University of West Bohemia
Faculty of Applied Sciences
Department of Cybernetics

BACHELOR THESIS

Modeling and Estimation of Glycolysis Reingulation in Escherichia coli

Pilsen 2014 Veronika Řežábková

Prohlášení

Předkládám tímto k posouzení a obhajobě bakalářskou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni dne 16. května 2014

. . . . . .

Declaration

I hereby declare that this bachelor thesis is completely my own work and that I used only the cited sources.

Abstrakt

Glykolýza je důležitou součástí metabolických drah bakterie Escherichia coli. Metabolické dráhy zajišťují rozkladem glukózy tvorbu energie v podobě ATP a v přední řadě buněčný růst. Dráhy jsou pevně spjaty a proto regulace glykolýzy nezávisí pouze na jejích vlastních reakcích, ale na celé metabolické síti. Byl navžen jednoduchý optimalizační model chování metabolické sítě a porovnán s výsledky experimentů.

Klíčová slova: glykolýza, regulace, enzymy, matematický model, metabolické dráhy, optimalizace

Abstract

Glycolysis is an important part of the metabolic pathways of Escherichia coli. Metabolic pathways ensure by decomposition of glucose production of energy in the form of ATP and in the front row cellular growth. Pathways are tightly linked and therefore regulation of glycolysis depends not only on its own reaction, but on the whole metabolic network. A simple optimization model of metabolic network behavior was designed and compared with experimental results.

Keywords: glycolysis, regulation, enzymes, mathematical model, metabolic pathways, optimalization

Contents

1 Introduction

Synthetic biology is a new field of research that combines genetics, chemistry and engineering. It approach the creation of new biological systems from different perspectives, focusing on finding how life works.

This work focuses on synthetic biology in terms of Cybernetics, when any biological object can be seen as a system. Such a biological object can be mathematically describe, to model and simulate, thus we achieve a better understanding of the behavior of biological objects.

2 Biological background

If we want to model metabolic pathways, it's important to know essential biological regularities, which controls the metabolic network. These regularities or we can say rules will form the basis for modeling of metabolic network.

2.1 The Glycolitic pathway

Glycolysis is the main pathway of central metabolism. In this work provides introductory concept relevant to the remainder of this work.

2.1.1 Metabolism

Metabolism is a set of chemical and physical processes in an organism, which produces and destroys substances and produces energy. Metabolism can be divide into anabolism and catabolism. Anabolism is the process, by which complex molecules are form smaller molecules, whereas catabolism is the breakdown of complex molecules.

Central metabolic pathways provide the precursor metabolites to all other pathways. The central pathways for decomposing carbohydrates are the Embden-Mayerhof-Parnas pathway (also called glycolysis), the Pentose Phosphate pathway and the Entner-Doudoroff pathway. All three pathways use different mechanisms to convert glucose into phosphoglyceraldehyde and same mechanisms to convert phosphoglyceraldehyde to pyruvate. [5]

2.1.2 Glycolysis

Glycolysis is a sequence of ten enzymatic reactions, which convert glucose into pyruvate. The free energy released during this process is used for production of the high-energy compound ATP (adenosine triphosphate) and NADH (nicotinamide adenine dinucleotide). Glycolysis is used to produce ATP in aerobic and anaerobic conditions. [5]

2.1.3 Reactions of glycolysis

As many metabolic pathways, glycolysis is divided into separate phases with differing roles.

Preparatory phase

This phase consumes energy to destabilization and cleavage of glucose in the blood and thereby increase its intake. The following reactions represents this phase:

1.) $Glucose + ATP \longrightarrow G6P$
In the first reaction 1 ATP is consumed for conversion of glucose into G6P (glucose-6-phosphate). G6P is also an essential metabolite for cellular growth.

2.) $G6P \rightleftharpoons F6P$
Followed change in structure from G6P to F6P (fructose-6-phosphate) - the next metabolite important for cellular growth.

3.) $F6P + ATP \longrightarrow F16BP$
This is another energy-loss reaction.1ATP is consumed to convert F6P into F16BP (fructose-1,6-bisphosphate).

4.) $F16BP \rightleftharpoons GA3P + DHAP$
The next is splitting of F16BP to two molecules - GA3P (glycealdehyde-3-phosphate) and DHAP (dihydroxyacetone phosphate). GA3P is one of the significant metabolites for cellular growth.

5.) $DHAP \rightleftharpoons GA3P$
Now is creating a second GA3P conversion of DHAP.

Pay-off phase

The second part of glycolysis generates energy in the form of ATP and NADH.

6.) $GA3P + NAD^+ \rightleftharpoons 13BPG$
GA3P is converted into 13BPG (1,3-bisphosphoglycerate) by consump-

tion of NAD+. NADH is formed which is also important metabolite in cellular growth.

7.) $13BPG + ADP \rightleftharpoons 3PG$
This reaction is important because ATP is formed during conversion of 13BPG into the next metabolite important for cellular growth 3PG (3-phosphoglycerate).

8.) $3PG \rightleftharpoons 2PG$
2PG (2-phosphoglycerate) is creating from 3PG.

9.) $2PG \rightleftharpoons PEP$
In the ninth step is formed PEP (phosphoenolpyruvate) from 2PG. PEP is metabolite used for cellular growth.

10.) $PEP + ADP \longrightarrow Pyruvate$
The final step is the second energy-yielding reaction. ATP is formed during conversion of PEP into P (pyruvate), which is also essential for cellular growth.

[6]

2.2 Enyzmatic reactions and Enzymes

Enzymatic reactions are the fundamental building blocks of the models introduced in the next chapter.

Almost all enzymatic reactions in living organisms are made possible through the catalytic effect of biological catalysts - enzymes. Enzymes are simple or complex proteins.

2.2.1 Enzymatic reaction

An enzymatic reaction is a biochemical reaction that must be catalyzed by an enzyme in order to proceed under normal conditions. The enzyme reduces the activation energy of the reaction by creating an intermediate enzyme-substrate complex. The substrate binds to the active site of the enzyme, a three-dimensional shape made up of amino acid residues, where undergoes

the corresponding chemical transformation. We can imagine that the substrate has the form of a key that fits only in an enzyme whose cutout shape corresponds to the key lock. Once bound, the enzyme is able to adapt to the shape of the key further securing the bond. This principle is called the lock and key theory and we can see it in the picture below.

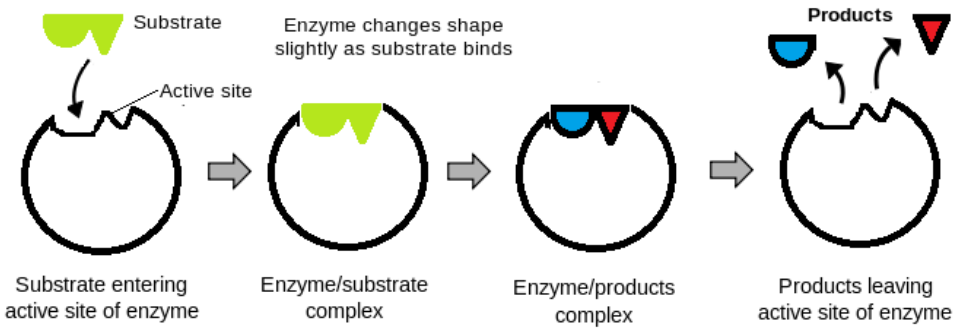


Figure 2.1: The theory of key and lock - princip of enzymatic reaction.

[6]

2.2.2 Enzymes of glycolysis

The central metabolism modeled in the next chapter includes three types of reactions:

Category 1:

$$Substrate1 \quad + \quad Substrate2 \quad \overset{Enzyme}{\longrightarrow} \quad Product$$

This category includes the following enzymes:

Kinase: A kinase is a type of enzyme that transfers phosphate groups PO_3 from high-energy donor molecules, such as ATP, to specific substrates, a process called phosphorylation. Kinases are a part of the larger family of phosphotransferases.

- **Hexokinase** is an allosteric enzyme which is strongly inhibited by the product of glucose-6-phosphate. In the first reaction hexokinase adds a phosphate group to glucose to form glucose-6-phosphate.
- **Phosphofructokinase-1(PFK-1)**is a key enzyme of glycolysis, which regulates the rate of sequence. It's allosteric regulation and inducible enzyme. PFK-1 catalyzes the addition of phosphate group to fructose-6-phosphate and ATP to create fructose-1,6-bisphosphate and ADP.
- **Phosphoglycerate kinase** is an enzyme that catalyzes releasing of phosphate from 1,3-bisfosfoglycerátu and adding it to ADP to form ATP. The product of reaction is 3-phosphoglycerate.
- **Pyruvate kinase** catalyzes the transfer of a phosphate group from phosphoenolpyruvate to ADP yielding one molecule of stable pyruvate and one molecul of ATP.

Dehydrogenase: A dehydrogenase is an enzyme that catalyses the removal of hydrogen.

- **Glyceraldehyde phosphate dehydrogenase** is the pyridine enzyme with NAD^+. It catalyzes the irreversible reaction involving dehydrogenation and phosphorylation, which forms 1,3-bisphosphoglycerate and $NADH + H$

Category 2:

$$Substrate \overset{Enzyme}{\longrightarrow} Product$$

The following enzymes representing this category.

Isomerase: An isomerase is an enzyme that catalyzes the structural rearrangement of isomers. The structural changes is called isomerization.

- **Glucose-6-phosphate isomerase** in an enzyme that catalyzes the conversion of glucose-6-phosphate into fructose-6-phosphate.
- **Triosephosphate isomerase** is an enzyme that catalyzes the reversible isomerization reaction when both phosphorylated triose transferred from one to another - dihydroxyacetone phosphate to D-glyceraldehyde 3-phosphate.

Mutase: A mutase is an enzyme that catalyzes the transferring of a functional group from one position to another within the same molecule.

- **Phosphoglycerate mutase** is an enzyme that catalyzes the internal transfer of a phosphate group from 3-phosphoglycerate to 2-phosphoglycerate.

Lyase: A lyase is an enzyme that catalyzes the breaking of various chemical bonds by means other than hydrolysis and oxidation, often forming a new double bond or a new ring structure.

- **Enolase** chipped off water from 2-phosphoglycerate and dehydrated product is very unstable phosphoenolpyruvate.

Category 3:

$$Substrate \overset{Enzyme}{\longrightarrow} Product1 + Product2$$

This category is represented with only one enzyme.

Lyase:
- **Aldolase** is an enzyme that occurs in several isoenzymes. It catalyzes an aldol splitting reaction - the substrate, fructose-1,6-bisphosphate is broken down into glyceraldehyde-3-phosphate and dihydroxyacetone phosphate.

[6]

2.3 Regulation of glycolysis

Glycolysis is regulated for generating of ATP and providing of building blocks for synthetic reactions. Potential sites of regulation are irreversible reaction catalyzed by kinases.

Two of the most common control strategies used by engineers are implemented in the regulation of this enzymatic reaction: feedback and feedforward. Feedback is through the pathway metabolite ATP. Feedforward is through secondary enzyme F26BP.

2.3.1 Regulation of fructose-1,6-bisphosphate

One of the irreversible steps (large decrease in Gibbs free energy) which determines the speed of glycolysis, is the phosphofructokinase reaction (fructose-6-phosphate to fructose-1,6-bisphosphate). This is the key regulatory point

of glycolysis. The activity of this allosteric enzyme is regulated by several factors whose effects are shown in Figure 2.2.

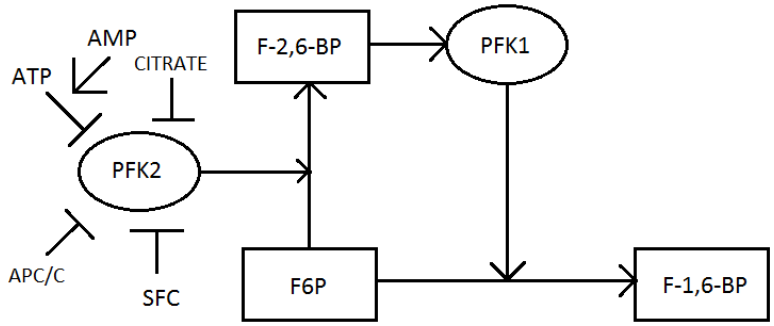


Figure 2.2: Feedforward in third reaction with regulators

On the block diagram is a model of third reaction including a regulation step by F26BP. Fructose-2,6-bisphosphate is a metabolite, which is synthesized by phosphorylation of F6P using ATP from PFK2 enzyme. F26BP strongly activate breaking down of glucose through allosteric modulation of PFK1. Increase of F26BP activates PFK1 so that increase enzyme's affinity for F6P and currently decrease his affinity for inhibit ATP and citrate. Another regulation factor is delivery of oxygen called Pasteurs efect. Plenty of oxygen contributes to save up of glucose. Nature of Pasteur efect is in inhibition efect of ATP to phosphofructokinase, when proportion of [AMP] : [ATP] is low. When there is deficiency of oxygen [ATP] decreases and glycolysis speeds up. AMP - adenoin monophosphate is formed by splitting of phosphate group from ADP. AMP works as indicator low-energy state of cell, when ATP is deficiency. Plenty of ATP prevents other consumtion of glucose as nutrients and thus it saves.

However, F26BP not continue nowhere further just goes back to F6P and therefore this regulatory point unusable.

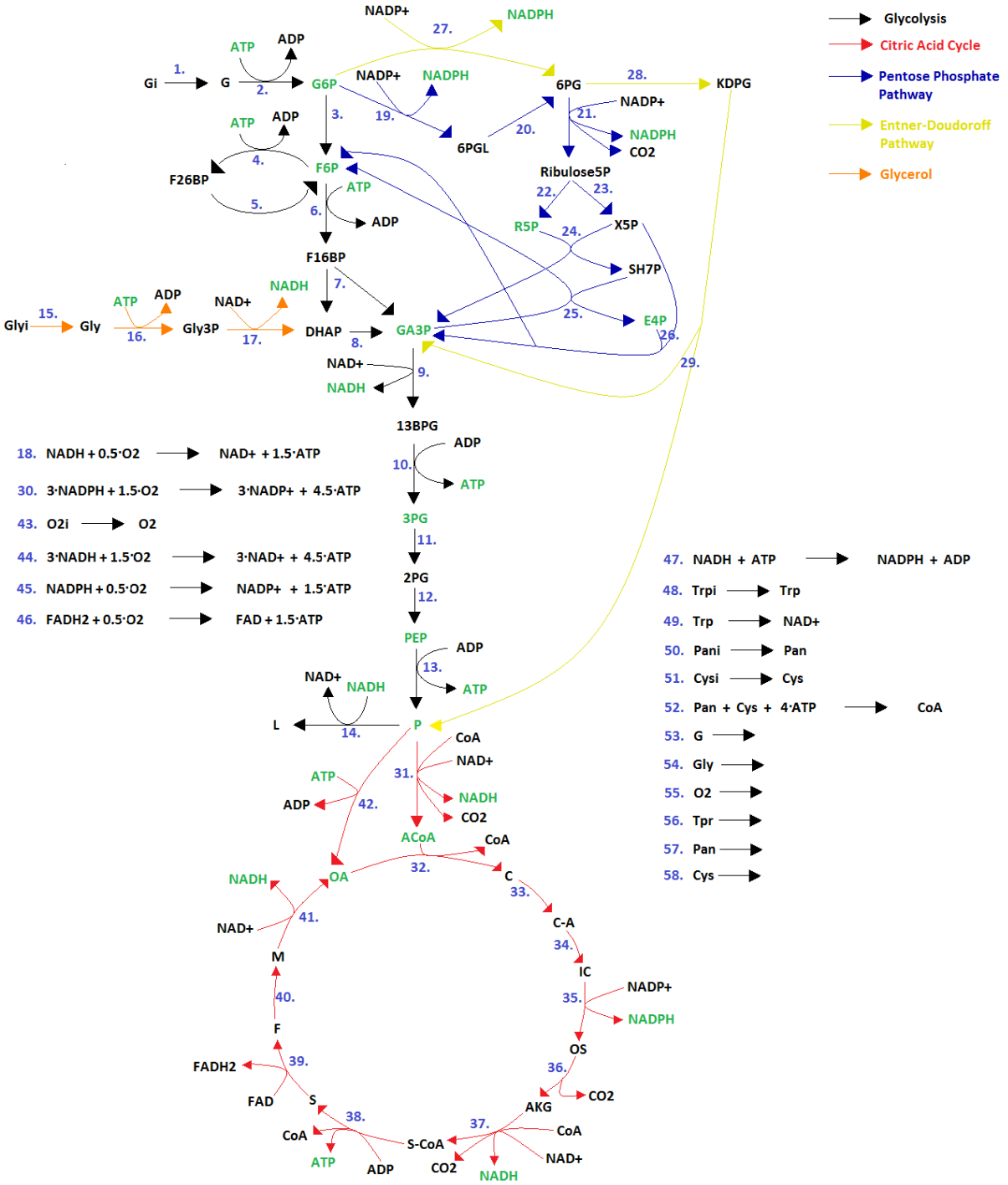
As can be seen, enzymatic reactions are complex. In practice, it is difficult to identify or measure all regulatory and kinetic parameters. Instead, optimization techniques are used in the next section to compute fluxes that are likely to be realized by live cells. [6]

3 The Steady state

In this chapter, a steady state representation of the Glycolytic pathway is derived. The steady state representation of the entire central metabolism (including the Glycolytic pathway) forms the parametric model used in the next Chapter to compute expected metabolic fluxes.

3.1 Biological model

Due to a change of approach (from control to optimize), we focused on complex central metabolism, not only on glycolysis. This led to the inclusion of other metabolic pathways such as the Pentose Phosphate Pathway, Entner-Doudoroff Pathway and Citric Acid Cycle. Was also added the second input glycerol. The entire biological model can be seen in the following figure.


Glycolysis
Citric Acid Cycle
Pentose Phosphate Pathway
Entner-Doudoroff Pathway
Glycerol
Gli
G
G6P
F6P
F26BP
F16BP
6PG
6PGL
KDPG
Ribulose5P
R5P
X5P
SH7P
E4P
Glyi
Gly
Gly3P
DHAP
GA3P
13BPG
3PG
2PG
PEP
P
L
ACoA
OA
C
C-A
IC
OS
AKG
S-CoA
S
F
FADH2
FAD
M
NADP+
NADPH
ATP
ADP
NAD+
NADH
CoA
CO2
18. NADH + 0.5·O2 → NAD+ + 1.5·ATP
30. 3·NADPH + 1.5·O2 → 3·NADP+ + 4.5·ATP
43. O2i → O2
44. 3·NADH + 1.5·O2 → 3·NAD+ + 4.5·ATP
45. NADPH + 0.5·O2 → NADP+ + 1.5·ATP
46. FADH2 + 0.5·O2 → FAD + 1.5·ATP
47. NADH + ATP → NADPH + ADP
48. Trpi → Trp
49. Trp → NAD+
50. Pani → Pan
51. Cysi → Cys
52. Pan + Cys + 4·ATP → CoA
53. G →
54. Gly →
55. O2 →
56. Tpr →
57. Pan →
58. Cys →

[5]

Figure 3.1: In the picture can easily recognize the different metabolic pathways through different colored arrows. Gi, Glyi, O2i, Trpi, Pani and Cysi takes as inputs the system. On the sides are reactions 18, 30, 44,45,46, 47 and 52, which balance the whole metabolic system and one reaction for balancing glycolysis conversion of pyruvate to lactate. Further reaction 53, 54, 55, 56, 57 and 58 are the output reactions for the case when the quantity of input material will not be consumed,so the residue was drained off. The green metabolites are metabolites which are necessary for cellular growth.

3.2 Flux Balance Analysis

Flux Balance Analysis is a modeling approach based on the constraints on the metabolic network. The most important limitation concerns the stoichiometric matrix. This approach narrows the range of options that a metabolic network can accept based on limitations which the cell must follow. Each of these constraints can be mathematically described as meaning that each flux of metabolic network represents dimension of the solution space, and therefor can be represented by the axes in the graph. The graph combined with the limitations of cell specify the solution space for the metabolic network of all possible states that the network may accept.

Flux Balance Models require to define all metabolic reactions and metabolites used in the biological model. This can be defined using mathematical model. [1]

3.2.1 Matematical model

To create a mathematical model should take into account the rules based on the chemical equations of the biological model. These chemical equations can be transformed the system of differential equations using the law of mass action.

3.2.2 Mass Action

Mass action law is based on the principle of conservation of mass and is used to describe the dynamic behavior of reaction networks. Metabolic network

in the steady state may be described by equations of conservation of mass. These equations also describe the changes in concentration over time, which is the difference of the rate at which a metabolite is formed and the rate at which a metabolite is consumed.

For some chemical reaction:

$$A + B \rightleftharpoons_{k_2}^{k_1} C$$

(This chemical reaction is reversible - it's marked by the double arrow.

k_1, k_2 represents the reaction rate, where

k_1 is a reaction rate of reaction: $A + B \xrightarrow{k_1} C$

and k_2 is a reaction rate of reaction: $C \xrightarrow{k_2} A + B$)

Then according to the law of mass action this reaction can be converted to a mathematical equation:

$$\dot{c}(t) = a \cdot b \cdot k_1 - c \cdot k_2$$

So we get the differential equation describing the formation of the product C. In general we can write: $\sum_i s_i \cdot x_i \xrightarrow{k} c$

$$\dot{c} = k \cdot \prod_i x_i^{s_i}$$

It is obvious that the sum of all the variables in the equations is equal to the outer fluxes and thus satisfies the principle of conservation of mass, thus no mass is lost during the experiment.

In this point I have applied the mass action law on the individual reactions of glycolysis and I got the following differential equations:

1.) $\underbrace{Glucose}_{x_1} + \underbrace{ATP}_{x_2} \xrightarrow{k_1} \underbrace{ADP}_{x_3} + \underbrace{G6P}_{x_4}$

$$\dot{x}_4(t) = x_1 \cdot x_2 \cdot k_1$$

2.) $\underbrace{G6P}_{x_4} \rightleftharpoons_{k_3}^{k_2} \underbrace{F6P}_{x_5}$

$$\dot{x}_5(t) = x_4 \cdot k_2 - x_5 \cdot k_3$$
$$\dot{x}_4(t) = x_5 \cdot k_3 - x_4 \cdot k_2$$

3.) $\underbrace{F6P}_{x_5} + ATP \xrightarrow{k_4} ADP + \underbrace{F16BP}_{x_6}$

$$\dot{x}_6(t) = x_5 \cdot x_2 \cdot k_4$$

4.) $F16BP \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \underbrace{GA3P}_{x_7} + \underbrace{DHAP}_{x_8}$

$$\dot{x}_7(t) = x_6 \cdot k_5 - x_7 \cdot x_8 \cdot k_6 = \dot{x}_8(t)$$

5.) $DHAP \underset{k_8}{\overset{k_7}{\rightleftharpoons}} GA3P$

$$\dot{x}_7(t) = x_8 \cdot k_7 - x_7 \cdot k_8$$

6.) $GA3P + \underbrace{NAD^+}_{x_9} \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} \underbrace{NADH}_{x_{10}} + \underbrace{13BPG}_{x_{11}}$

$$\dot{x}_{11}(t) = x_7 \cdot x_9 \cdot k_9 - x_{10} \cdot x_{11} \cdot k_{10}$$

7.) $13BPG + ADP \underset{k_{12}}{\overset{k_{11}}{\rightleftharpoons}} ATP + \underbrace{3PG}_{x_{12}}$

$$\dot{x}_{12}(t) = x_{11} \cdot x_3 \cdot k_{11} - x_2 \cdot x_{12} \cdot k_{12}$$

8.) $3PG \underset{k_{14}}{\overset{k_{13}}{\rightleftharpoons}} \underbrace{2PG}_{x_{13}}$

$$\dot{x}_{13}(t) = x_{12} \cdot k_{13} - x_{13} \cdot k_{14}$$

9.) $2PG \underset{k_{16}}{\overset{k_{15}}{\rightleftharpoons}} \underbrace{PEP}_{x_{14}}$

$$\dot{x}_{14}(t) = x_{13} \cdot k_{15} - x_{14} \cdot k_{16}$$

10.) $PEP + ADP \xrightarrow{k_{17}} ATP + \underbrace{P}_{x_{15}}$

$$\dot{x}_{15}(t) = x_{14} \cdot x_3 \cdot k_{17}$$

3.2.3 Michaelis-Menten kinetics

Michaelis-Menten kinetics is a model of enzyme kinetics, which describes the rate of enzymatic reaction related to the concentration of substrate S.

For basic elementary chemical equation:

$$\underbrace{S}_{x_1} + \underbrace{E}_{x_3} \rightleftharpoons_{k_2}^{k_1} \underbrace{C}_{x_2} \xrightarrow{k_3} \underbrace{P}_{x_4} + \underbrace{E}_{x_3}$$

relate the differential equations:

1.) $\dot{x}_1 = k_2 \cdot x_2 - k_1 \cdot x_1 \cdot x_3$

2.) $\dot{x}_2 = -k_2 \cdot x_2 + k_1 \cdot x_1 \cdot x_3 - k_3 \cdot x_2$

3.) $\dot{x}_3 = k_2 \cdot x_2 - k_1 \cdot x_1 \cdot x_3 + k_3 \cdot x_2$

4.) $\dot{x}_4 = k_3 \cdot x_2$

The fourth equation expressting the creation of the product does not feedback into three first equations and so we can drain this equation and solve the set of equation without this one (later after solving the system, integrating the equation we get x_4).

The sum of the second and third equation gives: $\dot{x}_2 + \dot{x}_3 = 0$ Since equation x_3 is a linear combination of the equation x_2 and so we can also one of the following equations (e.g., x_3) drain from solution.

According to laws of mass conservation we set up:

$E + C = x_3 + x_2 = E_T \quad \Longrightarrow \quad x_3 = E_T - x_2$

$S + C + P = x_1 + x_2 + x_4 = S_T$

We get the reduced system in the form:

1.) $\dot{x}_1 = k_2 x_2 - k_1 x_1 x_3 = k_2 x_2 - k_1 x_1 (E_T - x_2)$

2.) $\dot{x}_2 = -k_2 x_2 + k_1 x_1 x_3 - k_3 x_2 = -k_2 x_2 - k_3 x_2 + k_1 x_1 (E_T - x_2)$

The following applies: $E_T \ll S_T \Longrightarrow \frac{E_T}{S_T} = \varepsilon \ll 1$

For the quasi-steady state we need equations in the form:

$$\dot{x}_1 = f_1(x)$$

$$\dot{x}_2 = \frac{1}{\varepsilon} f_2(x)$$

Because $\varepsilon \ll 1$, then $\frac{1}{\varepsilon}$ is a big number and it follows that x_2 change takes place very quickly. Fast dynamics must be stable around a quasi-steady state and therefore $\varepsilon > 0$.

We choose the auxiliary variables:

$\tilde{x}_1 = \frac{x_1}{S_T} \quad \Longrightarrow \quad \dot{\tilde{x}}_1 = f_1(\tilde{x}) \quad \Longrightarrow \quad x_1 = S_T \cdot \tilde{x}_1$

$\tilde{x}_2 = \frac{x_2}{E_T} \quad \Longrightarrow \quad \dot{\tilde{x}}_2 = f_2(\tilde{x}) \quad \Longrightarrow \quad x_2 = E_T \cdot \tilde{x}_2$

Following applies to ε: $\quad \tau = \varepsilon \cdot t \quad \Rightarrow \quad \frac{d\tau}{dt} = \varepsilon \ \frac{d\tilde{x}}{dt} = \varepsilon \cdot \frac{d\tilde{x}}{d\tau}$

Substituting into the differential equations:

1.) $\frac{d\tilde{x}_1}{dt} \cdot \varepsilon = \dot{\tilde{x}}_1 \varepsilon S_T = k_2 \tilde{x}_2 E_T - k_1 \tilde{x}_1 S_T (E_T - \tilde{x}_2 E_T)$

2.) $\frac{d\tilde{x}_2}{dt} \cdot \varepsilon = \dot{\tilde{x}}_2 \varepsilon E_T = -k_2 \tilde{x}_2 E_T - k_3 \tilde{x}_2 E_T + k_1 \tilde{x}_1 S_T (E_T - \tilde{x}_2 E_T)$

Factor out E_T from both equations and express $\dot{\tilde{x}}_1$ and $\dot{\tilde{x}}_2$:

1.) $\dot{\tilde{x}}_1 = k_2 \tilde{x}_2 - k_1 \tilde{x}_1 S_T (1 - \tilde{x}_2)$

2.) $\dot{\tilde{x}}_2 = \frac{1}{\varepsilon}[-k_2 \tilde{x}_2 - k_3 \tilde{x}_2 + k_1 \tilde{x}_1 S_T (1 - \tilde{x}_2)]$

According to Michaelis Menten kinetics sets the $\dot{\tilde{x}}_2 = 0$:

2.) $0 = -k_2 \tilde{x}_2 - k_3 \tilde{x}_2 + k_1 \tilde{x}_1 S_T - k_1 \tilde{x}_1 S_T \tilde{x}_2$

$$\tilde{x}_2 = \frac{k_1 \tilde{x}_1 S_T}{k_2 + k_3 + k_1 \tilde{x}_1 S_T}$$

1.) $\dot{\tilde{x}}_1 = \tilde{x}_2 (k_2 + k_1 \tilde{x}_1 S_T) - k_1 \tilde{x}_1 S_T$

$$\dot{\tilde{x}}_1 = \frac{k_1 \tilde{x}_1 S_T (k_2 + k_1 \tilde{x}_1 S_T)}{k_2 + k_3 + k_1 \tilde{x}_1 S_T}$$

Using modifications we obtain:

$$\dot{\tilde{x}}_1 = -\frac{k_3 \tilde{x}_1}{\frac{k_2 + k_3}{k_1 S_T} + \tilde{x}_1}$$

By back replacement for the $\tilde{x}_1 = \frac{x_1}{S_T}$, we get:

$$\dot{x}_1 = -\frac{k_3 E_T x_1}{\frac{k_2+k_3}{k_1} + x_1}$$

Specify auxiliary variables: $x_1 = S, \quad K_m = \frac{k_2+k_3}{k_1}, \quad V_m = k_3 E_T$

$\dot{S} = \frac{V_m S}{K_m+S} \quad \Rightarrow \quad S \xrightarrow{\frac{V_m}{K_m+S}} P$

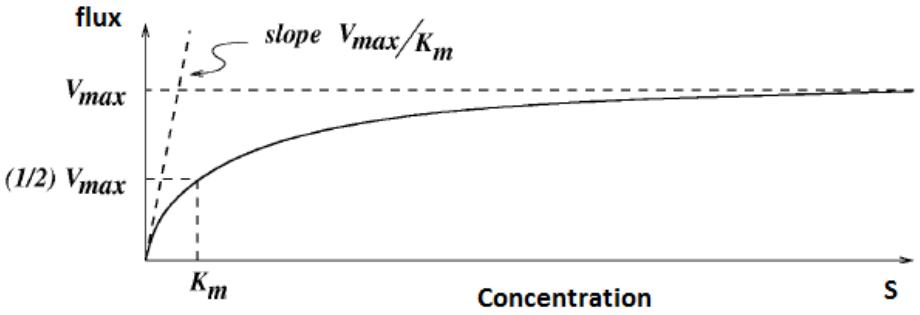


Figure 3.2: The dependence of the substrate concentration on the rate of product formation. At low concentrations, we get a linear growth.

[2]

By this we derive an expression for the rate of enzymatic reaction related to the concentration of substrate according to Michaelis-Menten kinetics.

Now need to write the state equations for the metabolic pathway:

$$U_1 \xrightarrow{1} X_1 \xrightarrow{\frac{V_1}{K_1+X_1}} X_2 \xrightarrow{3} X_3 \xrightarrow{4} X_4 \xrightarrow{5} X_5$$

$$X_1 + E_1 \rightleftharpoons C_1 \rightarrow E_1 + X_2$$

$$X_6 \xrightarrow{7} X_7 \xrightarrow{8} X_8 \xrightarrow{9} X_9 \xrightarrow{10} X_{10} \xrightarrow{11} X_{11}$$

Figure 3.3: Simplified metabolic system. U_1 represents the input (glucose).

3.2.4 The Steady state

The most important assumption of FBA is that we need to have a model in the steady-state, it means that model is in equilibrium. This assumption is important because the metabolic transitions are much faster than the rate of cell growth and, therefore, stabilize the state of the whole system will within a few seconds. On the other hand, changes in metabolism will be visible after a few minutes or even hours - this assumption is known as quasi-steady state.

To obtain a steady state model can be system of differential equations obtained in the Mass Action Law described stoichiometric matrix and vector Xr. The stoichiometric matrix S is a mxn matrix, where m corresponds the number of metabolites and n is the number of chemical reactions of the metabolic network. Matrix S represents the relationship between the metabolites and products, in that case how much amount of the metabolite is necessary to produce a particular product. Vector of fluxes Xr is a vector describing a rate at which the product is formed.

Stoichiometric matrix S with the vector Xr describing glycolysis:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \\ \dot{x}_8 \\ \dot{x}_9 \\ \dot{x}_{10} \\ \dot{x}_{11} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}}_{S} \cdot \underbrace{\begin{pmatrix} X_1 r_1 \\ X_2 r_2 \\ X_3 r_3 \\ X_4 r_4 \\ X_5 r_5 \\ X_6 r_6 \\ X_7 r_7 \\ X_8 r_8 \\ X_9 r_9 \\ X_{10} r_{10} \\ X_{11} r_{11} \end{pmatrix}}_{Xr}$$

This model can be transformed into a steady state where we assume that the change in concentration over time will be approximately zero. Thus we can write:

$$\dot{x} = 0 = S \cdot Xr$$

The metabolic pathway operates optimally and therefore logically requires the optimal model results and the optimal path through metabolic networks that optimize certain target. In our case, we want to optimize bacterial growth and maximum yield of energy (ATP). [2]

3.3 Optimalization

The basic problem of optimal control theory is the search for a sequence or a strategy of control that achieves the desired objectives while minimizing (or maximizing) defined system criteria. From this system criteria or optimality criteria also expect that allow comparison of different solutions to problems that are available and choose the best of them.

The problem of optimal control can be divided into four interrelated parts:

1.) Definition of the objective.

2.) Knowing the current state due to the target.

3.) Knowledge of environmental factors affecting the present and future.

4.) Determining the best strategy for generating control based on provided objective, current state and environment.

For the solving of the problem of optimal control is therefore necessary first set a target, which has to be optimized. This requires an appropriate definition of the problem in the real world and transform this description into mathematical terminology. In the same way, the real process converted into mathematical description. The term system will always be understood as an abstract system, which is a mathematical model. In our case, we have now two objectives: to maximize the output energy (ATP) and to optimize cell growth. [4]

3.3.1 ATP yield

Glycolytic network to transfer into the mathematical description, which will include all the ATP molecules that were consumed and produced during glycolysis. The resulting description is determined by the corresponding fluxes

of metabolic network:

$$Y = -Xr(2) - Xr(4) - Xr(6) + Xr(10) + Xr(13)$$

From equation easy to see that in reactions 2,4 and 6 occurred ATP consumption while in reactions 10 and 13 ATP was produced.

3.3.2 Biomass function

Our second objective is to optimize cellular growth. In order to able to realize cellular growth, it is necessary the presence of metabolites that are listed in the following table, with the required quantity.

Metabolite	**Demand(mmol)**
ATP (Adenosine triphosphate)	41.2570
$NADH$ (Nicotinamide adenine dinucleotide)	-3.5470
$NADPH$ (Nicotinamide adenine dinucleotide phosphate)	18.2250
$G6P$ (Glucose-6-phosphate)	0.2050
$F6P$ (Fructose-6-phosphate)	0.0709
$R5P$ (Ribose-5-phosphate)	0.8977
$E4P$ (Erythrose-4-phosphate)	0.3610
$GA3P$ (Glyceraldehyde-3-phosphate)	0.1290
$3PG$ (3-phosphoglycerate)	1.4960
PEP (Phosphoenolpyruvate)	0.5191
P (Pyruvate)	2.8328
$ACoA$ (Acetyl coenzyme A)	3.7478
OA (Oxaloacetate)	1.7867
AKG (α - ketoglutarate)	1.0789

Cellular growth we take as a single reaction, in which all of the above metabolites converted to one gram of biomass and $3.547 \cdot NADH$ (which has a negative value in the table and this means that it is created). To derive the mathematical description we only have one reaction, it means that we optimize the flux of this reaction and call it Biomass function.

However, to realize the cellular growth, metabolic networks need to include metabolites necessary for growth in the table. Metabolites of ATP, $NADH$, $G6P$, $F6P$, $GA3P$, $3PG$, PEP and P are already part of the metabolic network of glycolysis. For obtaining $NADPH$, $R5P$ and $E4P$ was connect the Pentose Phosphate Pathway and for obtain $ACoA$, OA and

AKG was added Citric Acid Cycle. For completeness metabolic networks was still added Entner-Doudoroff Pathway and due fulfillment of the task will add a second input glycerol. The whole modified metabolic network is shown in the figure 3.1. [3]

3.3.3 CVX

The optimization problem we solved in Matlab with a CVX superstructure. CVX is a modeling system for convex optimization, which allows the definition of objectives and constraints specified using the standard syntax in Matlab. Constraints and objectives are expressed by rules and are automatically converted to a canonical form and solved.

The structure of the optimization problem:

- variables: $x_1,\ x_2,\ \ldots\ x_n$
- constraints: $f_i(x) \leq 0,\ \ i = 1, ..., m$
- cost function: $J(x)$ - have to be convex
- minimize $J(x)$ - for all x satisfying the conditions

In the example there is the syntax of cvx optimization In Matlab. Includes defining of variables, which are in our case fluxes Xr. Follows the limiting condition for the varialbes and also a condition for cost function. The last point is maximalization of the last reaction flux.

```
cvx_begin
    cvx_quiet(true)
    variables Xr(nrea)
    cvx_solver sdpt3

    Xr >= 0
    for i = 1:ninp
        Si(i,:)*Xr == inp{i}{2}
    end
    S*Xr == 0
    maximize(Xr(end))
```

```
cvx_end
```

4 Simulations and Experiments

For experimental verify the model, it is necessary to perform a series of simulations.

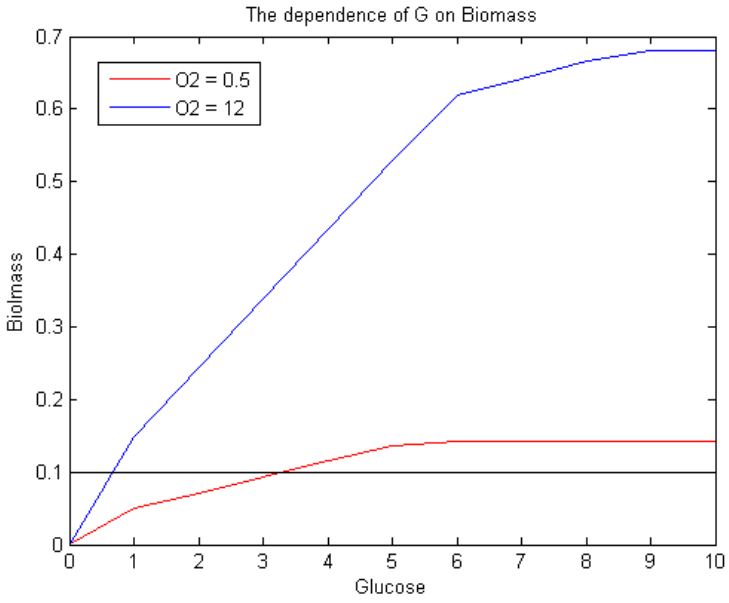
4.1 Simulation

Simulation is a process of formation a model of real system and implementing experiments with this model for a better understanding of studied system or assess different variants of system activity.

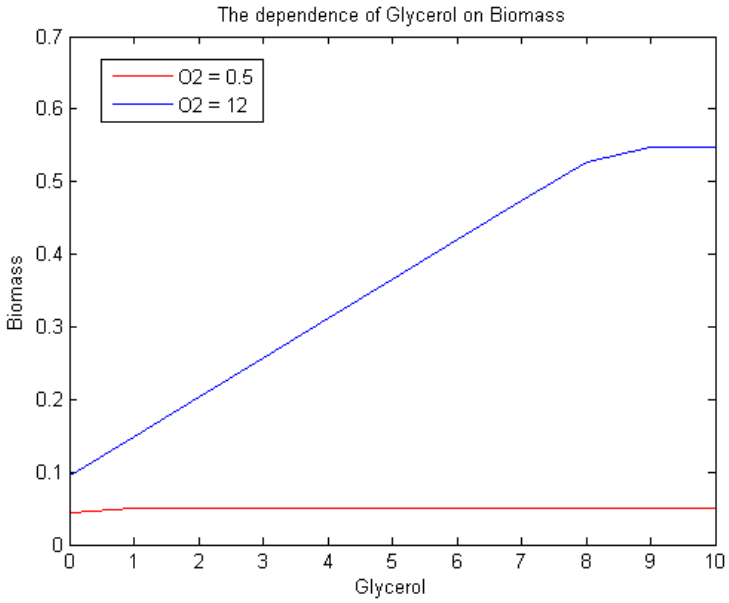
Different changes of input parameters, we followed the behavior of the model and the dependencies between parameters. We came to the following predictions:

4.1.1 Results of Simulations

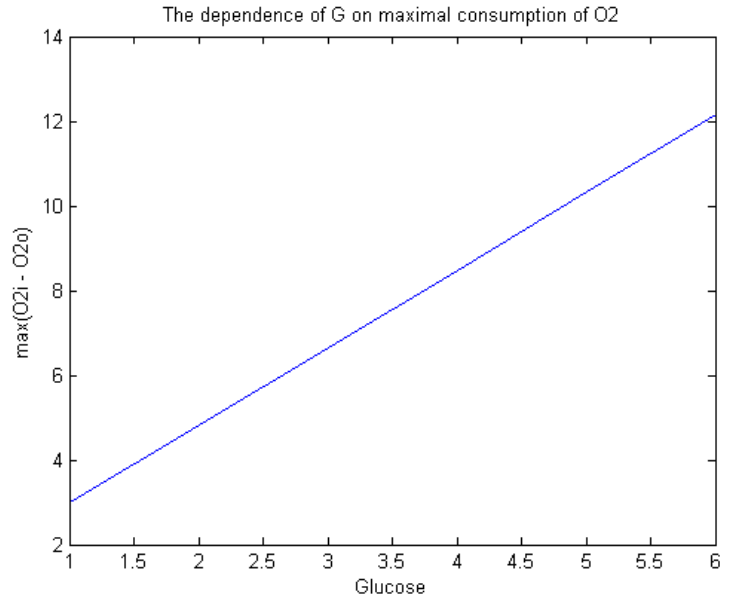
1. **The dependence of glucose on cellular growth** - we wanted to know, how will the model behave in aerobic and anaerobic conditions. The following figure shows the dependence of glucose in cellular growth for maximal and minimal amount of oxygen. We can easy see, that the growth is slower and sooner stable for small amount of oxygen than for large amount of oxygen. From this we concluded prediction, that for the same growth is necessary different amount of glucose in aerobic and anaerobic conditions.



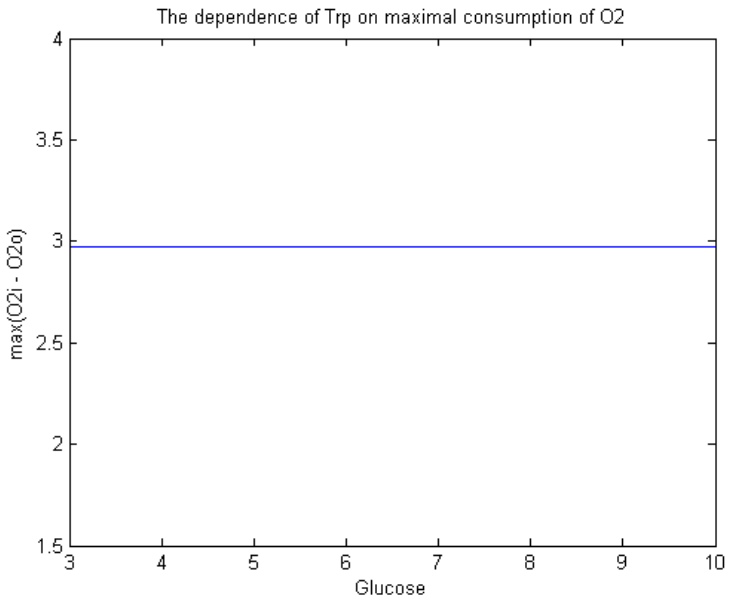
2. **The dependence of glycerol on cellular growth** - we tried there the same prediction such as in previous case. On the following figure we can see, that the dependence end up similarly thus slower growth for smaller concentration of oxygen and faster growth for larger concentration of oxygen. But we can't to say here, that for the same growth we need different amount of glycerol. It is caused by imperfections in the model, in which is missing gluconeogenesis. Gluconeogenesis is a metabolic pathway, which produce glucose e.g. from pyruvate, lactate or glycerol when amount of glucose is insufficient for growth.



3. **The dependence of glucose on maximal consumption of oxygen** - On the figure below is shown, what is the maximum consumption of oxygen for a given amount of glucose. From the figure is easy to see, that this dependence is linear. This means, that the bigger concentration of glucose, the higher oxygen consumption.



4. **The dependence of tryptophan on maximal consumption of oxygen** - Tryptophan is amino acid, which is involved on the regeneration of NAD^+. On the figure we can see, that this dependence is constant. This means, that addition of any amount of tryptophan does not affect on consumption of oxygen.



4.2 Experiments

The task of the experiment is to verify or disprove the findings about model, which come from simulations.

4.2.1 Draft of experiment

Experiment can be designed on the basis of predictions, which are obtained from simulations. The first prediction was, that for the same rate of growth in aerobic and anaerobic conditions is necessity different concentration of

glucose. The experiment was designed precisely on the basis of different concentrations of glucose and on aerobic and anaerobic conditions.

4.2.2 Protocol

It was necessary to guarantee different concentrations of glucose. This was achieved via growth medium YPD, when instead of standard 2% of glucose was added 4% glucose. Then through dilution series by using YP (medium without glucose - 0%) was obtained concentrations 4%, 2%, 1%, 0.5% and 0%. To these concentrations was added YPG, growth medium with glycerol.

Available was 12x8-well plate (12 rows and 8 columns), which is divided into two parts, one for aerobic condition and second for anaerobic condition. Filling of plate can be seen in the following table:

	A B C D E F	G H
1	$4\%YPD + 10\mu l$ cells	$4\%YPD$
2	$2\%YPD + 10\mu l$ cells	$2\%YPD$
3	$1\%YPD + 10\mu l$ cells	$1\%YPD$
4	$0.5\%YPD + 10\mu l$ cells	$0.5\%YPD$
5	$0\%YP + 10\mu l$ cells	$0\%YP$
6	$YPG + 10\mu l$ cells	YPG
7	$4\%YPD + 10\mu l$ cells + oil	$4\%YPD$ +oil
8	$2\%YPD + 10\mu l$ cells + oil	$2\%YPD$ + oil
9	$1\%YPD + 10\mu l$ cells + oil	$1\%YPD$ + oil
10	$0.5\%YPD + 10\mu l$ cells + oil	$0.5\%YPD$ + oil
11	$0\%YP + 10\mu l$ cells + oil	$0\%YP$ + oil
12	$YPG + 10\mu l$ cells + oil	YPG + oil

Each well was filled with 100 ml media and added 10ml cells, whereas last two columns G and H was without cells. First six rows represents aerobic conditions and second six rows represents anaerobic conditions, which are seal with mineral oil.

The plate was measured kinetically at the required temperature 30°C and by absorbance at 600nm, at which is measured cellular growth.

4.2.3 Results of Experiments

This experiment was repeated several times. Here are the results from the two measurements.

First measurement

Measurement process took 3 hours and growth was measured at 10 minute intervals. The measurement can watch on two following graphs.

On the first graph we can see process of measurement in aerobic condition. It is obvious, that the rate of growth is the biggest for 4% concentration of glucose.

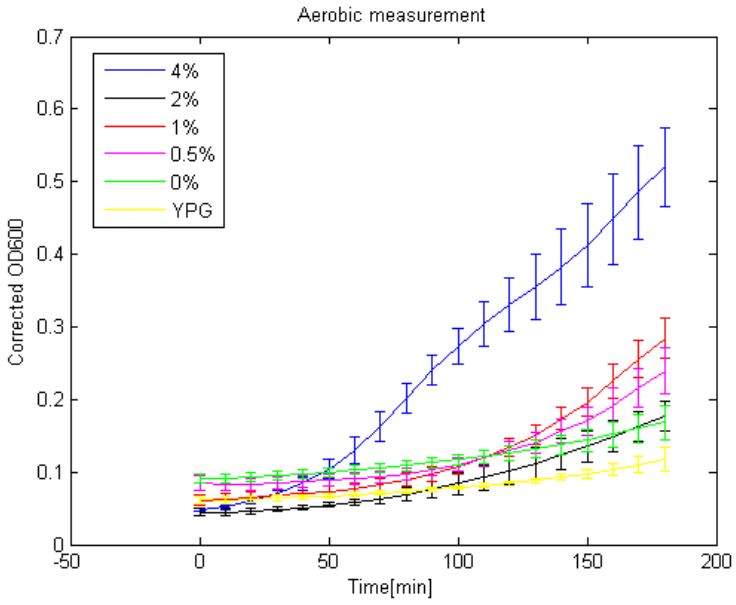


Figure 4.1: The graph for measurement in aerobic conditions. Individual concentrations are color coded.The largest growth is noticeable for a concentration of 4%

If each curve is approximated in his last part using the straight line, according to the slope can determine the rate of growth of different concentrations.

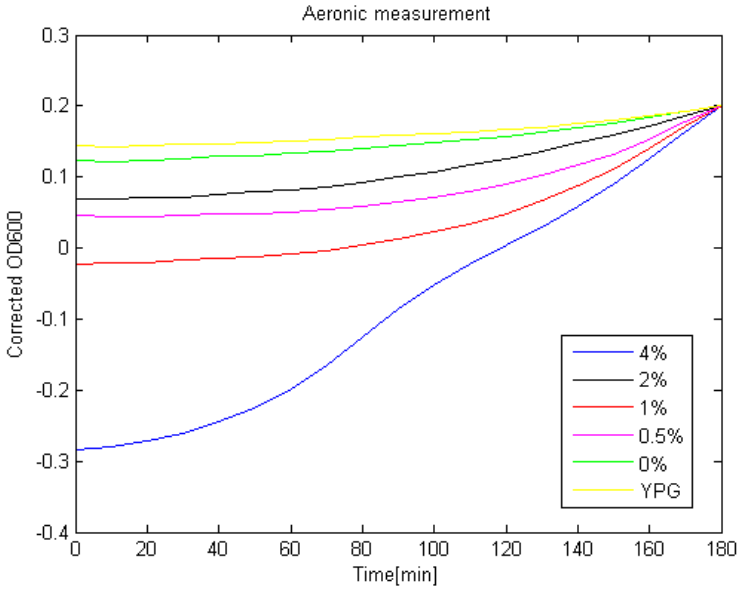


Figure 4.2: The graph for measurement in aerobic conditions. Individual concentrations was approximated using the straight line.

Released the following order:

1	2	3	4	5	6
4%	1%	0.5%	2%	0%	YPG

This order doesn't work just for 2% concentration of glucose . The error was probably caused by the preparation of 2% medium. The rest of the order went as expected.

On the second graph is shown measurement in anaerobic condition.

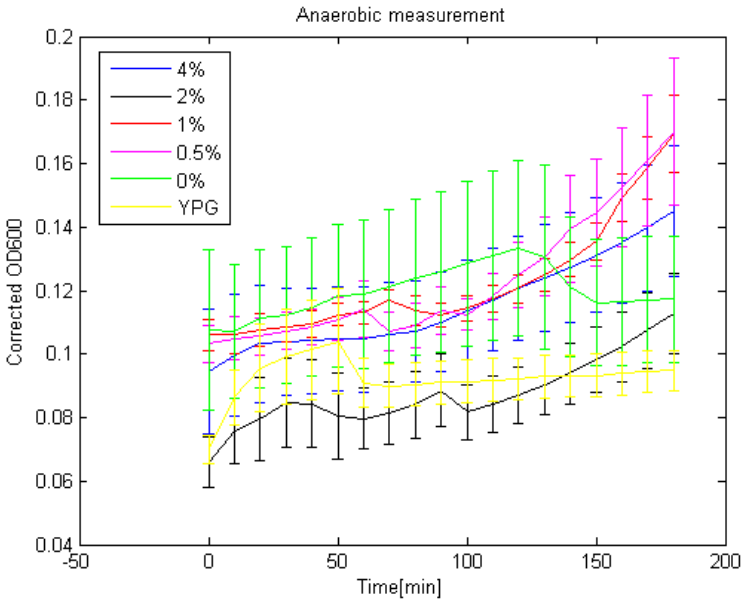


Figure 4.3: The graph for measurement in anaerobic conditions. Individual concentrations are color coded.

The fluctuations were probably caused by the rapid growth of cells and the rapid depletion of oxygen. Therefore, part of the cells began to die and then the rest of oxygen was sufficient for the rest of the cells, which started to grow again. Approximation we get the following chart:

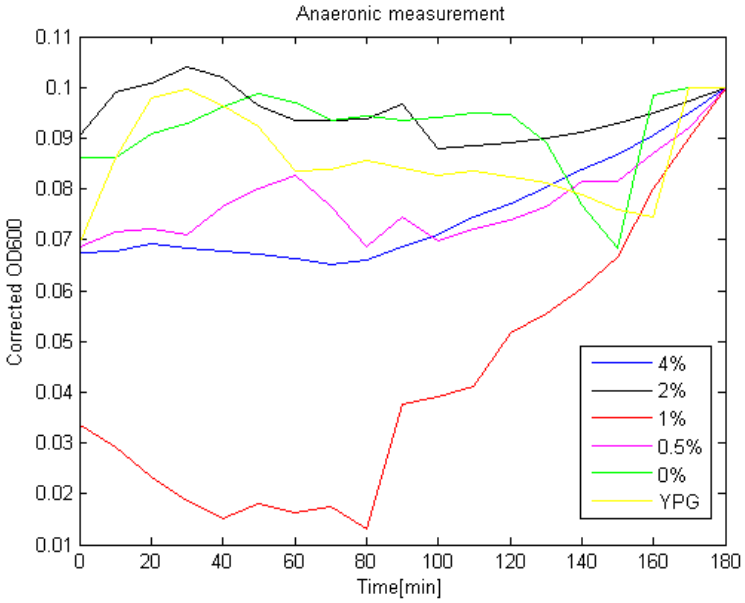


Figure 4.4: The graph for measurement in anaerobic conditions and approximated with straight line.

In this case, the growth rate following:

1	2	3	4
1%	0.5%	4% and 2%	0% and YPG

This order is not surprising. 4% and 2% concentration proved to be too high for such a small amount of oxygen and become toxic. While the 1% concentration proved to be ideal for anaerobic conditions.

4.3 Comparison

The object of this experiment was to found for the same rate of growth in aerobic and anaerobic conditions different concentration of glucose. More precisely lower concentration for aerobic and higher concentration for anaerobic condition. In one case this prediction was proved.

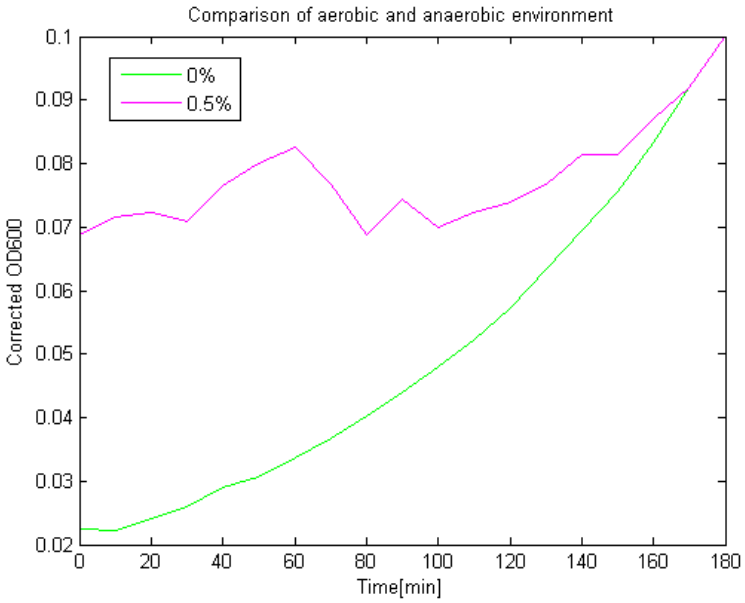


Figure 4.5: The comparison of 0.5% concentration in anaerobic condition with 0% concentration in aerobic condition.

On the Figure we can see the comparison of 0.5% concentration in anaerobic condition with 0% concentration in aerobic condition. We suppose, that adding of cells with 2% medium into 0% YP to leave little concentration of glucose for growth.

Bibliography

[1] J. Tanis, *Metabolic Engineering, Flux Balance Analysis*, July 2006

[2] E. D. Sontag, *Lecture Notes on Mathematical Systems Biology*, Rutgest Univeryity 2010

[3] B. Palson, *Flux-balance analysis: Basic concept*, Lecture 6, September 2003

[4] M. Flídr, *Skripta OPS*

[5] D. White, *The Physiology and Biochemistry of Prokaryotes*, Oxford 2012

[6] M. Ledvina, *Biochemistry for students of medicine 1*, 2005