). Furthermore, dFBA simulations were subject to a sinusoidal light regime with a period of 24 hr by constraining the light exchange reaction to the output of a sinusoidal function. To avoid simulating the production of light, the sinusoidal function was limited to a maximum value of 0. The sinusoidal function had an amplitude of -30 with neither an x nor y offset.

As for simulating growth during the night, previous work has tracked the glycogen content during a 24 hour period [13]. This data showed that there was not a constant linear consumption rate during the night, so we therefore divided the night time glycogen content into three intervals, fit each interval with a linear model and used the slope as the consumption rate for that interval. The glycogen data and calculated consumption rates are shown in **fig. 5.6**

Glycogen
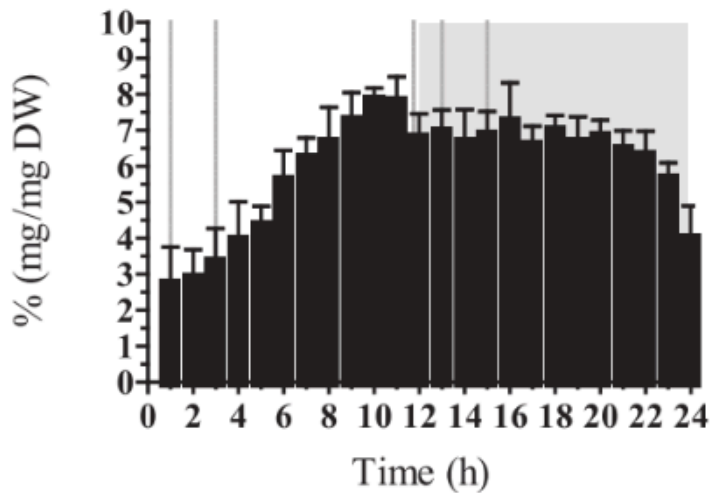
Interval	Consumption Rate (n
12-13	$-1.44 \times$
13-22	$-6.42 \times$
22-24	$-2.18 \times$

<figure id=fig56<img src=images/figures/Model9_5.6.png class=module-figure-image><figcaption class=module-figure-text>Figure 5.6. <i>The graph of glycogen content on the left was taken from Angermayr et al. 2016 [13]. Three night time consumption rates were calculated from this data for the intervals shown in the table on the rate. In order to calculate these slopes in units appropriate for dFBA, this data was first converted from %(mg mgDW −1) to mmol gDW −1. Three different linear models were then fitted to the converted data specified by the interval and the consumption rate was taken as the slope of the linear model.</i></figcaption></figure>

<p>All scripts and models used in this project are open source and freely distributed. FRUITS was conducted according to default settings. The additions to PyCultivator that we developed in this project have also been added to the repository of the Molecular Microbial Physiology group at https://gitlab.com/mmp-uva/pycultivator . </p>

References

 Feist AM, Zielinski DC, Orth JD, Schellenberger J, Herrgard MJ, Palsson BO. (2010). Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. Metabolic Engineering.12:173–86.
 Erdrich P, Knoop H, Steuer R, Klamt S. (2014) Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. Microbial Cell Factories.13:128.
 Allahverdiyeva Y, Mustila H, Ermakova M, Bersanini L, Richaud P, Ajlani G, et al. Flavodiiron proteins Flv1 and Flv3 enable cyanobacterial growth and photosynthesis under fluctuating light. Proc Natl Acad Sci U S A. 2013;110:4111–6.
 Branco dos Santos F, Du W, Hellingwerf KJ. <i>Synechocystis</i>: not just a plug-bug for CO2, but a green <i>E. coli</i>. Front Bioeng Biotechnol. 2014;2.
 Kim J, Reed JL. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. BMC Syst Biol. 2010;4:53.

 Burgard AP, Pharkya P, Maranas CD. OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. Biotechnol Bioeng. 2003;84:647–57.
 Wei Du1, Joeri A. Jongbloets1, Coco van Boxtel2, Hugo Pineda Hernández1, David Lips1, Brett G. Oliver2,3, Klaas J. Hellingwerf1 and Filipe Branco dos Santos1. (2017). Alignment of microbial fitness with engineered product formation: Obligatory coupling between acetate production and photoautotrophic growth. Under Review.
 Nogales J, Gudmundsson S, Knight EM, Palsson BO, Thiele I. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. Proc Natl Acad Sci. 2012;109:2678–83.
 Du W, Angermayr SA, Jongbloets JA, Molenaar D, Bachmann H, Hellingwerf KJ, et al. Nonhierarchical flux regulation exposes the fitness burden associated with lactate production in <i>Synechocystis</i> sp. PCC6803. ACS Synth Biol. 2016.
 Jones PR. Genetic instability in cyanobacteria - an elephant in the room? Front Bioeng Biotechnol. 2014;2:12.
 Bachmann H, Molenaar D, Branco Dos Santos F, Teusink B. Experimental evolution and the adjustment of metabolic strategies in lactic acid bacteria. FEMS Microbiol Rev. 2017.
 Rajib Saha, Deng Liu, Allison Hoynes-o Connor, Michelle Liberton, Jingjie Yu, and Maitrayee Bhattacharyya-pakrasi. "Diurnal Regulation of Cellular Processes in the Cyanobacterium". mBio 7.3 (2016), pp. 1–14.
 Angermayr, Dicle Hasdemir, Gertjan Kramer, Muzamal Iqbal, Wilmar van Grondelle, Huub C. Hoefsloot, Young Hae Choi, and Klaas J. Hellingwerf. "Culturing of Synechocystis sp. PCC6803 with N 2 /CO 2 in a diel regime shows multi-phase glycogen dynamics and low maintenance costs". Applied and Environmental Microbiology 82.14 (2016), AEM.00256–16.
 Scanlan DJ, Sundaram S, Newman J, Mann NH, Carr NG. Characterization of a zwf mutant of Synechococcus sp. strain PCC 7942. Journal of Bacteriology. 1995; 177:2550–2553.
 Ren ́e H. Wijffels, Olaf Kruse, and Klaas J. Hellingwerf. "Potential of industrial biotechnology with cyanobacteria and eukaryotic microalgae". Current Opinion in Biotechnology 24.3 (2013), pp. 405–413.
 Quinn Straub and Juan Ordonez. "A Methodology for the determination of the Light Dis- tribution Profile of a Micro-Algal Photobioreactor". In: Proceedings of the ASME 2011 5th International Conference on Energy Sustainability. Washington, DC, USA: ASME, 2011.

-
- **
A. Varma and B. O. Palsson. "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110". Applied and Environmental Microbiology 60.10 (1994), pp. 3724–3731.
**