UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR

TRABAJO FIN DE MÁSTER

Degradación de plásticos: Análisis de simulaciones de bacterias con FLYCOP

Máster Universitario en Bioinformática y Biología Computacional

Autora: del Ramo Galián, Ana

Tutores: Nogales Enrique, Juan; San Leon Granado, David

Departamento de Biología de sistemas,

Grupo de Biotecnología de sistemas, CNB

Febrero, 2021

Resumen

El plástico es uno de los materiales más usados en la sociedad actual. Se espera que tanto su producción como su uso sigan en aumento, lo que pone en el punto de mira la necesidad de encontrar métodos para su reciclaje, dado que actualmente solo se recicla un pequeño porcentaje del plástico producido. Mientras, el resto es incinerado o depositado en mares y océanos, con los consecuentes problemas que esto acarrea para la fauna y flora marina durante los cientos de años de media que tarda en descomponerse. De entre los tres métodos de reciclaje existentes – mecánico, químico y biorreciclaje – solo este último supone un reciclaje limpio y verde, puesto que no emite sustancias tóxicas ni requiere de ellas para el proceso, tan solo de microorganismos o sus enzimas. Sin embargo, es necesario buscar más opciones que promuevan el desarrollo del biorreciclaje de plásticos, dado que actualmente los ejemplos son escasos. En este trabajo se aborda el estudio computacional de tres estrategias diferentes de revalorización de ácido tereftálico (TPHA) y etilenglicol (EG), monómeros procedentes de la degradación de tereftalato de polietileno (PET) en polihidroxialcanoato (PHA), un biopolímero con propiedades plásticas producidos por la bacteria *Pseudomonas putida*. Dos de estas estrategias comprenden el uso de consorcios microbianos en lugar de un cultivo puro, pues, a pesar de que no son muy utilizados en la industria actualmente, exhiben un gran potencial para realizar tareas complejas. El uso de consorcios microbianos se basa en que en la naturaleza los microorganismos no se encuentran aislados, sino que se relacionan con el medio y con otros microorganismos, lo que influencia su comportamiento. Para el estudio de estas estrategias se ha utilizado FLYCOP, una herramienta que permite simular y optimizar consorcios bacterianos con un determinado propósito. Finalmente, en este trabajo se presenta FLYCOP como una herramienta potente para entender el metabolismo bacteriano, además de esclarecer estrategias no solo para degradar completamente un plástico, sino para producir, a su vez, un bioplástico, con todas las ventajas que suponen frente a los plásticos convencionales.

Palabras clave: Plástico, simulación, bioplástico, consorcio microbiano.

Abstract

Plastic is one of the most used materials in today's society. It is expected that both its production and consumption will keep increasing, which puts in the spotlight the need to find new methods for recycling it, since currently only a small part is recycled. Meanwhile, the rest is incinerated or abandoned in seas and oceans, with the consequent problems this entails for the marine fauna and flora during hundreds of years in average - time which plastic needs to degrade naturally. Within the three existent recycling methods – mechanical, chemical and biorecycling – only the latter means clean and green recycling, since it does not emit toxic substances nor uses them in their procedure, using for this purpose microorganisms or just their enzymes. However, it is necessary to look for more options to promote the development of biorecycling since currently, examples of this kind of procedure are scarce. Thus, three strategies are discussed in this project for revaluating terephthalate acid (TPHA) and ethylene glycol (EG), monomers from polyethylene terephthalate (PET) degradation into Polyhydroxyalkanoate (PHA), a biopolymer with plastic properties produced by the bacterium *Pseudomonas putida*. Two of the said three strategies include the use of microbial consortia rather than using a culture of a single microorganism because, although they are not widely used in industry nowadays, they exhibit great potential to perform complex tasks. The use of microbial consortia is based on the fact that, in nature, microorganisms are found in communities and establish relationships with the media as well as other microorganisms, which influences their behaviour. For the study of these strategies, FLYCOP, a tool that allows the simulation and optimization of microbial consortia with a certain goal has been used. Overall, in this project FLYCOP is presented as a potent tool to understand microbial metabolism, in addition to clarify strategies not only to completely degrade plastic, but to produce bioplastic at the same time, with all the advantages they have compared to conventional plastics.

Keywords: Plastic, simulation, bioplastic, microbial consortium.

Agradecimientos

En primer lugar, agradecer a Juan, el director de este proyecto, que me haya dado la oportunidad de realizar este proyecto a pesar de las circunstancias de la pandemia. Han sido unos meses de trabajo en los que he aprendido mucho y que me han ayudado a desarrollarme para mi futuro profesional.

Quisiera también agradecer a David, puesto que sin él este trabajo no habría sido posible. Gracias por enseñarme los conocimientos técnicos necesarios para desarrollar este trabajo, pero sobre todo ser paciente con mis dudas y resolverlas tan amablemente.

Gracias a mi familia por permitirme realizar un máster fuera de mi ciudad natal, con todo el esfuerzo que eso supone, y a Pablo por ser mi punto de apoyo y estar ahí siempre.

Index

1. Introduction

1.1. The ticking bomb of plastic wastes and its accumulation

With climate change on our heels, it is mandatory to change mankind's production model to a sustainable one. Any global development made without this initial premise will lead us to the destruction of our planet, as scientists have been warning for years.

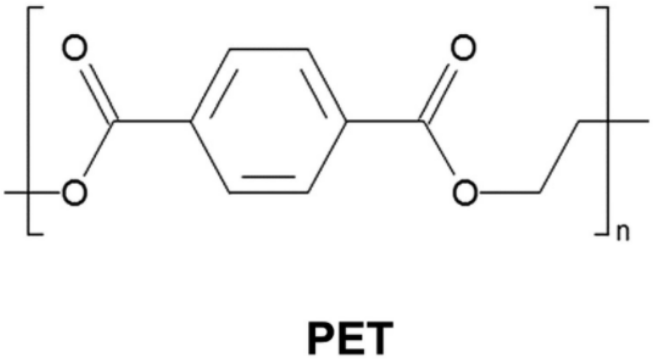
Since the plastic industry was established in the early 1900s, it quickly became a real revolution that would change the quality of our lives, as modern life is unthinkable without it. But plastic accumulation has become an environmental issue. Although its production rate in Europe has been declining since 2017, it still overcame 350 million tons in 2018. This situation reaches the need to find sustainable alternatives or a solid way to degrade this kind of material, as only 9.4 of those 350 million tons were recycled globally, which corresponds to a mere 2.6 percent of total plastic production (Plastic Europe., 2019). Furthermore, it is estimated that around 50% of plastic eventually makes its way to the beaches or even the oceans (Pathak *et al.,* 2014), where it can persist for hundreds of years, therefore becoming the world's largest plastic dump (EU, 2013). From their starting retail size, these larger pieces of plastic progressively become smaller, reaching a point when its dimension is lower than 1mm, the moment from which they're denominated 'microplastics'. This kind of plastic raises another issue. Being this small, microplastic are much harder to remove from the oceans due to the fact that they're extremely difficult to spot (Meng *et al.,* 2020). Moreover, microplastic can be potentially dangerous to marine fauna. Different studies have shown that exposure to microplastics can cause different conditions such as inflammation of the digestive system (Von Moos *et al.,* 2012), reduced nutrient uptake (Hurley *et al.,* 2017) which may lead to reduced growth (Sussarellu *et al.,* 2016).

Incineration is the go-to option for countries without enough land for depositing plastic, as Japan does. This process contributes in energy recovering but it is harmful to the environment because it produces and release to the atmosphere toxic compounds as

dioxins and polycyclic aromatic hydrocarbons, which are classified as potential carcinogens (Li *et al.,* 2001 & Samak *et al.,* 2020)

The data doesn't seem to get any better, as it's expected the global population reaches 9700 million by 2050, thus requiring an even higher plastic production, which will lead to even more plastic waste (EU, 2013).

This task is harder than it seems. We tend to talk about plastic as a single, common material, but unfortunately, that is not the case. Plastic is actually a family of materials, meaning each plastic has unique properties that vary substantially between subfamilies.



PET

Figure 1. PET molecular structure. Extracted from Wei & Zimmermann, 2017.

For this project, we are going to focus on a plastic that belongs to the thermoplastic family, polyethylene terephthalate (PET), a copolymer of terephthalate acid (TPHA) and ethylene glycol (EG) (Figure 1), whose main properties are that it is highly flexible, colourless, dimensionally stable and impact resistant. It also has variable consistency, ranging from semi-rigid to rigid, depending on its processing.

Due to these properties PET is one the most used thermoplastic, mainly as textile fibers and packaging, accounting for 8-12% of global waste. Having that into consideration, it is no surprise that PET is also one of the most recycled plastic (George & Kurian, 2014).

1.2. Recycling plastic

Among the different techniques of recycling, the mechanical way is the most used one, which consists in collecting, sorting, washing and grinding the plastics. The recycling of PET is typically the main example when talking about mechanical recycling of plastics (Ragaert *et al.,* 2017).

Another recycling technique is the chemical one, mostly used for heterogeneous and contaminated plastic wastes, consisting in converting the polymers into smaller molecules in a process called depolymerisation (Ragaert *et al.,* 2017). There are different types of

depolymerisation depending on the chemical compound used to cleave PET's chain (George & Kurian, 2014), as glycolysis, methanolysis, hydrolysis, etc. (Sinha *et al.,* 2010). However, chemically recycled materials are more expensive than virgin ones, thus reducing the applicability of this type of recycling. Furthermore, these processes generally require high temperatures and generate polluting components, such as dioxins and carbon monoxide nitrogen as a result (Marshall & Todd, 1953).

Compared to chemical recycling, the mechanical way is much simpler, but it has some disadvantages as well: the loss of colour and clarity because of the paper labels PET packages often have, or the easier degradation PET suffers if it is moist, among others. Besides, recycled PET obtained from mechanical recycling cannot be used to produce bottles or films due to its viscosity level being modified by the recycling process and other properties being also compromised after mechanical recycling. In contrast, chemical recycling may be preferred, due to the fact that PET recycled this way is accepted for food related usage (George & Kurian, 2014).

Biodegradation is also a way to recycle plastic, although not all plastics have the same reaction to this type of recycling. In fact, plastics with hydrolysable chemical bonds, such as PET or polyurethane, are more easily biodegraded than plastic without these ester bonds. However, this is a complicated process, as their insolubility in water restricts its absorption by microorganisms. Moreover, enzymatic degradation is a process that depends on the surface properties of the plastic polymers, making this process very difficult to accomplish (Wei & Zimmermann, 2017).

There aren't many examples of PET degrading enzymes in the literature, and the existing ones still don't have a yield as high as needed for industrial applications. There are different methods for increasing enzymatic activity, such as mutagenesis of the active site, introducing ions or improving thermal stability, but despite these efforts, PET degradation activities still remain low (Joo *et al.,* 2018). Moreover, the majority of PET hydrolytic enzymes described don't have a high substrate specificity, meaning they show hydrolytic activity versus a variety of substrates (Taniguchi *et al.,* 2019).

It has been observed, though, that by decreasing PET particle sizes, the hydrolysis rate increases (Gamerith *et al.,* 2017). PET monomers are linked by ester bonds, which

enables their hydrolysis by different hydrolytic enzymes (Hiraga *et al.,* 2020). The hydrolysis of PET results in the two compounds that form its backbone: TPHA and EG, which is going to be the start point of this project.

An example of PET biodegradation was accomplished by Yoshida *et al.,* 2016 when they isolated a new bacterium, *Ideonella sakaiensis*, which uses PET as its major carbon source at a moderate temperature and presenting high PET specificity. The yield observed from this degradation was higher than the one obtained by other existing enzymes. Two enzymes were discovered in this novel microorganism: PETase, which presents PET hydrolytic activity and MHETase, which belongs to the tannase family, that carries out the hydrolysis of the product of PETase, into TPHA and EG (Yoshida *et al.,* 2016). The proposed PET metabolic pathway is schematized in Figure 2.

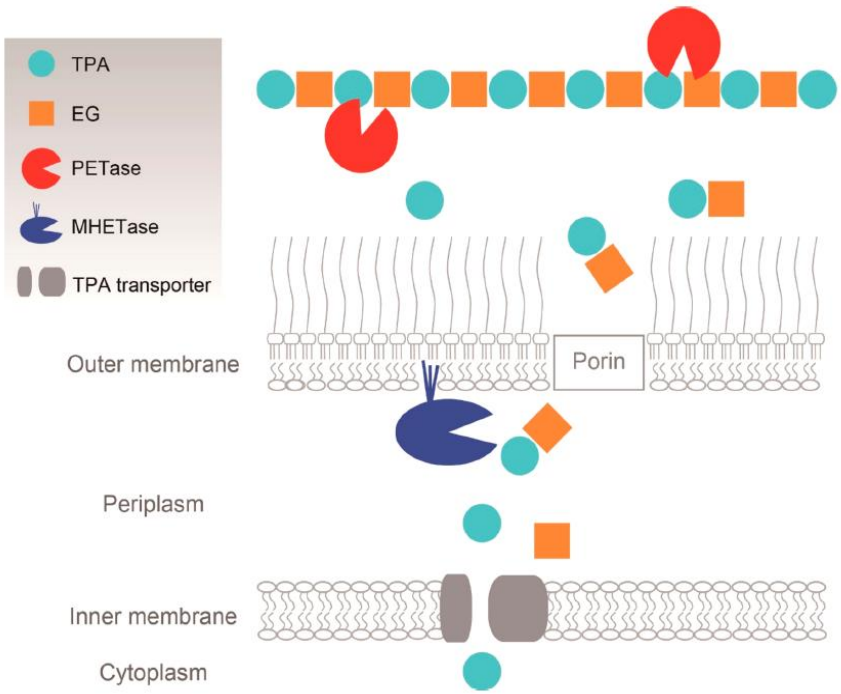


Figure 2. PET metabolic pathway. Extracted from Taniguchi et al., 2019.

Until that point, there were little known microorganisms able to do this process, being the best characterized examples from the genera *Thermobifida* (Müller *et al.,* 2005) or *Thermomonspora*, both belonging to the gram-positive phylum of *Actinobacteria* (Danso *et al*., 2019). The hydrolases discovered in those thermophilic microorganisms have a diverse substrate pool, meaning they are not specific for PET degradation and, as their

host, are thermophilic, as they need high temperatures for a correct enzymatic functionality (Wei *et al.,* 2014).

There are several advantages of developing plastics biodegradation rather than continue to recycle them using the mechanical or chemical techniques just mentioned. Biodegradation establishes a new form of recycling plastic in a sustainable and green way, which would help reduce the emissions caused by this huge industry.

Hiraga *et al.,* 2020, proposed using a consortium, in fact, the one in which they discovered *Ideonella sakaiensis,* for this purpose, as it enables to recapture energy and PET derived monomers by creating a circular process, as seen in Figure 3 (Hiraga *et al.,* 2020).

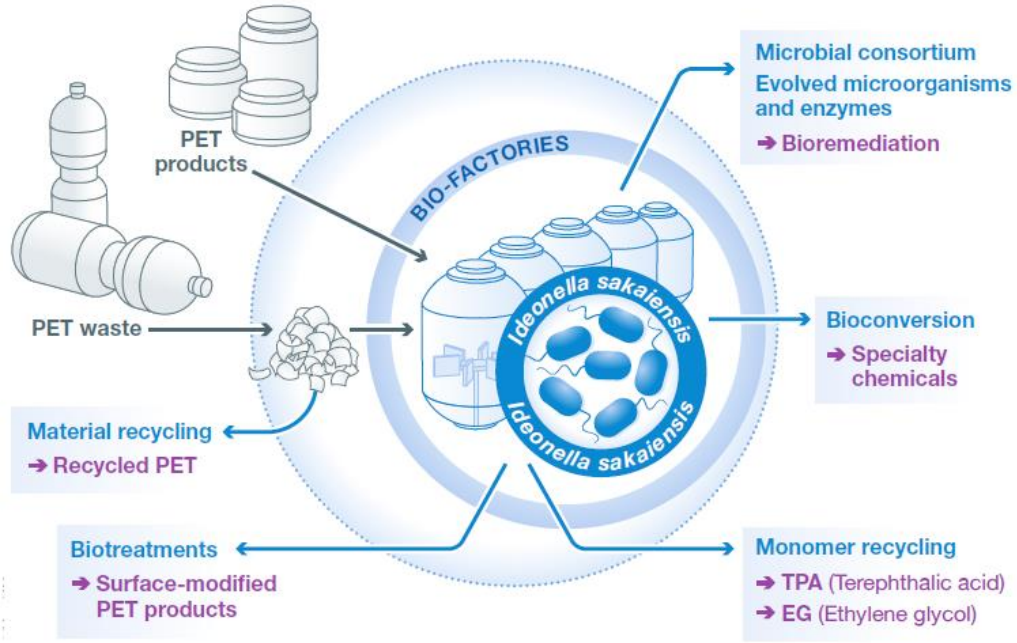


Figure 3. Scheme representing a cycle for PET recycling. Extracted from Hiraga et al., 2020.

Having found a microorganism capable of degrading PET shows that recycling plastic using microorganisms is possible. In order to find more genes that may have similar functionality, genome sequencing would be really helpful, as it can find similar sequences in other microorganisms. Moreover, knowing the 3D structure of the *Ideonella sakaiensis* enzymes can be really helpful as well, as they can be modified to improve their stability and activity (Hiraga *et al.,* 2020). Additionally, synthetic biology could be used to create genetically modified microorganisms that can carry out this function.

The current interest in recycling plastics is not only to reduce waste and pollution, but also producing another material, such as bioplastics, that have market value, which, besides, can be produced using consortia instead of pure culture, thus reducing its cost (Jiang *et al.,* 2017).

1.3. Using microbial consortia

Ideonella sakaiensis is the only microorganism found to be able to carry out PET depolymerisation, which may reflect the need of using a microorganism consortium - multiple interacting microbial populations (Brenner *et al.,* 2008) -, instead of trying to find a single microorganism that can complete this process with good results. This idea relies on the knowledge that we are unable to cultivate the vast majority of microorganisms with the current technology. Moreover, using microbial consortia may have benefits coming from the interaction between microorganisms, since it is known that microorganisms can build synergic relationships by exchanging metabolites or detecting each other's chemical signals; or they can interact negatively by competing for limiting resources, although the most common is a mixture between positive, negative and neutral interactions (Street & Bey, 2014; Keller & Surette,2006; Zomorrodi & Maranas, 2012). Many of these interactions are mediated by the constant production and uptake of metabolites by the members of the consortia, inducing continuous changes within the extracellular space, thereby modifying the signals sensed by the microorganisms, what translates into interaction between microorganisms being dependant on the state of the environment, which changes constantly (Harcombe *et al.,* 2014).

From an evolutionary point of view, as the reductionist approach of using a single microbial strain does not generally happen in nature, it is necessary to take into account the interactions between cells. In fact, to get this would require considering that microbial social interactions manage biological processes. The point is that, although community fitness is expected to increase over generations, in an evolutionary frame, cell fitness does not depend only on its genotype, but on the whole population. Interestingly, by using microbial communities, cellular stress is reduced, which enhances cell fitness, which, at the same time, helps the bioprocess yield. This effect is more noticeable when the biosynthetic

pathway is long or the expression of some gene produces cellular stress on the host (Cavaliere *et al.,* 2017, Zhang & Wang, 2016).

One of the advantages of using microbial consortia is that consortia resist better environmental fluctuations than a pure culture, which may facilitate the process of interest (Brenner *et al.,* 2008). In Figure 4, we can see the difference between using a pure culture or a consortium. If wanting to make a product P, it may take multiple steps to obtain it from the substrate S through the sequential transformation of S to some intermediate metabolites (X_1, X_2). In Figure 4.a we see this process carried out by a single population, meaning it has to make all the steps of the whole process. On the other hand, in Figure 4.b the same process is done by a consortium, where each population is given a single reaction. Using a consortium such as in Figure 4.b has several potential advantages: the number of exogenous elements in each population is limited, hence reducing the metabolic imbalance, which at the same time improves microbial growth. Also, the division of the process simplifies its optimization by enabling working on each reaction separately (Brenner *et al.,* 2008). While for mono-cultures, the expression level of the pathway is generally optimized by modulating the promoter, the ribosome binding site, varying DNA copy number, post translational balancing, etc. (Jones *et al.,* 2015); for consortia this process can be done separately, thus, optimizing each pathway module apart in its host strain, resulting in a more precise optimization process (Jawed *et al.,* 2019).

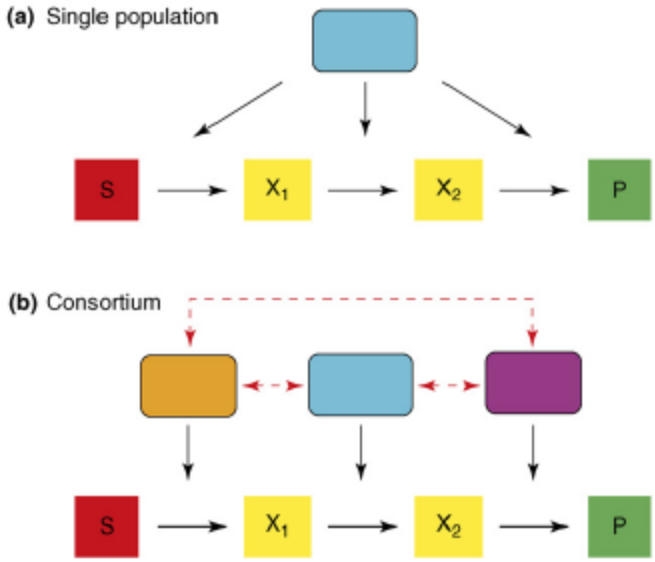


Figure 4. Schematized representation of a process carried out by (a) single microbial population and (b) a consortium. Extracted from Brenner et al., 2008.

Moreover, using microbial consortia could perform thermodynamically unfavourable reactions by isolating toxic compounds, such as enzymes or intermediates, into different cell types; or by creating concentration gradients between them. Another potential advantage may be an increased robustness from the consortium in comparison to a single strain culture, due to the fact that a biosynthetic pathway may contain enzymes with different biochemical properties that need concrete environmental conditions to be expressed or to be functional. A single strain can only provide a particular environment, which may not be suitable for the expression of all the pathway genes. In contrast, having different microbial strains offers a diversity of environments, providing every enzyme the conditions it needs to be active and functional (Brenner *et al.,* 2008; Zhang & Wang, 2016).

There are examples in which a microbial consortium was used to perform a bioprocess more efficiently than with a single microorganism species, as *Zuroff & Curtis.,* 2012*,* in which they developed a symbiotic consortia capable of producing biofuel from lignocellulose.

Although industry has typically used single strain cultures, its constant evolution, that is to say, designing ways to produce or degrade more complex compounds each time, is making it harder to use pure cultures. Using a microorganism as a host for a metabolic pathway to produce a desired product has succeeded thanks to the development of metabolic engineering. Despite these efforts, as the desired product becomes more and more complex, it is very difficult to achieve its production by a single strain. Thus, there have been large efforts toward the use of synthetic microbial consortia in order to improve the production rate, which consist in dividing the pathway in its singular components, as we have seen in Figure 4.b, and use a different microorganism to host each of those components (Zhang & Wang, 2016).

Great efforts are being made to develop commercial use of microbial consortia, such as MIXed plastics biodegradation and Upcycling using microbial communities project (MIX-UP), where this Master's Thesis partakes in. More information about the project can be found in its site. As its name suggests, this project tries to apply the approach of using

microbial consortia to the recycling of plastic, with the ultimate goal of establishing a circular (bio)-economy for plastics. To fulfil this purpose, MIX-UP, as an interdisciplinary project, makes use of different omics to try to obtain promising results that assure us a better future (MIX-UP, 2020).

Although there are several advantages in using microbial consortia over single strain cultures, they have to overcome challenges, one of them being the maintenance of the co-existence of several strains. The consortia members may not necessarily be compatible, often becoming competitors for the media nutrients, therefore resulting in an outgrowth of one of the strains at the expense of the other consortia members. Optimizing growth conditions, such as pH, temperature or oxygen level may moderately alleviate this problem. Another potential solution would be creating consortia of the same species or species that grow in similar environmental conditions (Zhang & Wang, 2016).

Creating a microbial consortium, a process usually called consortia engineering, is a major challenge, due to the large number of data that needs to be integrated -microbial metabolisms and their interactions-. Bearing this in mind, computational analyses and simulations are required to perform this task. The integration of synthetic biology with metabolic engineering is needed in order to be able to build and optimize synthetic consortia for biotechnology applications.

1.4. Computational approaches

With the growing complexity of the experiments and its respective increase in the data to be analysed, the need to use adequate tools to optimize the research process prevails. Therefore, here are presented the computational approaches used in order to perform this project.

1.4.1. Genome-scale metabolic model (GEM)

Genome-scale metabolic models computationally describe the entire metabolic genes in an organism, as well as the gene-protein-reaction relationship. They can be used to simulate and predict metabolic fluxes using different optimization techniques, such as flux balance analysis (FBA). Specifically, GEMs describe metabolic reactions based on

stoichiometry and mass balance, which are obtained from genome annotation and empirical - as the experimental - data.

Since the first GEM of *Haemophilus influenzae* was built in 1999, GEMs have played an important role for metabolic studies, as the functional part of the models is still an active part of research, yet they have become one of the most used approaches since they can be used to predict cellular behaviour (Schellenberger & Palsson, 2009).

As sequencing techniques got better through the years, more GEMs have been reconstructed, from model organisms such as *Escherichia coli* and *Saccharomyces cerevisiae* to multicellular organisms, such as plants and humans (Gu *et al.,* 2019).

Nowadays, GEMs have become a fundamental part for analysis and prediction of cellular metabolism. Moreover, there are plenty of successful examples of producing metabolites using metabolic engineered strains, such as *Lee et al., 2016* in which they produced homo-succinic acid with high yield using *in silico* genome scale metabolic flux analysis and different omics (Lee *et al.,* 2016).

1.4.2. Flux balance analysis (FBA)

FBA is a mathematical approach for analysing the flow of metabolites through a metabolic network, thus enabling the prediction of the growth rate of a microorganism, as well as the production rate of a metabolite of interest. It is commonly used for studying the metabolic flux at a particular steady state of the system (Orth *et al.,* 2010; Becker *et al.,* 2007; Mahadevan *et al.,* 2002)

The first step of FBA is to mathematically represent metabolic reactions. For that purpose, the S matrix (stoichiometric matrix) is established. The S matrix has a size of m x n; m being the number of compounds in the system, and n being the reactions of the system. The cells in each column (reactions) are the stoichiometric coefficients of the metabolites that take part of a reaction. Stoichiometric coefficients are negative if the metabolite is consumed in the reaction, and positive is the metabolite is created. If the metabolite doesn't participate in a reaction, it is represented with a zero. As every reaction only involves a few metabolites, the S matrix is mainly filled with zeros, being, therefore, a sparse matrix. It is important to note that if a metabolite exists in multiple cellular

components, it must have a row for each said component - for example, if a metabolite exists in cytosol and periplasm, there must be two rows, one for the metabolite in each compartment. These two rows will be treated as two separate metabolites, though they can be 'linked' by a transport reaction if it occurs in the system (Orth *et al.,* 2010; Becker *et al.,* 2007).

There is also a vector, v, which represents the flux of every reaction, so it has a length of n - the number of reactions of the system -. The metabolite concentration is represented by vector x, with length m.

The system of mass balance equations at steady state is:

$$Sv = 0$$

In a metabolic model true to reality, there are more reactions than metabolites ($n > m$), so, as there are more unknown variables than equations, there are many solutions to the equation system.

FBA seeks to maximize or minimize an objective function, $Z = c^T v$, which can be any combination of fluxes. c is a vector of weights that indicate how much each reaction contributes to the objective function. In case of the objective function being biomass reaction, c would be the weight each reaction contributes to growth rate. c is usually a vector of 0 except for the position corresponding to the reaction of interest. The output of FBA is a flux distribution, v, that maximizes or minimizes the objective function (Orth *et al.,* 2010).

Dynamic FBA is an extension of FBA that allows to study the transience of metabolism due to metabolic reprogramming, by incorporating rate of change of flux constraints. While FBA cannot be used to predict metabolic concentrations, it is possible to do it using dynamic FBA instead (Mahadevan *et al.,* 2002).

One of the first steps in *in silico* design of microbial consortia was performed by Hanly & Henson., 2011, when they utilized dynamic flux balance modelling to simulate and optimize co-cultures for microorganisms that consume glucose and xylose. Each flux balance model had the form of a standard linear program:

$$\min_{v_i} \mu_i = w_i^T v_i$$
$$A_i v_i = 0$$
$$v_{i,\min} \leq v_i \leq v_{i,\max}$$

where *i* represents the species, A_i is the matrix of stoichiometric coefficients (matrix S); v_i is the vector of reaction fluxes; μ_i is the growth rate and w_i is a vector of weights which represent the contribution of each flux to biomass formation. The linear problem was solved using Matlab (Mathworks, Natick, MA) and Mosek optimization toolbox (Mosek ApS, Copenhagen, Denmark) (Hanly & Henson, 2011). This represents one of the first examples of using simulations instead of laboratory work; which is needed when dealing with complex data.

An application derived from FBA is flux variability analysis (FVA) (Mahadevan & Schilling, 2003), which is used to find the maximum and minimum flux in the system while still satisfying the imposed constraints the system has, thus providing a method for studying flux distributions under non optimal growth rates (Reed & Palsson, 2004) or studying the system's flexibility or robustness (Thiele *et al.,* 2010). FVA also allows optimization for producing a metabolite of interest (Bushell *et al.,* 2006).

1.4.3. Methods for metabolic model analysis: COBRA

The COnstraint-Based Reconstruction and Analysis (COBRA) is an integrated set of software tools that allow analyses of metabolic methods *in silico*. It is applicable to any system that can be represented through chemical reactions. COBRA computes fluxes, which, as said before, are a representation of the biochemical reactions in the system.

COBRA is based on the constraints that can be used to define the states a biochemical reaction can achieve and finally, the possible phenotypes a biological system can have in that particular condition. Some of the used constraints include osmotic balance, mass conservation or thermodynamic directionality, besides transcriptome data, used to reduce the possible states.

As most of the methods rely on the principles of FBA, its central piece is the S matrix. In Figure 5.a we can see a biochemical pathway and in Figure 5.b, its representation

in the form of an S matrix. It should be noted that the last column of the S matrix shown in Figure 5 is named 'Ex_glc' since it corresponds to an exchange reaction that allows the pertinent metabolite to enter and leave the system. In addition to the S matrix, COBRA needs to establish an upper and lower bound for each reaction, as seen in Figure 5.c, which allows the flux through said reaction. Accordingly, an irreversible reaction - reactions that only occur in one direction - must have the lower bound down to zero, indicating the direction of the flux. On the other hand, reversible reactions can have a variety of numbers as their bounds, always fulfilling that the lower bound has to be numerically lower than the upper one.

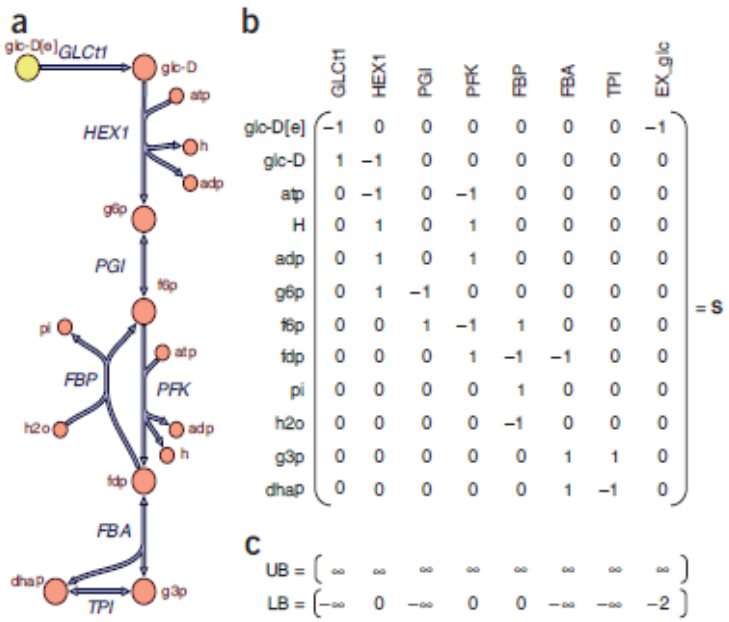


Figure 5. Scheme of biochemical pathway and its representation in a S matrix and its corresponding bounds. Extracted from Becker et al., 2007.

For the simulation to be successful, at least one reaction in the model must have a constrained bound. To sum up, the upper and lower bounds are quantitative *in silico* representation of the media conditions, which usually are experimentally determined.

A system may not have a unique solution, but a set of solutions that give insight about the biological state, called 'solution space', including all the solutions that satisfy the imposed constraints (Schellenberger & Palsson, 2009), and can guide the hypothesis

development. The solution space provides a set of so-called "equivalent solutions" that may be assessed experimentally for biological relevance, as it is needed to treat simulations as hypotheses that need experimental verification (Becker *et al.,* 2007; Schellenberger *et al.,* 2011).

The COBRA Toolbox supports models in the systems biology mark-up language (SBML) format and has a Python library, called COBRApy, which is one of the libraries used in this project (Ebrahim *et al.,* 2013). In COBRApy, the approach followed was using object-oriented programming, due to the fact that this kind of programming method is best suited for using complex biological models. Therefore, Python classes are used to represent organisms ('*Model'* class), biochemical reactions ('*Reaction'* class) and biomolecules (such as '*Metabolite'* or '*Gene'* classes). This library also contains a package whose purpose is reading and writing Systems Biology Markup Language (SMBL) models and COBRA Toolbox MATLAB structures, called '*cobra.io'*. It also includes '*cobra.flux_analysis'*, among other COBRA utilities, used to carry out FBA operations, such as FVA (available in the COBRApy method *'cobra.flux_analysis.variability'*), which can be used to identify problems in a model structure. Besides, COBRApy does not require commercial software, as COBRA does, as it depends on MATLAB.

As stated before, the main classes of COBRApy are '*Model'*, '*Metabolite'*, '*Reaction'* and '*Gene'*, and they and their relationships are schematized in Figure 6. The '*Model'* class is used as a container for a system of chemical reactions, as well as its associated metabolites and genes (Figure 6.a). Metabolites can be modified by one or many reactions, which at the same time can be spontaneous or catalysed by genes (Figure 6.b). This mechanism works as a whole and is aware of the other classes, in a way that if using a method on a specific metabolite, you can know in which reactions it participates in. This can be possible due to the '*Metabolite'* class itself containing information about its reactions. The object-oriented design offers the user the possibility to modify directly attributes of each object, while it is not possible to do the same if using COBRA Toolbox for Matlab, as the biological entities and their attributes are stored separately, in different lists, so for accessing information about in which biochemical reactions a particular

metabolite participates in, you must query several tables. And what is more, to update this information you must change it in every table (Ebrahim *et al.,* 2013).

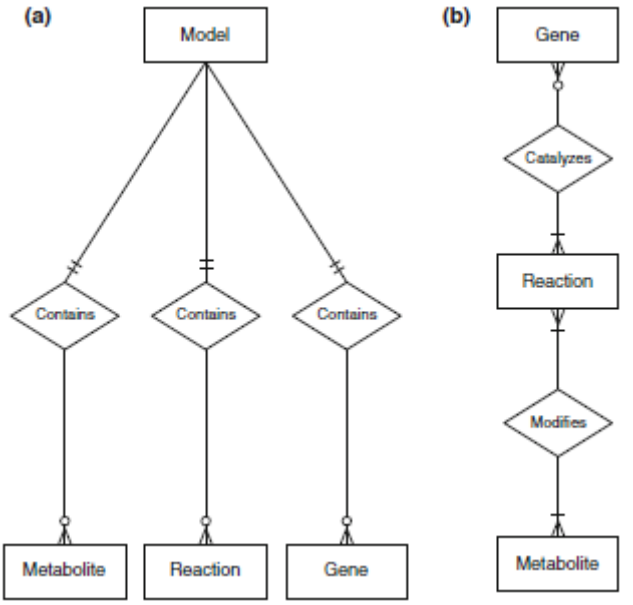


Figure 6. Schematized representation of the structure of COBRApy classes and its relationships. Extracted from Ebrahim et al., 2013.

1.4.4. Modelling approaches: COMETS

Different computational approaches exist to help designing microbial consortia. According to Perez-García *et al.* (2016), four different types of approaches have been developed regarding modelling microbial interactions in environmental processes using stoichiometric metabolic networks (SMN), a type of model where cellular biochemistry is represented as linear equations and whose dimension can go up until genome scale (Kuepfer, 2014). The first and less complex being lumped networks, where a microbial community is treated like a single organism, thus making the diversity of metabolites and reactions a unique set, for example, a single S matrix. The next approach following complexity level is compartment per guild networks (also known as multi-compartment). In this approach, exchangeable metabolites can be transferred between different microorganisms, due to the fact that each microbial unity (or guild) is modelled as a different compartment. These compartments, at the same time, are embedded inside a

bigger compartment, which represents the extracellular space, where metabolites can also stay. Hence, the extracellular space is shared between all the microbial compartments.

The third approach is called bi-level optimization, where an objective function is needed to describe the entire community, or different optimization problems for each microorganism are used. This approach is based on the assumption that universal fitness criteria that works for all the microbial does not exist, as the environment plays a key role in the physiology of microbial communities (Zomorrodi, & Maranas, 2012), thus, it uses fitness criteria at a community level. Moreover, it works by simulating successively the community in order to detect inter-community interactions. It should also be noted that it applies an optimization simulation (for example, FBA), to each compartment separately and next the output data is mined looking for the interactions, as noted before. The last and more complex approach is called dynamic-SMN, also known as hybrid SMN or dynamic FBA, although it is preferred not to use the last name as it creates the perception that intracellular flux distribution can only be achieved with FBA, which is not realistic. With dynamic-SMN, reaction rates and metabolite concentrations can be predicted for a time period, which is a huge advantage for optimizing models. Dynamic-SMN models are formed by three types of equations: kinetic, differential and stoichiometric reaction equations. Initial conditions of the simulation must be defined before starting, like the initial concentration of metabolites of the media, for example. The time interval in which the simulation is carried out is divided into discrete intervals that must also be defined. It is computationally draining, so it is best suited for communities of two to five guilds (Perez-Garcia *et al., 2016*). Every approach is focused for certain kinds of scenarios, but it is clear that dynamic-SMN is the most potent among all four, consequently offering a work frame for modelling complex conditions, as well as microbial consortia (García-Jiménez *et al.,* 2018).

One of the best hybrid SMN tools available is called Computation of Microbial Ecosystems in Time and Space (COMETS). It is able to compute ecosystem and spatiotemporal dynamics, which are based on intracellular metabolic stoichiometry. COMETS uses dynamic FBA in a mesh, which is able to track cellular dynamics in complex microbial consortia environments with genome resolution (Harcombe *et al.,* 2014).

COMETS follows two main steps, the first one being modelling cellular growth as an increase of biomass in all different spatial locations, using for this purpose a hybrid kinetic dynamic FBA algorithm. The second step consists in an approximation using finite differences of the diffusion of extracellular metabolites and of biomass expansion. COMETS has successfully been used to simulate microbial consortia, which converged into a steady state even from different starting conditions. Every microorganism is simulated based on its own objective in the corresponding environmental context and metabolite availability. As COMETS simulates in each spatial location, each microorganism may use the resources differently depending on where it is located, which is a functionality that differentiates COMETS from other approaches, as COMETS does not optimize based on group interests (Harcombe *et al.,* 2014).

1.4.5. SMAC

Sequential Model-based Algorithm Configuration (SMAC) is a tool that allows the optimization of the parameters of an algorithm or process that can be run automatically. This algorithm can be formally stated as:

"Given a parameterized algorithm *A (*the *target algorithm)*, a set or distribution of problem instances *I* and a cost metric *c*, find parameter settings of *A* that minimize *c* on *I."* (Hutter *et al.,* 2011).

This algorithm tries to find the best suited model configuration, and in this context, *c* would be the quality of the solution.

SMAC can be described as a more sophisticated Sequential Model-Based Optimization (SMBO) framework in which configurations are selected based on a model rather than randomly. SMAC models are based on random forests, a machine learning tool for regression and classification. They are similar to regression trees but have real values rather than class values.

SMAC manages to overcome SMBO limitations, allowing it to work with categorical parameters and configuring sets of problem instances. This is possible due to four technical advances: A new intensification mechanism to employ blocked comparisons between configurations, the use of random forests; being able to handle categorical parameters and

multiple instances, and an optimization procedure to select the most promising parameter configuration in a large space solution (Hutter *et al.*,2011).

1.4.6. FLYCOP

One of the main limitations for designing microbial consortia by programs using dynamic-SMN approach is that they simulate each scenario separately, thus they do not optimize the whole configuration for a consortium.

Flexible sYnthetic Consortium OPtimization (FLYCOP) (García-Jiménez *et al.,* 2018) is an open source computational framework created to automatize microbial consortia modelling and improve its understanding, as well as to solve the limitations other programs present. It uses COMETS as the computational tool to carry out the simulations due to the fact that it enables to model multiple scenarios using different configurations automatically. It also allows integration, analysis and optimization of GEMs.

FLYCOP is a flexible tool, as its own name indicates, since it can be used in a large variety of scenarios composed of different consortia. Therefore, as it can be used in a huge diversity of scenarios, it provides insight about how the consortia works and can save up time and resources, since otherwise, consortia would have to be tested in the laboratory, a process carried out by trial-and-error attempts, which could be time consuming and expensive, and, ultimately, it may not have any conclusive solution. Designing and optimizing a consortium for a biotechnological task requires the evaluation of thousands of consortium configurations and multiple criteria, leaving no other choice but to use *in silico* approaches. Hence, FLYCOP was created to alleviate this task and provide a framework for testing multiple random consortium configurations. FLYCOP uses a stochastic local search algorithm to find the best consortium configuration. Its algorithm works as follows: for a given instance of a combinatorial problem, the search for solutions takes place within the solution space, which may include partial solutions too. The local search starts by selecting a first candidate solution, then iteratively it moves from that candidate solution to another one randomly (Hoos & Stützle, 2004).

Figure 7 represents the FLYCOP algorithm. The inputs are the GEMs of the microorganisms present in the consortium, as well as certain parameters describing the

consortium configuration and their corresponding suitable values. Each consortium configuration is then evaluated depending on how that set of values affect the community behaviour, forming an iterative process in which each configuration is evaluated following the same criteria. When a certain configuration is chosen, GEMs have to be modified updating the information from that configuration, a process involving COBRA. Then, the simulation is run and evaluated when finished. The evaluation consists in calculating a quality measure, usually the fitness. The optimization algorithm is carried out by SMAC. The whole process finishes when the maximum number of cycles is reached.

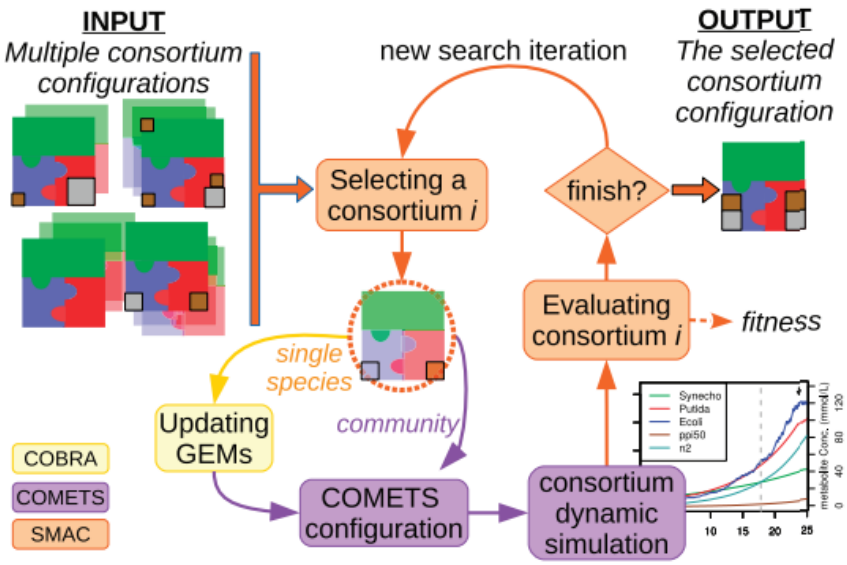


Figure 7. Schematized flowchart of FLYCOP.

The fitness function must be selected carefully, as it is a critical point on FLYCOP optimization, hence the need of choosing it meticulously. The fitness value of each configuration is calculated as the mean of several COMETS runs, due to the random nature of COMETS (García-Jiménez *et al.,* 2018).

Although FLYCOP returns a unique solution, there may be alternative solutions with high fitness values closer enough to the best one that may be worth studying. That is why FLYCOP carries out a report by data-mining all the possible solutions and offering a summary of all the evaluated configurations.

Because of FLYCOP being a general work frame, it is useful in a wide variety of purposes, such as maximizing growth rate, yield, minimizing degradation time of a certain

metabolite, etc. Another advantage is that new optimization goals can be defined besides the ones that are already established in the program, allowing FLYCOP to be adaptable to fit the experiment needs.

2. Objectives

This project has the main purpose of performing computational simulations regarding microbial metabolic behaviour for degrading PET constituent monomers (TPHA and EG) and producing Polyhydroxyalkanoate (PHA), a bioplastic; thus providing insight regarding important parameters for its correct functionality, such as initial biomass or essential media components. In order to do that, FLYCOP tool needs to be updated.

Herein, the main objective of this project is:

1) Update FLYCOP in the context of COMETSpy library and generalizing it.

2) Designing a microbial consortium providing optimal performance towards TPHA and EG revalorization into PHA for helping subsequent laboratory work.

3. Methodology and workflow

3.1. FLYCOP example script update

FLYCOP is a software comprising a variety of scripts, one of them being exclusive for the current simulation, in which the model and the parameters to evaluate are passed onto the simulation. That script is written in Python language (Van Rossum & Drake, 2009) and the first thing done to become familiarized with both FLYCOP and COMETS was updating an already existent example with the new functionalities from the COMETSpy library.

Since FLYCOP was created, a COMETS library for python was developed, COMETSpy, enabling the integration of COMETS directly into the python script, with the advantages it implies. One of these advantages is that it is programmed using object-oriented programming (OOP) instead of structured programming. Object-oriented

programming is a programming style that uses interacting objects to perform computational tasks. It is modular, because of the objects being separate entities, extensible, as objects can be added new functionality if required, and reusable, as objects can be used in other application – as in this project, where the COMETSpy library is imported into FLYCOP script, and the objects from the library are totally useable. Moreover, another benefit from using OOP is its easier maintenance, since its modular design allows to update each module without having to update the whole software.

COMETSpy, following the OOP programming style, is structured in four classes: '*Model'*, '*Layout'*, '*Params'* and '*Comets'*. This library facilitates the use of COMETS by allowing us to interact directly with it, thereby facilitating its use. In the original FLYCOP script, the metabolic model passed was in Matlab format, and it had to be converted to COMETS format manually – because its software uses its own, different format - using a previously created function. After the changes, that function no longer exists, since the model is passed in SBML format and the necessary changes are made using the '*change_bounds*' method of the '*Model'* class, that changes the bounds of a certain reaction. When all the modifications have been done, the model is saved directly in COMETS' required own format by another method called *'write_comets_model'*. This way, you can control directly what modifications are made directly within the new FLYCOP script, without the need to rely on external media.

Another advantage of the COMETSpy library is its '*Params'* class, which permits the user to interact and change the simulation parameters directly. One of the difficulties found when adapting the script was that some default parameters had changed, without warning, between the main COMETS software and the library. To fix this, the default parameters were restored from the software and introduced in the script.

COMETS also needs a layout in order to run the simulation. Said layout can be built in three different ways: by initializing the '*Layout'* class with a model; with a layout file, or initiating the '*Layout'* class empty and filling it with the data. For this project, the layout was built from a text file, as it is more complete and easier than the other ways and also allows more information to be included. The layout encapsulates the core of the experiment, as it includes the name of the model/s used in the simulation, the grid size,

the media composition and its quantities, and the initial population among several other parameters that were not of interest for this project. The media metabolites have the same name that they do in the model, and, as the model nomenclature was changed because of the SBML format, it was necessary to update the layout file and modify the corresponding metabolite names to match the model.

Then, the '*Comets*' class is initialized with the '*Layout'* and the '*Params'* classes previously configured as its arguments. This class has a main method, '*run'*, that runs the simulation.

The simulation results are stored inside the '*Comets'* class, which, again, makes working with them easier, since they are all gathered within the same object inside the script, not needing to import them from external media. Each simulation results are consulted to obtain different parameters – such as the final biomass of the microorganism, the concentration of certain metabolites or the flux of a certain reaction – needed for optimizing the simulation, the final goal of FLYCOP.

The next step is building a graph out of the results. With the old script, this process was done by calling a bash and a R script called '*plot_ biomassXN_ vs_ YmediaItem.sh*' and '*plot.biomassXN.vs.Ysubstrate.r*' respectively, where N proceeded from the number of strains in the experiment, and Y from the number of metabolites to be plotted. This was replaced in the new script with a python function (called '*make_ df_ and_ graph*') that, apart from plotting the results, also creates a data frame and a text file containing the numeric results of the parameters of interest, such as biomass or key metabolite concentration, which will be used later in the script for determining the end cycle of the simulation and the biomass yield. Then, the fitness is calculated according to the type of fitness function specified, as well as metabolite uptake if wanted regarding the experiment.

Once again, these modifications were performed in the context of a specific experiment (available here) which consisted of a long-term growth of two *E.coli* strains taking glucose, acetate and oxygen consumption as parameters, with the aim of understanding the process of the script and familiarizing with the language and COMETSpy library classes and methods to further optimize the script.

3.2. FLYCOP script generalization

The next step was to create a general script that could be used with any experiment adjusting as little parameters as possible. In the previous step, while modifying the example mentioned, every function was planned according to the experiment parameters and data. That is to say, the graph was prepared to receive four parameters (the two biomasses, glucose and acetate concentrations), and the metabolite uptake was set to the glucose, so the position number of the glucose exchange reaction in the model used was already established.

The input of this script are the values of the parameters to evaluate, the name of the fitness function chosen and how many iterations to do. The figure 8 represents the flowchart of the script, in which the step is the modification of the models according to the parameters evaluated by FLYCOP. Then, the layout is read and the parameters set. The simulation is run as many times as the parameter '*Repeat*' states, due the random nature of COMETS. The following steps are making a graph and calculating the fitness and associated parameters for each simulation, to finally calculate the average and standard deviation of those parameters. The final output of the script is the average fitness and its standard deviation.

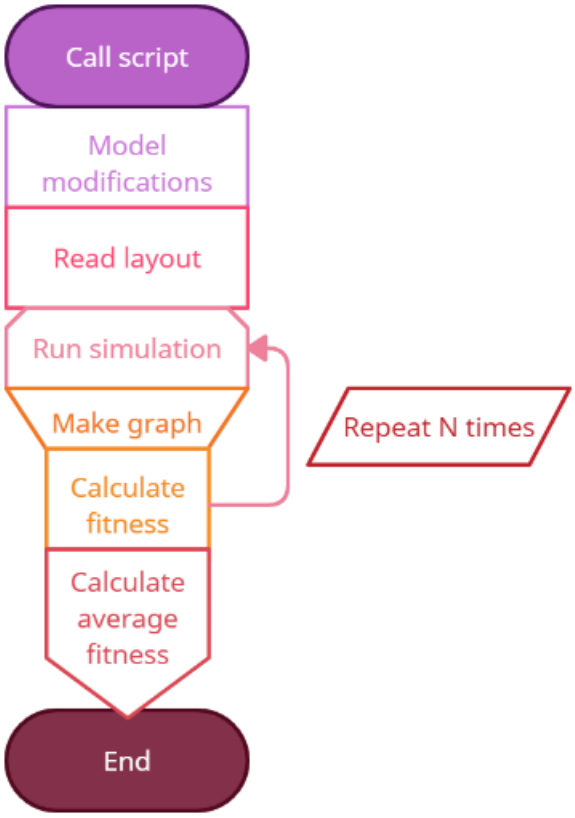


Figure 8. Flowchart of the main FLYCOP script.

These generalizations were possible by understanding the COMETSpy library structure and how to work with it. The final script accepts different numbers of models in the experiment as well as arguments for making the graphic; it is also possible to calculate the uptake of any metabolite present in the model, thanks to the fact that the function used searches the flux associated with the exchange reaction of the quoted metabolite.

When adapting the script to a certain experiment, as done later in the project, the number of parameters to configure are minimized, hence enabling a more user friendly configuration and use of FLYCOP.

3.3. Designing plastic degradation strategies

After finishing the FLYCOP script update, it is time to dive into the second part of the project: simulating different strategies surrounding plastic degradation. For that purpose, the models have to be established, as well as the layout and the script with the selection of parameters to optimize.

Three strategies were evaluated according to how they produce PHA degrading TPHA and EG for it.

The first strategy consists of a single strain which could use TPHA and EG as carbon sources, meaning there is no consortium, just a single strain culture. The second and third strategies explore the option of using a consortium instead of a single strain. Taking this into account, the second strategy consists of a default strain – that can use EG as a carbon source – and a modified strain that is able to use TPHA as a carbon source but cannot consume EG. With this strategy, the division of labour is implemented, as each strain will only consume one metabolite. Meanwhile, the third strategy includes again a default strain and a modified strain that can use both metabolites as a carbon source, but cannot end the degradation pathway of EG, hence an intermediate metabolite is expelled to the media and can be grasped by the default strain and used as well as a carbon source. Division of labour also applies in this strategy, though it goes one step further, as the steps

of the whole process are more distributed between the consortium members. In figure 9 we can see a schematized view of the three strategies.

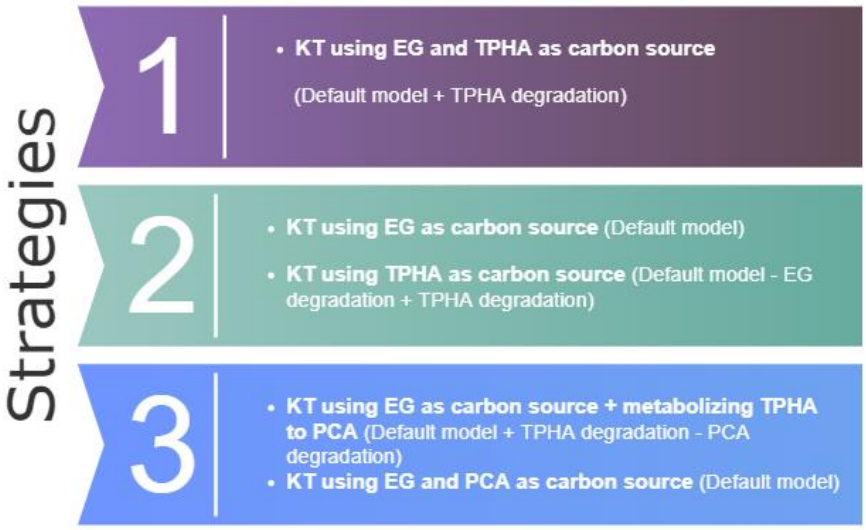


Figure 9. Summary of the three strategies evaluated in this project.

In addition to what was said above, each strategy was tested in three different oxygen conditions: without restriction, moderate oxygen restriction and severe oxygen restriction, to test how the strategy behaves under each of those scenarios.

3.4. Model modifications with COBRApy

The first step in order to perform a simulation is to create the models according to each strategy. The GEM used in this project is a modified version from the model *i*JN1411 (Nogales *et al.*, 2017) used in Manoli (2020); which consist of *P.putida* KT2440 recombinant growth-coupled PHA overproducing strain. *i*JN1411 (available here) was modified with the pathways involved in metabolism of PET. The resulting model, called *i*MM1412, was then simplified in order to reduce the computational complexity of the process. Thus, this model will be the default model stated before.

The reactions and metabolites required to enable the use of TPHA as a carbon source in the *Pseudomonas putida* GEM were introduced in the corresponding models using the COBRApy library, which enables the addition of reactions to an existing model. These reactions can be seen in Table 1.

Name	Description	Reaction
TPHAtex	Terephthalate transport via diffusion extracellular to periplasm	tpha_e ⇌ tpha_p

TPHAtpp	Terephthalate transport in via proton symport periplasm	h_p + tpha_p → h_c + tpha_c
TPHADOX	Terephthalate 1,2-dioxygenase	tpha_c + nadh_c + o2_c + h_c → dcd_c + nad_c
DCDDH	DCD dehydrogenase	nad_c + h2o_c → 34dhbz_c + nadh_c + h_c + hco3_c
EX_tpha_e	Exchange terephthalate	tpha_e ⇌

Table 1. Introduced reactions regarding tpha metabolism.

As previously mentioned, the strategy two also requires having the metabolic route of EG depleted. Table 2 shows the corresponding reactions that had their upper and lower bound changed to 0 in order to make the microorganism not able to use it.

Name	Description	Reaction
GCALDR	Glycoaldehyde reductase	gcald_c + h_c – nadph_c ⇌ glycol_c + nadp_c
EX_glycol_e	Exchange glycol	glycol_e ⇌
GLYCOLDHpp	Glycol dehydrogenase forward	glycol_p + pqq_p → gcald_p + pqqh2_p

Table 2. Depleted reactions in mutated model of strategy 2.

Furthermore, the strategy three consists in a default strain plus a strain that can use TPHA as a carbon source, thus containing the reactions stated in Table 1. Also, the reactions of the biological pathway of TPHA degradation must be depleted when Protocatechuate acid (PCA) is metabolized. In Table 3 we can see the reactions participating in this pathway that had their bounds changed to 0 as done with the reaction in Table 2 for their corresponding strategy.

Name	Description	Reaction
PYCADOX	Protocatechuate 3,4-dioxygenase	34dhbz_c + o2_c → CCbuttc_c + 2h_c

Table 3. PCA degradation reaction depleted in strategy 3.

After adding the corresponding reaction or altering them, every model had glucose removed as a carbon source by changing the bounds of its exchange reaction to 0.

Another set of modifications with the objective of maximizing PHA production was implemented, in this case, a set of reactions was deleted from all the strategies. In Annex 1 we find the names of the reactions cancelled for this purpose.

Pseudomonas putida naturally stores PHA in cytoplasm, but with COMETS it is not possible to calculate intracellular metabolite concentrations. Thus, PHA transport related reactions were introduced in the model in order to have PHA expelled to the extracellular media and computing its production by each strategy. It is necessary to specify that these reactions will not be in the future laboratory experiments as there are *in vivo* methods for calculating intracellular metabolite concentrations. The reactions introduced for simulating *in silico* are represented in Table 4.

Name	Description	Reaction
C80aPHAtpp	C80aPHA transporter periplasm	C80aPHA_c → C80aPHA_p
C80aPHAtex	C80aPHA transporter extracellular	C80aPHA_p → C80aPHA_e
EX_C80aPHA_e	C80aPHA exchange	C80aPHA_e →

Table 4. Reactions included for expelling PHA in in silico simulations.

Finally, the COBRApy method '*summary'* was used in order to analyse the model solved with FBA and check if the model presented the expected metabolites in the uptake and secretion sections. Then, by running again the '*summary'* method, we could check again the uptake and secretion of the model and, if everything looks good, the model would be saved in SBML and COMETS format directly.

3.5. Layout file preparation

In order to perform a COMETS simulation, it is needed to establish a layout. For this project, the layout was prepared in a text file later passed to the corresponding script. COMETS requires the layout to be in a specific format in order to be able to read it correctly. Certain things, such as the names of the models used or their initial population,

can be done manually, but it is best to automate the process of media composition, as there could be several hundreds of metabolites and their quantity is to be assessed too. Thus, a script using the tools provided by the COMETSpy library was created in order to automate this process. Firstly, the model in COMETS format is read, and a layout object is created using the model. This is done first because the class '*Layout'* provides an attribute, '*media'*, that lists all the metabolites present in the media, precisely the ones that must be written on the layout file, leaving out the intracellular ones. In order to fill the layout file, several lists are created grouping the metabolites by their desired media quantity, leaving the zero concentration without listing as it is the group the majority of metabolites are in. Then, the script reads the media metabolites and obtains the concentration that matches the list in which the metabolite is present, being zero if the metabolite is not in any. Finally, with the data obtained, a data frame is created with two columns, the metabolite names along with their concentration. For higher comfort while handling data, the data frame can be ordered according to the quantities, so if later in the experiment it is needed to change a concentration for simulating purposes, it would be easy as it would be on the top of the media metabolites list.

3.6. Testing and verifying the models

After finishing the creation of the models and the layout, it is necessary to test that the experiment works, that is to say, microorganisms grow and the products of interest – EG, TPHA and PHA – are consumed or synthesized by them. For this purpose, it is best to run a simple COMETS simulation and make a graph out of the result to visualize better if the microorganisms are behaving as they should. If the strains are not growing nor the products decrease or increase, there would be a need to check the models using the COBRApy method '*summary'* again, and checking all the introduced modifications in the models.

Once the test simulation turns out according to our expectations – the microorganisms grow and EG and TPHA are being consumed and PHA produced –, the model can be considered adequate and we can proceed to the next step.

3.7. Modifying FLYCOP scripts according to the parameters to optimize

As said before, a general version of the main FLYCOP script was updated, and so it needed to be slightly modified to accommodate a new experiment. For this purpose, the target parameters to optimize with FLYCOP needed to be studied and decided, in order to incorporate them into the main script as well as in SMAC.

For this project, the parameter to be evaluated and optimized is going to be the starting biomass of each of the microorganisms in each strategy. The COMETSpy class that manages this parameter is the *'Layout'*, with the '*initial_pop'* attribute. Therefore, to modify the biomass, we have to initialize the 'Layout', then change the '*initial_pop*' with the biomass value assigned by SMAC for the current FLYCOP cycle.

SMAC – the solver FLYCOP uses to optimize the fitness function – needs two files to run: the scenario and the parameters file, the latter being included in the former. In the scenario file, it is also specified the algorithm to optimize – a script in this case –. The parameters file includes the name of the parameters along with all the possible values it can take. For this experiment, the parameters file consists of a range of values the initial biomass can take, starting from 0.0001 until 0.1.

It was decided that the fitness function would be maximizing PHA production, since it is the main goal of this project. The model used has PHA production coupled to its growth, so, by maximizing PHA production, we are maximizing its growth too. Using this model, then, makes it easier to establish the fitness function – critical part of FLYCOP – by not needing to set a more complex function.

To sum up, in figure 10 we can see a flowchart indicating the path followed in this project, starting with the modification of the models to create the strains according to the experiment carried out; followed by creating the corresponding layout and testing the simulation to check everything works as expected. When satisfied on how the test simulation performs, the pertinent modifications of FLYCOP scripts can be done. Then, FLYCOP is run setting a number of cycles, N, in which it evaluates N different configurations of the selected parameters and retrieves the best solution found, while also

keeping the rest of the solutions for further studies if wanted. Noted, this solution may be a local maximum/minimum, not necessarily a global one.

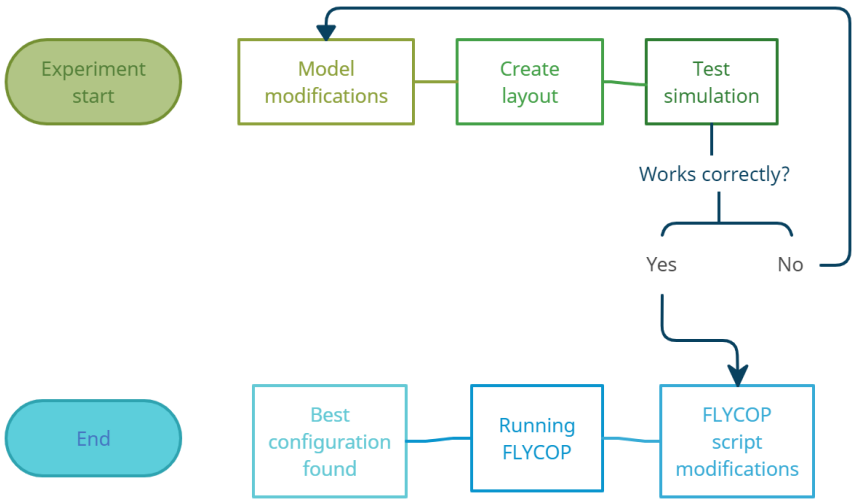


Figure 10. Flowchart of the cycle of an experiment carried out with FLYCOP.

3.8. Materials

With the purpose of using FLYCOP all the necessary programs and software are listed below, as well as the initial configuration if needed.

1. Equipment

- Personal computer with, at least, 50gb for storage and 8gb for ram.
- Linux as operative system (this project was performed under Ubuntu 20.04)
- Python programming language 3.7 or higher (this project was performed under python version 3.8)
- FLYCOP (GPL-3.0 License) (available in https://github.com/SBGlab/FLYCOP)
- SMAC software v.2.1.0
- COMETS software (Old version: 2.9.1. New version used: 2.10.0)
- Anaconda (Anaconda Software Distribution, 2020. This project was performed under version 1.7.2)
- The following python libraries: COMETSpy (v.0.4.1), COBRApy (v.0.20.0), pandas (v.1.1.4), matplotlib (v.3.3.3), numpy (v.1.19.4).

- All the scripts and data related to the plastic simulations can be found here.

2. Configuration
    1. Install Linux on the computer
    2. Install Anaconda
    3. Create a Conda environment for the project
    4. Install Python and its libraries using Anaconda
    5. Install SMAC and COMETS
    6. Install FLYCOP
    7. Launch FLYCOP or create an experiment following the methodology presented in this project

4. Results

4.1. Time saved with FLYCOP update

One of the reasons the update was made was to check out if the process time could be reduced, as it can be slow when testing over a hundred cycles, and finding the best solution may take many cycles.

Running FLYCOP					
Old version			**Updated Version**		
Nº runs	Time (s)	Memory (MB)	**Nº runs**	Time (s)	Memory (MB)
10	704.48	20	**10**	696	17
25	1743.32	43	**25**	1572	24
50	3673	37	**50**	3164	46
100	6614	39	**100**	6175	72

Table 5. Time and memory from a variety of runs performed under the old and updated versions of FLYCOP

In Table 5, we can see the time FLYCOP took to complete some sets of cycles with the first example – in which the update was firstly based on (available here) – before and after the update.

All the runs were performed under the same computer, under the same conditions and with all unnecessary processes closed.

As we can see, although with few cycles the time does not improve as much, the more cycles are performed the more time saves the updated version, up to almost 8 minutes, showing that the update improves this aspect. Time can be further optimized in a future where the whole process is done entirely in Python, without requiring R or bash.

Regarding memory usage, the FLYCOP update seems to use less memory when performing a low number of runs, but increases as the cycles grow. Nevertheless, memory used is still low, so its increase is not a problem.

4.2. Strategy 1: single strain simulations

The results obtained in Strategy one (Figure 11) shows that it is, in fact, possible to degrade TPHA and EG and produce bioplastic PHA using *Pseudomonas putida* KT2440. In this case it is noted that, when increasing oxygen limitation, the process becomes progressively slower, and with a severe oxygen limitation, the biomass needed to maximize the production of PHA is increased. Another interesting thing is that EG is consumed first in every scenario, and TPHA is only consumed when EG is exhausted.

Thereby, for the first scenario, with no oxygen limitation (Oxygen uptake set to 18.5), FLYCOP returned that the best configuration was 0.025g/L for the biomass, taking into account that uptakes of EG and TPHA are fix to 10 and 5 respectively. In this first scenario with such conditions a PHA production of 3.34mM is reached.

For the second scenario, a moderate oxygen limitation (Oxygen uptake set to 15), the configuration returned by FLYCOP is the same as in the previous scenario, 0.025g/L for the biomass, and the quantity of PHA produced is, again, 3.34mM. The difference between these two scenarios is the time they take, as the second scenario takes a bit longer to complete the process than the first one, when there was no oxygen limitation.

Finally, for the third scenario, with a severe oxygen limitation (Oxygen uptake set to 10), the optimum biomass according to FLYCOP is 0.1g/L, the maximum the starting biomass can take. Although the PHA production is almost the same, reaching 3.33mM. Overall, the difference between the third scenario and the other two is that when the oxygen limitation is severe, it is needed more starting biomass to maximize the production of PHA, though the latter does not vary much among the three scenarios.

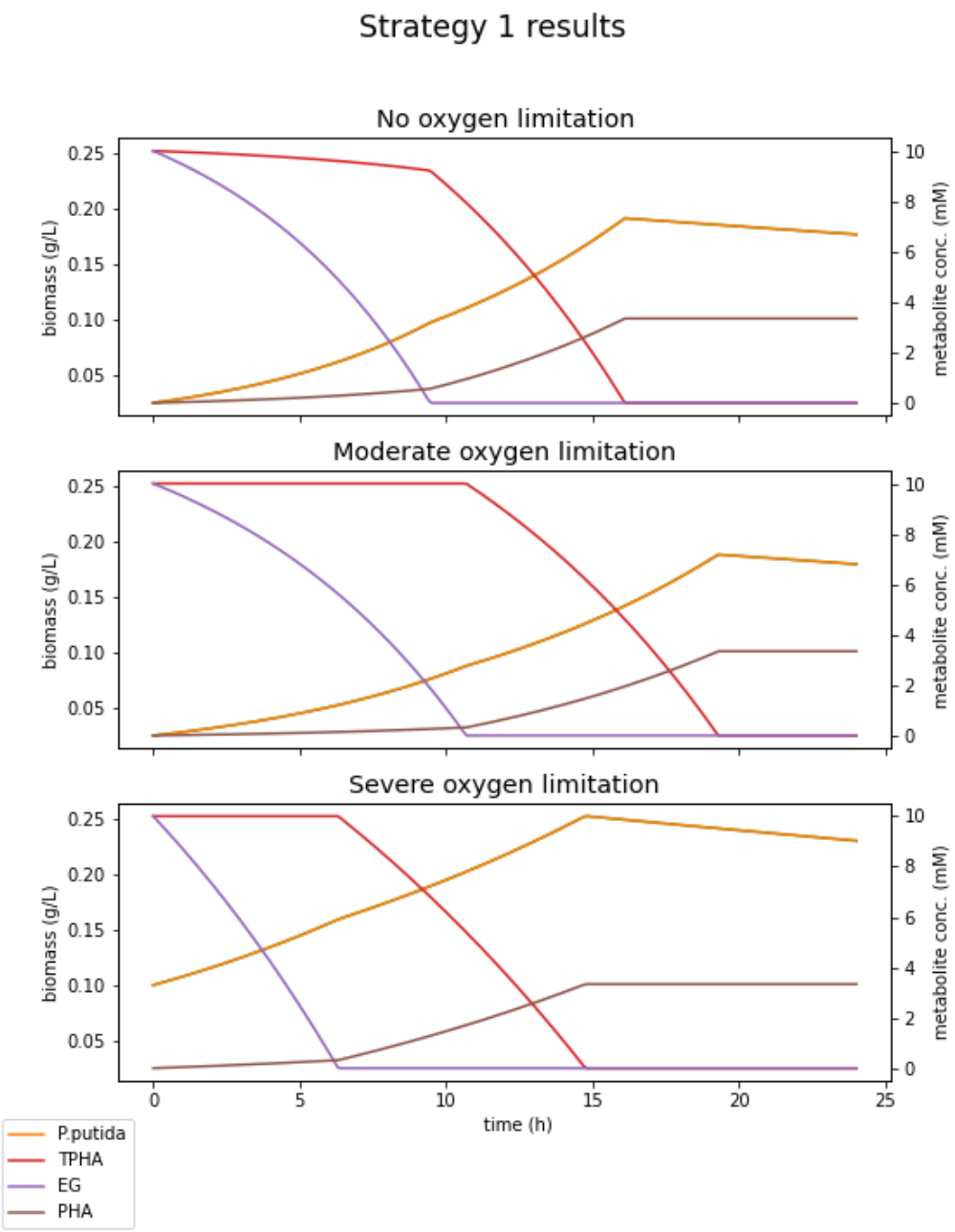


Figure 11. Results obtained using FLYCOP for strategy 1. It is shown a graph for each scenario - no oxygen limitation, moderate limitation or severe limitation -.

Regarding the yield – understood as g/L of PHA monomers produced for gram of gained biomass – it is obtained that the best scenario would be the third one, with severe oxygen limitation, as the yield obtained is 1.74, compared with 1.59 for the non-oxygen limitation scenario, and 1.62 for the second scenario.

4.3. Strategy 2: consortium simulations

Regarding strategy 2 (Figure 12, Annex 2), in contrast to the observed in the previous strategy, TPHA was consumed before EG in all the scenarios tested. Although they start to be consumed at the same time, TPHA is exhausted at a higher rate than EG does, likely due the higher starting biomass of TPHA-consuming strain in the best configuration of all the scenarios. Another interesting result was that the majority of PHA produced comes from the TPHA-consuming strain. This observation applies for all the scenarios, regardless of whether there is oxygen limitation or not. This difference cannot be attributed only to the different number of carbons between the two metabolites, – TPHA has eight carbons while EG has two –. In fact, the PHA produced from TPHA is more than four time higher than the produced using EG, thus PHA production is significantly favoured from TPHA under this strategy, likely due to the metabolic network structure of *P. putida*.

Focusing on the numbers obtained, for the first scenario, without oxygen limitation (Oxygen uptake set to 18.5), FLYCOP returned the optimal biomass as 0.05g/L for the TPHA strain and 0.01g/L for the EG strain. This configuration produces 3.32mM of PHA.

This exact results for the biomass were also obtained for the second scenario, where the availability of oxygen is slightly reduced (Oxygen uptake set to 15), only changing, in comparison with the first scenario, the time taken to reach the maximum PHA produced and increase a bit the PHA production, reaching 3.34mM.

For the third scenario, with a severe oxygen limitation (Oxygen uptake set to 10), we find a similar situation as in the previous scenario, in the sense that biomass needs to increase in order to reach the maximum PHA produced. This time, the best biomass for this goal according to FLYCOP are 0.1g/L for the TPHA strain and 0.03g/L for the EG strain. This configuration, under this scenario yields 3.34mM of PHA, the same value as the last scenario. Here, the time is still increased even though the biomass was increased from the optimal values obtained for the preceding scenarios.

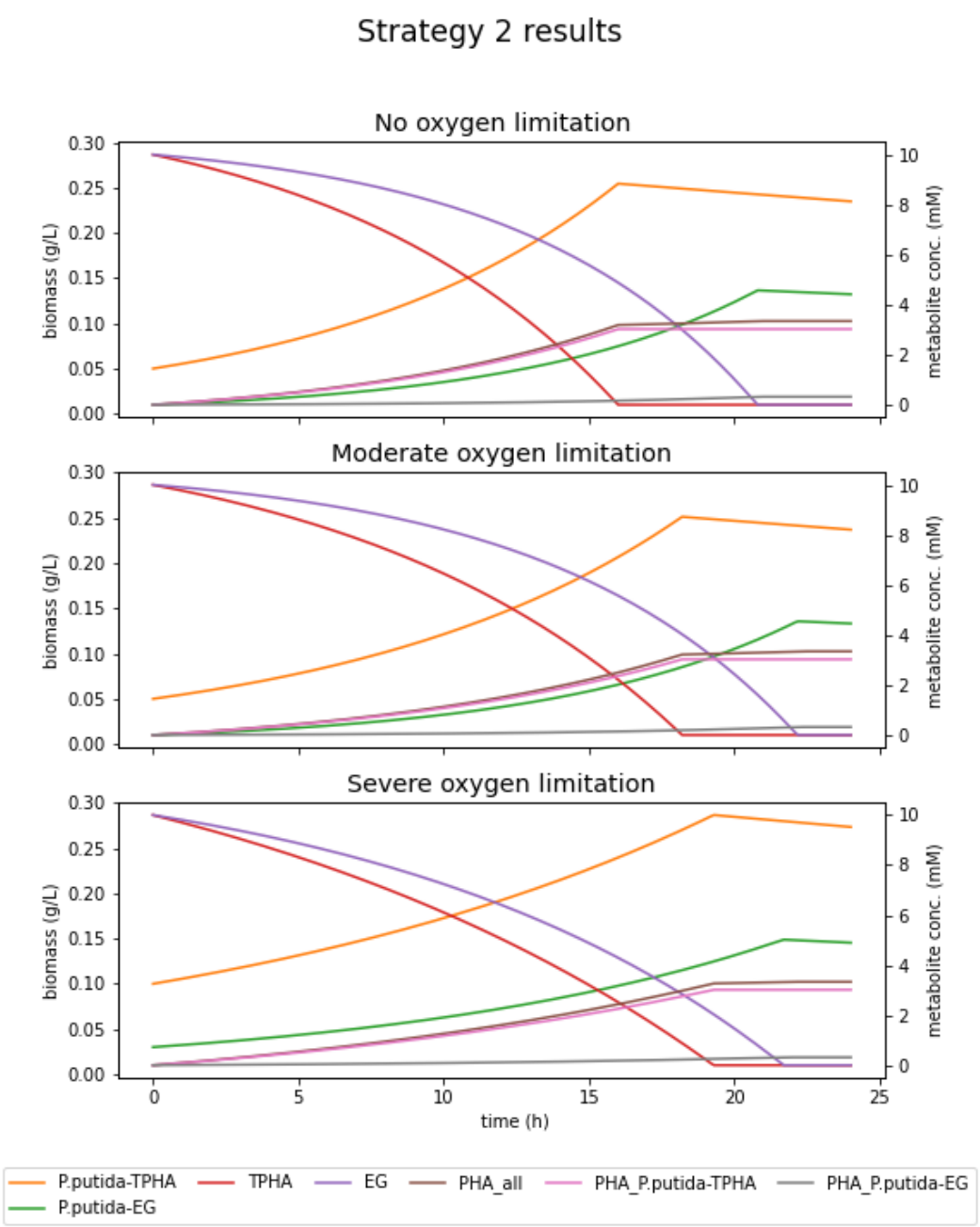


Figure 12. Results obtained using FLYCOP for strategy 2. It is shown a graph for each scenario - no oxygen limitation, moderate limitation or severe limitation -.

In terms of yield, with this strategy it is obtained a yield of 1.6 for the scenario in which there is no limitations; 1.62 for the scenario with moderate oxygen limitation and finally, 1.73 for the scenario with severe oxygen limitation. Again we find two similar yields for the first two scenarios and a higher yield for the more restricted one, resulting from the biomass not growing as much as in the other scenarios, and therefore can allocate more carbons to produce PHA.

4.4. Strategy 3: consortium simulations

Strategy 3 accounted for the consortia with a more complex division of labour. A strain consumed EG and metabolized TPHA into PCA while the second strain consumed EG and the PCA being produced. Overall we found that the optimal initial biomasses retrieved by FLYCOP have increased with respect to the two previous scenarios (Figure 13, Annex 3 and 4).

For the first scenario, without oxygen limitation, FLYCOP returned optimal initial biomasses of 0.1g/L in both strains – the largest value allowed. This configuration is obtained for the three scenarios, suggesting this strategy needs higher biomass amount to maximize the production of PHA. The amount of PHA produced for the three scenarios were 3.47mM for the first one and 3.41 for both scenarios with oxygen limitation.

We found in all scenarios that EG was consumed at a higher rate than TPHA did, likely due to the fact that EG was consumed by the two in silico strains. When EG became exhausted, the EG and TPHA-consuming strain stopped growth meanwhile continued producing PCA from TPHA.

Overall this strategy is on the quicker side except for the scenario in which oxygen is severely limited. The main difference being the speed at which TPHA is consumed and, as seen in Annex 3, just like in strategy 2, the main PHA producer is the strain consuming PCA. In this case, the metabolite obtained from TPHA metabolism. At the same time TPHA is slowed down, so is the production of PHA.

We can also see that, as soon as PCA is produced, it is captured by the EG and PCA-consuming strain and transformed into PHA, not allowing it to accumulate in the media.

In absolute numbers, the strategy 3 manages to produce a greater quantity of PHA, suggesting that the use of consortia can be helpful when dealing with complex metabolites. Having said that, in terms of yield, this strategy obtains a better result than the other scenarios. The yields obtained in this strategy are 1.79 g/L of PHA monomers in the scenario with no oxygen restrictions, 1.82 in the moderate restricted one and 1.96g/L of PHA monomers in the more restricted scenario. We can see that the pattern of yield we

had seen in the two previous strategies is maintained: the yields of the first two scenarios are still very similar to each other while the yield obtained in the scenario with the severe limited oxygen is greater than them. This could be due to the fact that growth is restricted when oxygen is, so more carbons are assigned to PHA production.

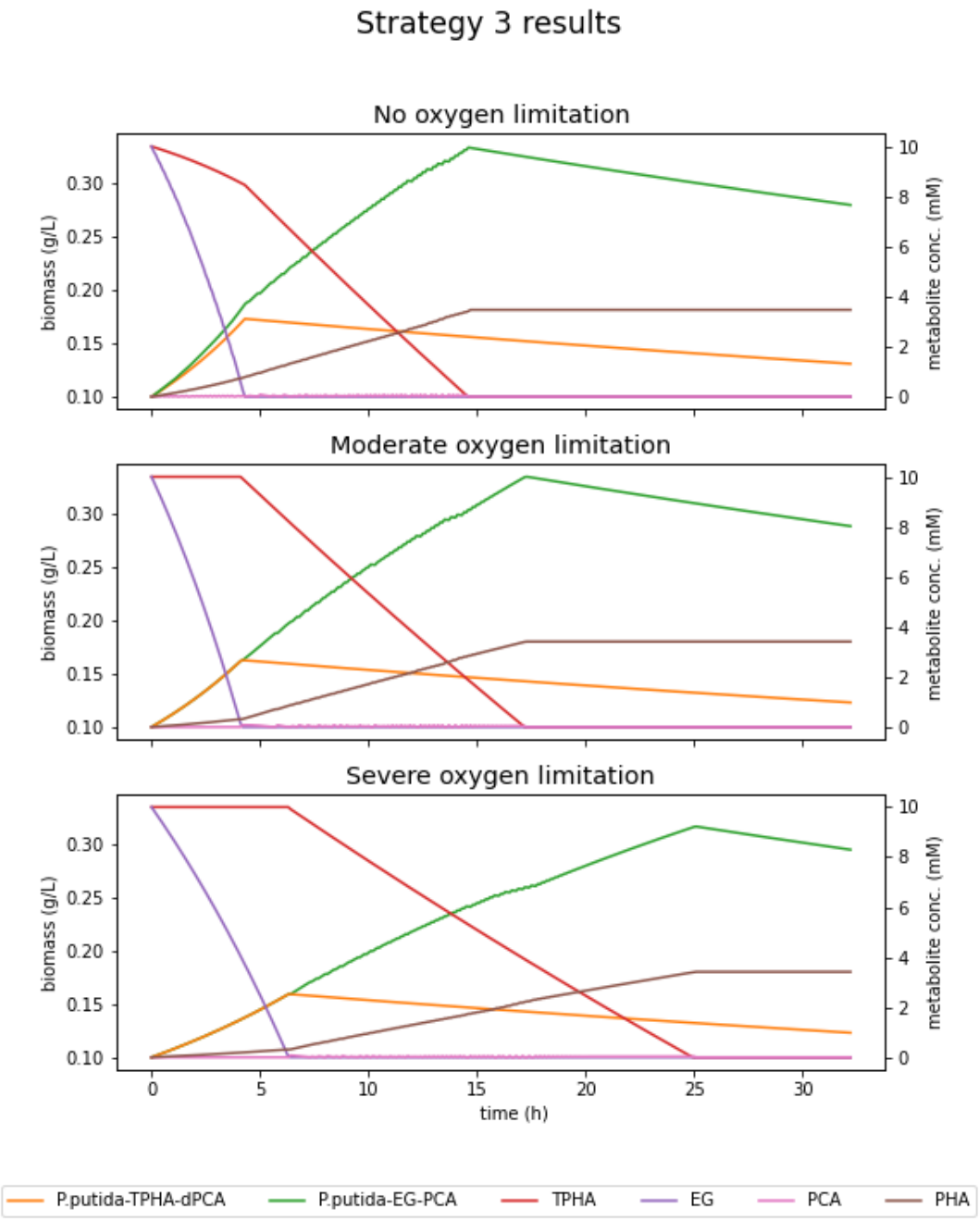


Figure 13. Results obtained using FLYCOP for strategy 3. It is shown a graph for each scenario - no oxygen limitation, moderate limitation or severe limitation -.

4.5. Overall results

To put in context all the results obtained and to be able to visualize more clearly the differences between the strategies tested in this project, in table 6 and 7 we can see all the best configurations retrieved by FLYCOP optimization as well as their results regarding PHA production.

FLYCOP opmitized biomass					
Strategy 1					
Severe oxygen restriction		Moderate oxygen restriction		No oxygen restriction	
0.1		0.025		0.025	
Strategy 2					
Severe oxygen restriction		Moderate oxygen restriction		No oxygen restriction	
P.putida-TPHA	*P.Putida- EG*	*P.putida-TPHA*	*P.Putida- EG*	*P.putida-TPHA*	*P.Putida- EG*
0.1	0.03	0.05	0.01	0.05	0.01
Strategy 3					
Severe oxygen restriction		Moderate oxygen restriction		No oxygen restriction	
P.putida-dPCA	*P.Putida- EG*	*P.putida-dPCA*	*P.Putida- EG*	*P.putida-dPCA*	*P.Putida- EG*
0.1	0.1	0.1	0.1	0.1	0.1

Table 6. Optimized biomass obtained for each strategy using FLYCOP. All values shown are in g/L

For the first two strategies, the biomass increases along the reduction of the oxygen availability, but this does not happen in the third strategy, as the biomass retrieved for all the scenarios is the same.

PHA production (mMol)								
Strategy 1								
Severe oxygen restriction			Moderate oxygen restriction			No oxygen restriction		
3.33			3.34			3.34		
Strategy 2								
Severe oxygen restriction			Moderate oxygen restriction			No oxygen restriction		
P.putida-TPHA	*P.Putida-EG*	*Total*	*P.putida-TPHA*	*P.Putida-EG*	*Total*	*P.putida-TPHA*	*P.Putida-EG*	*Total*
3.01	0.31	3.32	3.02	0.32	3.34	3.02	0.32	3.34
Strategy 3								
Severe oxygen restriction			Moderate oxygen restriction			No oxygen restriction		
P.putida-dPCA	*P.Putida-EG*	*Total*	*P.putida-dPCA*	*P.Putida-EG*	*Total*	*P.putida-dPCA*	*P.Putida-EG*	*Total*
0.16	3.25	3.41	0.16	3.25	3.41	0.16	3.3	3.47

Table 7. PHA obtained (mMol) with the different strategies tested.

It is interesting, though, that the vast majority of PHA production falls on the TPHA-consuming – in case of strategy 2 – or the PCA-consuming – in case of strategy 3 – strain. In both cases, PHA is mainly produced derived from TPHA metabolism, while EG metabolism only contributes around 10% in strategy 2 and 5% in strategy 3. Therefore, bearing these numbers in mind, the strategy 3 would be the most PHA producing strategy.

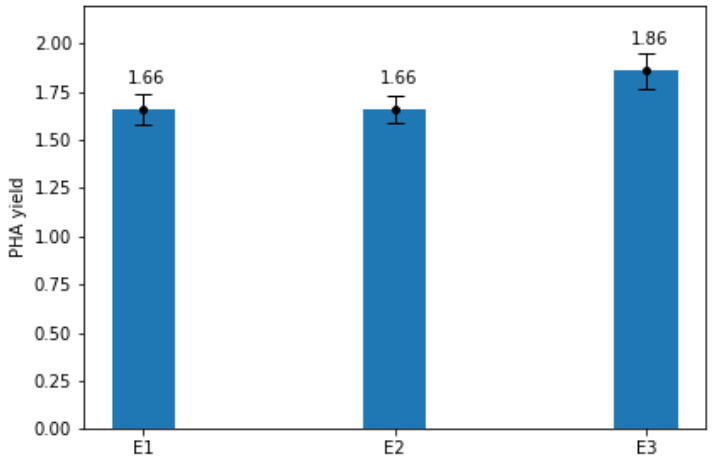


Figure 14. Mean yield of each strategy.

In terms of yield, as seen in Figure 14, the third strategy would be the best one again. As we are obtaining the yield based on the gained biomass and not the total biomass,

strategy 3 dedicates more carbons to PHA production than to grow, therefore obtaining a better yield than the other two strategies. Being, once more, the best strategy.

This strategy, however, shows less flexibility than the other, in the sense that, when studying all the solutions retrieved by FLYCOP, in the case of strategy 1 and 2, there were some answers almost as good as the first one, differing minimally from the best configuration. On the other hand, for the strategy 3, a lot less configurations are viable, showing a more noticeable difference between the best strategy and the rest, albeit the second best configuration solution still produces more PHA than the other two strategies (Annex 4). The starting biomass for this solution is significantly lower than the first solution, and they are collected in table 8.

Strategy 3 second best result					
Starting biomass (g/L)					
Severe oxygen restriction		Moderate oxygen restriction		No oxygen restriction	
P.putida-dPCA	*P.Putida-EG*	*P.putida-dPCA*	*P.Putida-EG*	*P.putida-dPCA*	*P.Putida-EG*
0.1	0.03	0.1	0.03	0.03	0.02
PHA production (mMol)					
3.33		3.35		3.37	

Table 8. Summary of biomass retrieved by FLYCOP in the second best solution of strategy three along with its production of PHA

As for the yield, the scenario with severe oxygen restriction has a yield of 1; the scenario with moderate oxygen restriction has a yield of 1.17. However, the scenario without oxygen restriction has a yield of 1.48, probably due to the fact that its starting biomasses are lower than the other scenarios. This yield being the highest obtained, meaning that even in terms of yields, the strategy 3 could be better than the others studied.

5. Discussion

Plastic waste management industry needs a twist, as of right now it is completely unsustainable, and its forecasts are even worse. Plastic is in our everyday life and even so, its recycling keeps being low, mainly being abandoned in oceans and seas. Even the most popular recycling techniques are not completely sustainable and green, as they produce toxic compounds or the recycled plastic cannot be used for every purpose.

In silico approaches have proven to be useful regarding the understanding and application of large-scale data, as is bacterial metabolism, with the creation of genomic-scale models that help us understand their metabolism. Moreover, thanks to the recent advances in algorithms allowing to simulate microbial behaviour with increasing precision each time, it is possible to use this simulation to understand, design and experiment complex biotechnological processes without needing to use *in vivo* work.

Degrading PET and its backbone monomers (TPHA and EG) is a process that can be carried out by microorganism satisfactorily, as found recently (Yoshida *et al.,* 2016*,* Müller *et al.,* 2005, Danso *et al*., 2019). In this project we present different approximations to this complicated task, although more fine tuning can be done to find new solutions or improve the existing ones. Moreover, not only is that process possible, but producing bioplastics, such as the one used in this project, PHA, can be accomplished with microorganisms in a whole single project covering both the degradation of the plastic and transforming it into a more sustainable, green product, such as bioplastic from unrelated carbon sources. In this case, compounds derived from PET degradation processes (Manoli, 2020).

In this project we have presented three successful approximations of degrading PET monomers and transforming them into PHA, a bioplastic using an engineered *Pseudomonas putida* KT2440 strain. Although the three strategies accomplish the project's main goal, it is clear that not all of them are equally good. In terms of plain numbers, the strategy three would be the best, meaning that incorporating microbial consortia to bioprocesses may be the key to unlock the hardest processes that cannot be done currently; or improving the results of the already existing ones.

Strategy 1 also shows promising results, but it is needed to consider that it could perform worse *in vivo* due to the large metabolic burden expected since the high number of genetic modifications required to address the PET-monomers revalorization into PHA in a single strain. Another problem it could lead to is catabolite repression, as some evidence suggest (Prieto et al., 2016). Strategy 2 also works fine and, although it performs slightly worse than strategy 3, could also be used as alternative.

Apart from this, a more profound investigation would be required to finish improving these strategies, because even if they perform well, it is certain there is room for improvement. Some of these improvements could come from studying the minimum number of mutations necessary for the model to maintain the maximum PHA produced, or adjusting fitness function or other parameters to see if significant changes appear.

The development section included in this project was also carried out successfully, managing to bring the idea of simplifying and making the future use of this program more accessible to the general user. Thanks to the COMETSpy library, this update occurred efficiently, improving execution times. Even so, it would be desirable in the future to finish updating FLYCOP completely, so that it uses only Python and does not rely on other softwares or programming languages. In addition, it also would be advantageous creating a graphical interface that makes FLYCOP more visual and intuitive. Due to the use of OOP, future updates to the main FLYCOP script that need to be carried out can be done in an easy and simple way.

6. Conclusions

1. FLYCOP has been updated successfully and its running time has been improved. Moreover, a series of scripts have been developed with the objective of trying to make this program more accessible and easier to use.
2. FLYCOP has demonstrated to be a powerful, adaptable and flexible tool that optimizes complex biotechnological processes simulations and allows to define an experiment with precision, which means that it can be used in a wide variety of situations and processes.

3. TPHA and EG can be degraded using *Pseudomonas putida* KT2440, allowing to explore this microorganism for biorecycling plastic.
4. Microbial consortia can be used to perform biotechnological processes with an equal or better result than using a pure culture, allowing to design more complex processes involving several strains with a modular approximation to overcome approaches that use only one strain, or open the door to processes that cannot be performed with pure cultures.
5. Using *Pseudomonas putida* KT2440, PHA can be produced using TPHA and EG as a carbon sources, creating a cycle of degrading plastic and producing a market value compound at the same time, in this case, a promising degradable and bioplastic.

7. Glossary of acronyms

- PET: Polyethylene terephthalate
- TPHA: Terephthalate acid
- EG: Ethylene glycol
- MIX-UP: Mixed plastic biodegradation and upcycling using microbial communities
- GEM: Genome-scale metabolic model
- FBA: Flux balance analysis
- FVA: Flux variability analysis
- COBRA: Constraint-based reconstruction and analysis
- SBML: Systems biology markup language
- SMN: Stoichiometric metabolic network
- COMETS: Computation of microbial ecosystems in time and space
- SMAC: Sequential model-based algorithm configuration
- SMBO: Sequential model-based optimization
- FLYCOP: Flexible synthetic consortium optimization

- PHA: Polyhydroxyalkanoate
- OOP: Object-oriented programming
- PCA: Protocatechuate acid

8. Bibliography

Anaconda Software Distribution. (2020). Anaconda Documentation. Anaconda Inc. Retrieved from https://docs.anaconda.com/

Becker, S. A., Feist, A. M., Mo, M. L., Hannum, G., Palsson, B. Ø., & Herrgard, M. J. (2007). Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox. *Nature Protocols*, *2*(3), 727–738. https://doi.org/10.1038/nprot.2007.99

Brenner, K., You, L., & Arnold, F. H. (2008). Engineering microbial consortia: A new frontier in synthetic biology. *Trends in Biotechnology*, *26*(9), 483–489. https://doi.org/10.1016/j.tibtech.2008.05.004

Bushell, M. E., Sequeira, S. I. P., Khannapho, C., Zhao, H., Chater, K. F., Butler, M. J., Kierzek, A. M., & Avignone-Rossa, C. A. (2006). The use of genome scale metabolic flux variability analysis for process feed formulation based on an investigation of the effects of the zwf mutation on antibiotic production in Streptomyces coelicolor. *Enzyme and Microbial Technology*, *39*(6), 1347–1353. https://doi.org/10.1016/j.enzmictec.2006.06.011

Cavaliere, M., Feng, S., Soyer, O. S., & Jiménez, J. I. (2017). Cooperation in microbial communities and their biotechnological applications. *Environmental Microbiology*, *19*(8), 2949–2963. https://doi.org/10.1111/1462-2920.13767

Danso, D., Schmeisser, C., Chow, J., Zimmermann, W., Wei, R., Leggewie, C., Li, X., Hazen, T., & Streit, W. R. (2018). New Insights into the Function and Global Distribution of Polyethylene Terephthalate (PET)-Degrading Bacteria and Enzymes in Marine and Terrestrial Metagenomes. *Applied and Environmental Microbiology*, *84*(8), e02773-17. https://doi.org/10.1128/AEM.02773-17

Ebrahim, A., Lerman, J. A., Palsson, B. O., & Hyduke, D. R. (2013). COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, *7*(1), 74. https://doi.org/10.1186/1752-0509-7-74

EU, 2013. Green Paper on a European Strategy on Plastic Waste in the Environment. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52013DC0123>. Recuperado 16 de diciembre 2020.

Gamerith, C., Zartl, B., Pellis, A., Guillamot, F., Marty, A., Acero, E. H., & Guebitz, G. M. (2017). Enzymatic recovery of polyester building blocks from polymer blends. *Process Biochemistry*, *59*, 58–64. https://doi.org/10.1016/j.procbio.2017.01.004

García-Jiménez, B., García, J. L., & Nogales, J. (2018). FLYCOP: Metabolic modeling-based analysis and engineering microbial communities. *Bioinformatics*, *34*(17), i954–i963. https://doi.org/10.1093/bioinformatics/bty561

George, N., & Kurian, T. (2014). Recent Developments in the Chemical Recycling of Postconsumer Poly(ethylene terephthalate) Waste. *Industrial & Engineering Chemistry Research*, *53*(37), 14185–14198. https://doi.org/10.1021/ie501995m

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., & Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biology*, *20*(1), 121. https://doi.org/10.1186/s13059-019-1730-3

Hanly, T. J., & Henson, M. A. (2011). Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. *Biotechnology and Bioengineering*, *108*(2), 376–385. https://doi.org/10.1002/bit.22954

Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., Bonilla, G., Kar, A., Leiby, N., Mehta, P., Marx, C. J., & Segrè, D. (2014). Metabolic Resource Allocation in Individual Microbes Determines Ecosystem Interactions and Spatial Dynamics. *Cell Reports*, *7*(4), 1104–1115. https://doi.org/10.1016/j.celrep.2014.03.070

Hiraga, K., Taniguchi, I., Yoshida, S., Kimura, Y., & Oda, K. (2019). Biodegradation of waste PET. *EMBO Reports*, *20*(11). https://doi.org/10.15252/embr.201949365

Hoos, H. H., & Stützle, T. G. (2005). *Stochastic local search: Foundations and applications*. Elsevier.

Hurley, R. R., Woodward, J. C., & Rothwell, J. J. (2017). Ingestion of Microplastics by Freshwater Tubifex Worms. *Environmental Science & Technology*, *51*(21), 12844–12851. https://doi.org/10.1021/acs.est.7b03567

Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2011). Sequential Model-Based Optimization for General Algorithm Configuration. In C. A. C. Coello (Ed.), *Learning and Intelligent Optimization* (Vol. 6683, pp. 507–523). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-25566-3_40

Jawed, K., Yazdani, S. S., & Koffas, M. AG. (2019). Advances in the development and application of microbial consortia for metabolic engineering. *Metabolic Engineering Communications*, *9*, e00095. https://doi.org/10.1016/j.mec.2019.e00095

Jiang, L.-L., Zhou, J.-J., Quan, C.-S., & Xiu, Z.-L. (2017). Advances in industrial microbiome based on microbial consortium for biorefinery. *Bioresources and Bioprocessing*, *4*(1), 11. https://doi.org/10.1186/s40643-017-0141-0

Jones, J. A., Toparlak, Ö. D., & Koffas, M. A. (2015). Metabolic pathway balancing and its role in the production of biofuels and chemicals. *Current Opinion in Biotechnology*, *33*, 52–59. https://doi.org/10.1016/j.copbio.2014.11.013

Joo, S., Cho, I. J., Seo, H., Son, H. F., Sagong, H.-Y., Shin, T. J., Choi, S. Y., Lee, S. Y., & Kim, K.-J. (2018). Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nature Communications*, *9*(1), 382. https://doi.org/10.1038/s41467-018-02881-1

Keller, L., & Surette, M. G. (2006). Communication in bacteria: An ecological and evolutionary perspective. *Nature Reviews Microbiology*, *4*(4), 249–258. https://doi.org/10.1038/nrmicro1383

Kuepfer, L. (2014). Stoichiometric Modelling of Microbial Metabolism. In J. O. Krömer, L. K. Nielsen, & L. M. Blank (Eds.), *Metabolic Flux Analysis* (Vol. 1191, pp. 3–18). Springer New York. https://doi.org/10.1007/978-1-4939-1170-7_1

Lee, J. W., Yi, J., Kim, T. Y., Choi, S., Ahn, J. H., Song, H., Lee, M.-H., & Lee, S. Y. (2016). Homo-succinic acid production by metabolically engineered Mannheimia succiniciproducens. *Metabolic Engineering*, *38*, 409–417. https://doi.org/10.1016/j.ymben.2016.10.004

Li, C.-T., Zhuang, H.-K., Hsieh, L.-T., Lee, W.-J., & Tsao, M.-C. (2001). PAH emission from the incineration of three plastic wastes. *Environment International*, *27*(1), 61–67. https://doi.org/10.1016/S0160-4120(01)00056-3

Mahadevan, R., & Schilling, C. H. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering*, *5*(4), 264–276. https://doi.org/10.1016/j.ymben.2003.09.002

Mahadevan, Radhakrishnan, Edwards, J. S., & Doyle, F. J. (2002). Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli. *Biophysical Journal*, *83*(3), 1331–1340. https://doi.org/10.1016/S0006-3495(02)73903-9

Manoli, M. T. (2020). Optimización del metabolismo de polihidroxialcanoatos de Pseudomonas putida KT2440 mediante abordajes de biología sintética y de sistemas.

Marshall, I., & Todd, A. (1953). The thermal degradation of polyethylene terephthalate. *Transactions of the Faraday Society*, *49*, 67. https://doi.org/10.1039/tf9534900067

Meng, Y., Kelly, F. J., & Wright, S. L. (2020). Advances and challenges of microplastic pollution in freshwater ecosystems: A UK perspective. *Environmental Pollution*, *256*, 113445. https://doi.org/10.1016/j.envpol.2019.113445

MIX-UP (2020). Retrieved from https://www.mix-up.eu/

Müller, R.-J., Schrader, H., Profe, J., Dresler, K., & Deckwer, W.-D. (2005). Enzymatic Degradation of Poly(ethylene terephthalate): Rapid Hydrolyse using a Hydrolase fromT. fusca. *Macromolecular Rapid Communications*, *26*(17), 1400–1405. https://doi.org/10.1002/marc.200500410

Nogales, J., Gudmundsson, S., Duque, E., Ramos, J. L., & Palsson, B. O. (2017). Expanding the computable reactome in *Pseudomonas putida* reveals metabolic cycles providing robustness [Preprint]. Systems Biology. https://doi.org/10.1101/139121

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature Biotechnology*, *28*(3), 245–248. https://doi.org/10.1038/nbt.1614

Pathak, S., Sneha, C. L. R., & Mathew, B. B. (2014). Bioplastics: its timeline based scenario & challenges. Journal of Polymer and Biopolymer Physics Chemistry, 2(4), 84-90.

Perez-Garcia, O., Lear, G., & Singhal, N. (2016). Metabolic Network Modeling of Microbial Interactions in Natural and Engineered Environmental Systems. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.00673

PET Plastic (Polyethylene Terephthalate): Uses, Properties & Structure. (s. f.). Omnexus. Recuperado 16 de noviembre de 2020, de https://omnexus.specialchem.com/selection-guide/polyethylene-terephthalate-pet-plastic

Prieto, A., Escapa, I. F., Martínez, V., Dinjaski, N., Herencias, C., de la Peña, F., ... & Revelles, O. (2016). A holistic view of polyhydroxyalkanoate metabolism in Pseudomonas putida. Environmental microbiology, 18(2), 341-357. https://doi.org/10.1111/1462-2920.12760

Plastics Europe. (2019, October 14). Plastics Europe. https://www.plasticseurope.org/application/files/9715/7129/9584/FINAL_web_version_Plastics_the_facts2019_14102019.pdf

Ragaert, K., Delva, L., & Van Geem, K. (2017). Mechanical and chemical recycling of solid plastic waste. *Waste Management*, *69*, 24–58. https://doi.org/10.1016/j.wasman.2017.07.044

Reed, J. L., & Palsson, B. Ø. (2004). Genome-Scale In Silico Models of E. coli Have Multiple Equivalent Phenotypic States: Assessment of Correlated Reaction Subsets That Comprise Network States. *Genome Research*, *14*(9), 1797–1805. https://doi.org/10.1101/gr.2546004

Samak, N. A., Jia, Y., Sharshar, M. M., Mu, T., Yang, M., Peh, S., & Xing, J. (2020). Recent advances in biocatalysts engineering for polyethylene terephthalate plastic waste green recycling. *Environment International*, *145*, 106144. https://doi.org/10.1016/j.envint.2020.106144

Schellenberger, J., & Palsson, B. Ø. (2009). Use of Randomized Sampling for Analysis of Metabolic Networks. *Journal of Biological Chemistry*, *284*(9), 5457–5461. https://doi.org/10.1074/jbc.R800048200

Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., Kang, J., Hyduke, D. R., & Palsson, B. Ø. (2011). Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0. *Nature Protocols*, *6*(9), 1290–1307. https://doi.org/10.1038/nprot.2011.308

Sinha, V., Patel, M. R., & Patel, J. V. (2010). Pet Waste Management by Chemical Recycling: A Review. *Journal of Polymers and the Environment*, *18*(1), 8–25. https://doi.org/10.1007/s10924-008-0106-7

Street, B., & Bey, M. (2014). Mixed microbial cultures for industrial biotechnology: success, chance, and challenges. Industrial Biocatalysis, 1, 205.

Sussarellu, R., Suquet, M., Thomas, Y., Lambert, C., Fabioux, C., Pernet, M. E. J., Le Goïc, N., Quillien, V., Mingant, C., Epelboin, Y., Corporeau, C., Guyomarch, J., Robbens, J., Paul-Pont, I., Soudant, P., & Huvet, A. (2016). Oyster reproduction is affected by exposure to

polystyrene microplastics. *Proceedings of the National Academy of Sciences*, *113*(9), 2430–2435. https://doi.org/10.1073/pnas.1519019113

Taniguchi, I., Yoshida, S., Hiraga, K., Miyamoto, K., Kimura, Y., & Oda, K. (2019). Biodegradation of PET: Current Status and Application Aspects. *ACS Catalysis*, *9*(5), 4089–4105. https://doi.org/10.1021/acscatal.8b05171

Thiele, I., Fleming, R. M. T., Bordbar, A., Schellenberger, J., & Palsson, B. Ø. (2010). Functional Characterization of Alternate Optimal Solutions of Escherichia coli's Transcriptional and Translational Machinery. *Biophysical Journal*, *98*(10), 2072–2081. https://doi.org/10.1016/j.bpj.2010.01.060

Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. Scotts Valley, CA: CreateSpace.

von Moos, N., Burkhardt-Holm, P., & Köhler, A. (2012). Uptake and Effects of Microplastics on Cells and Tissue of the Blue Mussel *Mytilus edulis* L. after an Experimental Exposure. *Environmental Science & Technology*, *46*(20), 11327–11335. https://doi.org/10.1021/es302332w

Wei, R., Oeser, T., & Zimmermann, W. (2014). Synthetic Polyester-Hydrolyzing Enzymes From Thermophilic Actinomycetes. In *Advances in Applied Microbiology* (Vol. 89, pp. 267–305). Elsevier. https://doi.org/10.1016/B978-0-12-800259-9.00007-X

Wei, R., & Zimmermann, W. (2017). Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: How far are we? *Microbial Biotechnology*, *10*(6), 1308–1322. https://doi.org/10.1111/1751-7915.12710

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., & Oda, K. (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science*, *351*(6278), 1196–1199. https://doi.org/10.1126/science.aad6359

Zhang, H., & Wang, X. (2016). Modular co-culture engineering, a new approach for metabolic engineering. *Metabolic Engineering*, *37*, 114–121. https://doi.org/10.1016/j.ymben.2016.05.007

Zomorrodi, A. R., & Maranas, C. D. (2012). OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. *PLoS Computational Biology*, *8*(2), e1002363. https://doi.org/10.1371/journal.pcbi.1002363

Zuroff, T. R., & Curtis, W. R. (2012). Developing symbiotic consortia for lignocellulosic biofuel production. *Applied Microbiology and Biotechnology*, *93*(4), 1423–1435. https://doi.org/10.1007/s00253-011-3762-9

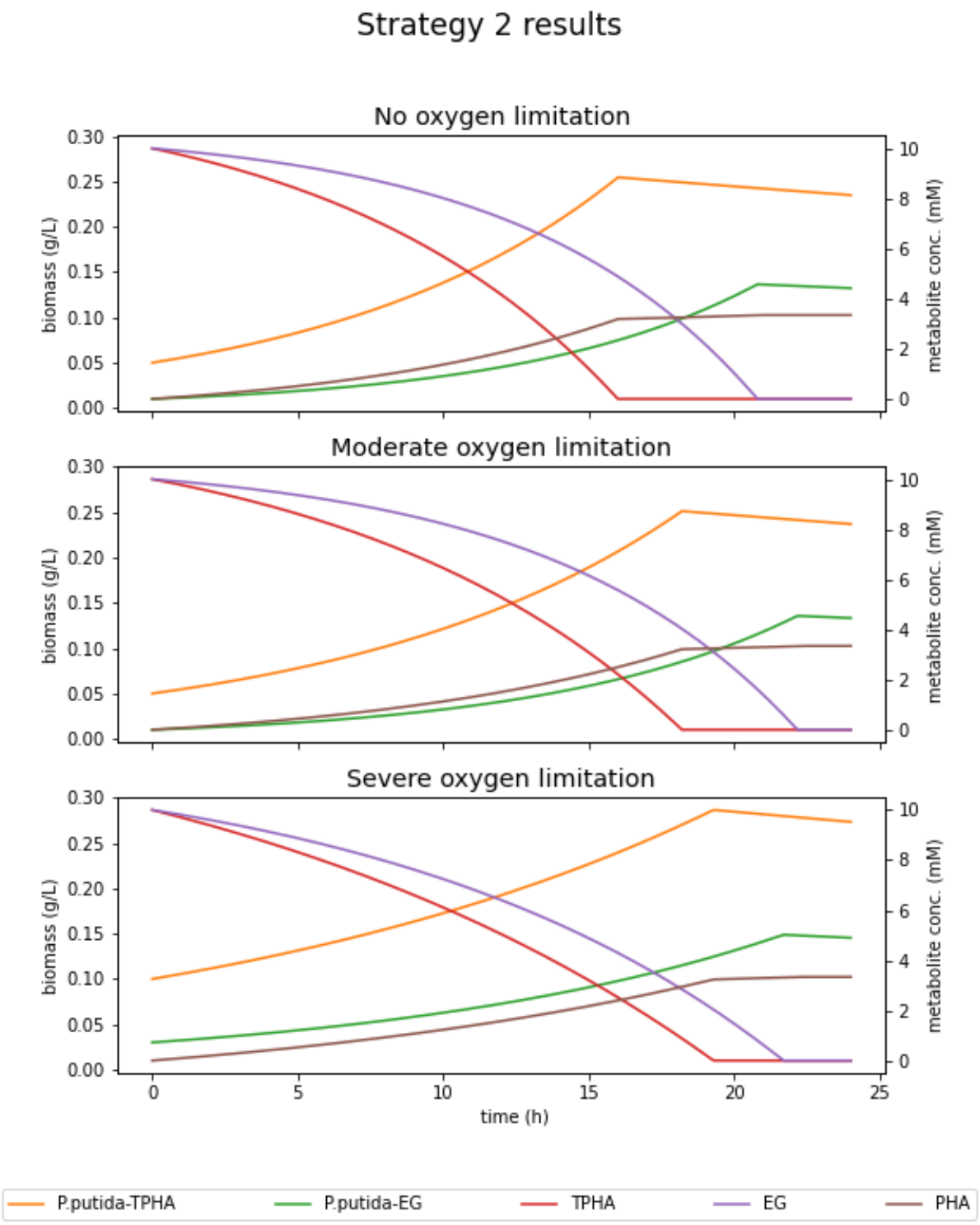
9. Annex

Annex 1. Reactions deleted from all the strategies in order to maximize PHA production. Their bounds were changed to 0 using COBRApy in order to make the bacteria not be able to use it.

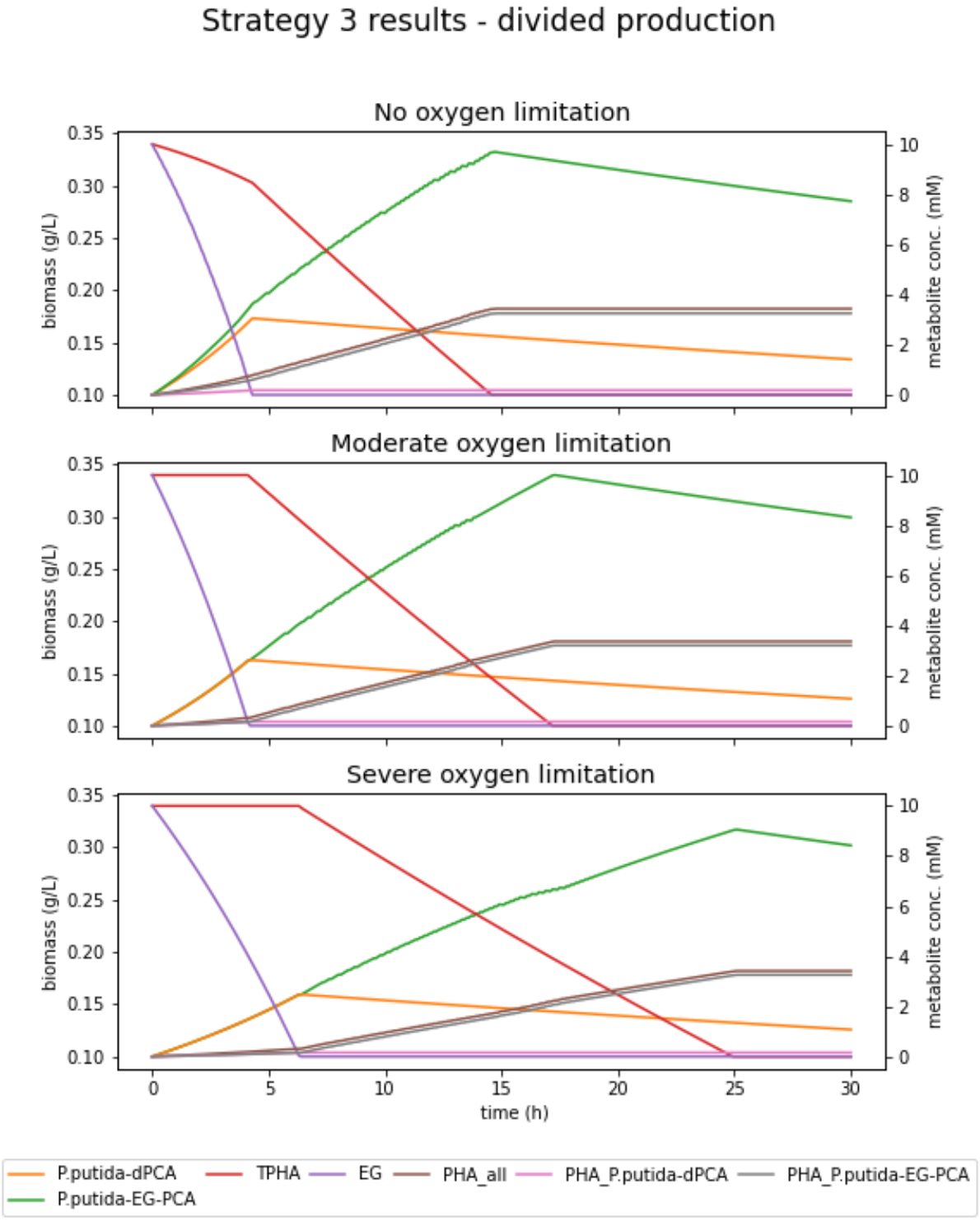
Name	Description	Reaction
PDH	Pyruvate dehydrogenase	coa_c + nad_c + pyr_c → accoa_c + co2_c + nadh_c
RHACOAE80	R-hydroxy-acyl CoA thioesterase C80	h2o_c + R_3hocoa_c ⇌ coa_c + h_c + R_3hocta_c
PHADPC80	Poly-3-hydroxyalkanoate depolymerase C80	h2o_c + C80aPHA_c ⇌ h_c + PHAg_c + R_3hocta_c
CYSS	Cysteine synthase	acser_c + h2s_c ⇌ ac_c + cys__L_c + h_c
NACODA	N-acetylornithine deacetylase	acg5sa_c + h2o_c ⇌ ac_c + glu5sa_c
ALDD2x	Aldehyde dehydrogenase	acald_c + h2o_c + nad_c ⇌ ac_c + 2.0 h_c + nadh_c
ABUTD	Aminobutyraldehyde dehydrogenase	4abutn_c + h2o_c + nad_c ⇌ 4abut_c + 2.0 h_c + nadh_c
ALAR	Alanine racemase	ala__L_c ⇌ ala__D_c
ALATA_L	L-alanine transaminase	akg_c + ala__L_c ⇌ glu__L_c + pyr_c
FORGLUIH2	N formimino L glutamate inimohydrolase	h_c + h2o_c + forglu_c ⇌ nh4_c + Nforglu_c
ICDHyr	Isocitrate dehydrogenase (NADP)	icit_c + nadp_c ⇌ akg_c + co2_c + nadph_c
ICL	Isocitrate lyase	icit_c ⇌ glx_c + succ_c
MALS	Malate synthase	accoa_c + glx_c + h2o_c ⇌ coa_c + h_c + mal__L_c
MCITL2	Methylisocitrate lyase	micit_c ⇌ pyr_c + succ_c
OARGDC	Oxoarginine decarboxylase	h_c + 5g2oxpt_c ⇌ co2_c + gdbtal_c
ORNCD	Ornithine cyclodeaminase	orn_c ⇌ nh4_c + pro__L_c

PDHbr	Pyruvate dehydrogenase (dihydrolipoamide) reversible	coa_c + adhlam_c ⇌ accoa_c + dhlam_c
PDHcr	Pyruvate dehydrogenase (dihydrolipoamide dehydrogenase) reversible	nad_c + dhlam_c ⇌ h_c + nadh_c + lpam_c
PPCSCT	Propanoyl-CoA: succinate CoA-transferase	ppcoa_c + succ_c ⇌ ppa_c + succoa_c
RBK	Ribokinase	atp_c + rib__D_c ⇌ adp_c + h_c + r5p_c
SUCOAS	Succinyl-CoA synthetase (ADP-forming)	atp_c + coa_c + succ_c ⇌ adp_c + pi_c + succoa_c
UPPN	B-ureidopropionase	2.0 h_c + h2o_c + cala_c ⇌ ala_B_c + co2_c + nh4_c
DM_C80aPHA_c	C80aPHA demand	C80aPHA_c ⇌

Annex 2. Figure of strategy 2 not showing PHA production separately in each strain for a clearer visualization.



Annex 3. Strategy three first result with the production of PHA divided by the two strains.



Annex 4. Second best results of strategy 3, according to FLYCOP.